Statistics for Industry and Technology

N. Balakrishnan
Erhard Cramer

The Art of Progressive Censoring

Applications to Reliability and Quality

Birkhäuser

Birkhäuser

Statistics for Industry and Technology

For further volumes:
http://www.springer.com/series/4982

N. Balakrishnan • Erhard Cramer

The Art of Progressive Censoring

Applications to Reliability and Quality

Birkhäuser

N. Balakrishnan
Department of Mathematics and Statistics
McMaster University
Hamilton, ON
Canada

Erhard Cramer
Institute of Statistics
RWTH Aachen University
Aachen, Germany

ISBN 978-0-8176-4806-0 ISBN 978-0-8176-4807-7 (eBook)
DOI 10.1007/978-0-8176-4807-7
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014939954

Mathematics Subject Classification (2010): 62N86, 62N05, 62N01, 62P30

Printed on acid-free paper

Springer Basel is part of Springer Science+Business Media (www.birkhauser-science.com)

To Sarah and Julia Balakrishnan

N.B.

Für meine Eltern

E.C.

Preface

Practitioners and statisticians are often faced with incomplete or censored data. In life testing, censored samples are present whenever the experimenter does not observe the failure times of all units placed on the life test. This may happen intentionally and unintentionally and may be caused, e.g., by time constraints on the test duration like in Type-I censoring, by requirements on the minimum number of observed failures, or by the structure of a technical system. Naturally, the probabilistic structure of the resulting incomplete data depends heavily on the censoring mechanism and so suitable inferential procedures become necessary.

Progressive censoring can be described as a censoring method where units under test are removed from the life test at some prefixed or random inspection times. It allows for both failure and time censoring. Many modifications of the standard model have been developed, but the basic idea can be easily described by progressive Type-II censoring which can also be considered as the most popular model. Under this scheme of censoring, from a total of n units placed simultaneously on a life test, only m are completely observed until failure. Then, given a censoring plan $\mathscr{R} = (R_1, \ldots, R_m)$:

- At the time $x_{1:m:n}$ of the first failure, R_1 of the $n-1$ surviving units are randomly withdrawn (or censored) from the life-testing experiment.
- At the time $x_{2:m:n}$ of the next failure, R_2 of the $n-2-R_1$ surviving units are censored, and so on.
- Finally, at the time $x_{m:m:n}$ of the mth failure, all the remaining $R_m = n - m - R_1 - \cdots - R_{m-1}$ surviving units are censored.

Note that censoring takes place here progressively in m stages. This scenario is illustrated in Fig. 1 which may be one of the most reproduced figures in the literature on progressive censoring. Clearly, this scheme includes as special cases the complete sample situation and the conventional Type-II right censoring scenario. The ordered failure times $X^{\mathscr{R}}_{1:m:n} \le \cdots \le X^{\mathscr{R}}_{m:m:n}$ arising from such a progressively

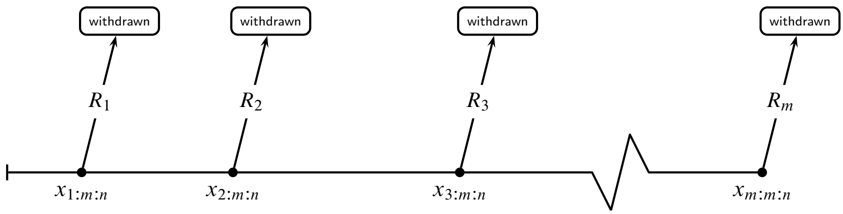

Figure 1 Illustration of progressive Type-II censoring.

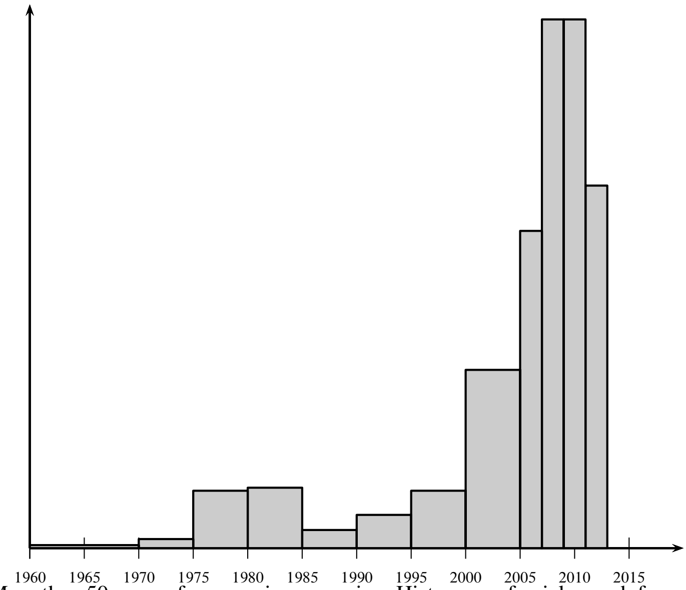

Figure 2 More than 50 years of progressive censoring: Histogram of quick search for publications (zbMATH) from 1961–2013.

Type-II right censored sample are called progressively Type-II censored order statistics. These are natural generalizations of the usual order statistics that have been studied quite extensively during the past century.

Progressive censoring has been termed a *relatively unexplored idea* in the monograph of Balakrishnan and Aggarwala [86] that appeared in 2000. Based on the outcome of a quick search in the zbMATH[1] database using the keyword `progressive* censor*`, we constructed a histogram on the number of published papers presented in Fig. 2.

It readily reveals that this statement was indeed true in the year 2000. But, the histogram also shows that the number of publications since this time has grown very fast and that it is still growing with about 200 papers in the last five years.

[1] http://zbmath.org.

Moreover, the topics addressed have by now become quite diverse. They range from distribution theory and various approaches in inference to modifications of the model, etc. All these new developments pertaining to progressive censoring, since the publication of the previous book by Balakrishnan and Aggarwala [86], have been carefully and systematically analyzed in the present book. Thus, it provides an up-to-date account on the state of the art on progressive censoring!

As mentioned above, the research on progressive censoring has grown fast in the recent past and great progress has been made with regard to distribution theory as well as in developing inferential procedures. In this book, we review the relevant literature and present a comprehensive, detailed, and unified account of the material. Due to the burgeoning literature on progressive censoring and many different developments that have taken place, we have presented here a detailed coverage of the following key topics which also reflect the structure of the book:

- **Distribution Theory and Models**
  After introducing the basic notion and models of progressive censoring, we present a comprehensive treatment of distributional properties of progressively censored order statistics. Even though the major part is devoted to progressive Type-II censoring, we give details on various other models like progressive Type-I censoring, progressive hybrid censoring, adaptive progressive censoring, and progressive censoring for nonidentical distributions and dependent variates. The material not only includes general results on joint, marginal, and conditional distributions and the dependence structure of the failure times, but also focuses on life distributions that are most important in applications (e.g., exponential and Weibull distributions). Further topics are moments, recurrence relations, characterizations, stochastic ordering, extreme value theory, simulation, and information measures like Fisher information and Shannon entropy.
- **Inference**
  The inferential topics cover linear, likelihood, and Bayesian inference in various models of progressive censoring. We discuss point and interval estimation for many life distributions as well as prediction problems. The discussion is completed by nonparametric inferential approaches and statistical tests including goodness-of-fit and precedence-type tests.
- **Applications in Survival Analysis and Reliability**
  Finally, applications in survival analysis and reliability are provided. The presentation ranges from acceptance sampling, accelerated life testing including step-stress testing, stress-strength models, and competing risks to optimal experimental design. These ideas also provide testimony to the usefulness and efficiency of progressive Type-II censoring as compared to conventional Type-II censoring.

The book provides an elaborate discussion on progressive censoring, with a special emphasis on Type-II right censoring. Even though we provide proofs of the results in most cases, it was not possible to include each detail especially when the derivations become quite technical. Further, we illustrate the methods and procedures by several plots and diagrams just to reveal the censoring mechanism and the differences between the models. All the inferential results are illustrated with several numerical examples and many tables are also provided. In this regard, it needs to be mentioned here that many of the tables of best linear unbiased estimators and optimal progressive censoring schemes that are in the book of Balakrishnan and Aggarwala [86] will still continue to be useful for practitioners. For illustrative purposes and the support of future research, we have included a great number of progressively censored data sets which have been analyzed in the literature and used to illustrate the inferential methods. Many generalizations to some other related censoring schemes like generalized order statistics and sequential order statistics and their applications are also highlighted. An extensive up-to-date bibliography on progressive censoring has also been included which reflects the current state of research. All these aspects of this book will make it a valuable resource for researchers and graduate students interested in the area of life testing and reliability and also serve as an important reference guide for reliability practitioners.

We have written this book with the sincere hope that more practitioners will recognize the versatility of progressive censoring and be tempted to employ in their work this type of censoring/sampling scheme and the methodologies based on it. We also hope that the mathematical ideas and results presented in this book will motivate the aspiring researchers (among the statistical and the engineering communities) to explore further into the theoretical aspects of progressive censoring.

The book has been written in a self-contained manner and, therefore, will be quite suitable either as a text for a graduate topic course, a text for a directed-reading course, or as a handbook on progressive censoring. Though a one-year mathematical statistics course at the undergraduate level will provide an adequate background to go over the introductory chapters of this book, a basic exposition to order statistics (such as the one based on the book

A First Course in Order Statistics

by Arnold, Balakrishnan, and Nagaraja [58]) will make the journey through this book a lot more pleasant! In order to show different roadmaps through the book, we have added two flow charts in Figs. 3 and 4. They illustrate several possibilities to go through the material. However, you may leave the path at any crossing, just to explore! Some advanced topics included in the book require a deeper knowledge of mathematics and statistics and may be mainly of interest to researchers and advanced practitioners. In this direction, the book serves as a comprehensive compendium on progressive censoring providing the background for research and applications of progressive censoring.

We express our sincere thanks to Allen Mann, Mitch Moulton, and Kristin Purdy (all of Birkhäuser, Boston) for their enthusiasm and keen interest in this project from the very beginning. Our thanks also go to Dharmaraj Raja (Project Manager at SPi Technologies India Pvt. Ltd) for helping us with the final production of the volume. Our final appreciation goes to the Natural Sciences and Engineering Research Council of Canada for providing research grants which certainly facilitated our many meetings during the course of this project, thus enabling the work to progress

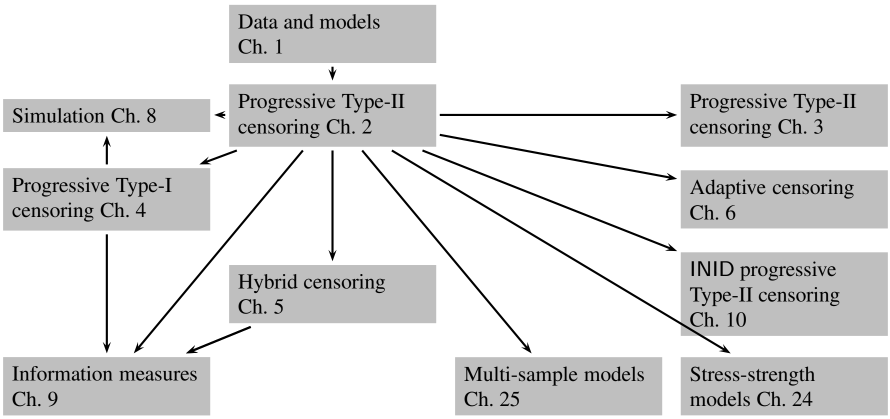

Figure 3 Journey through the book: the probability and model path.

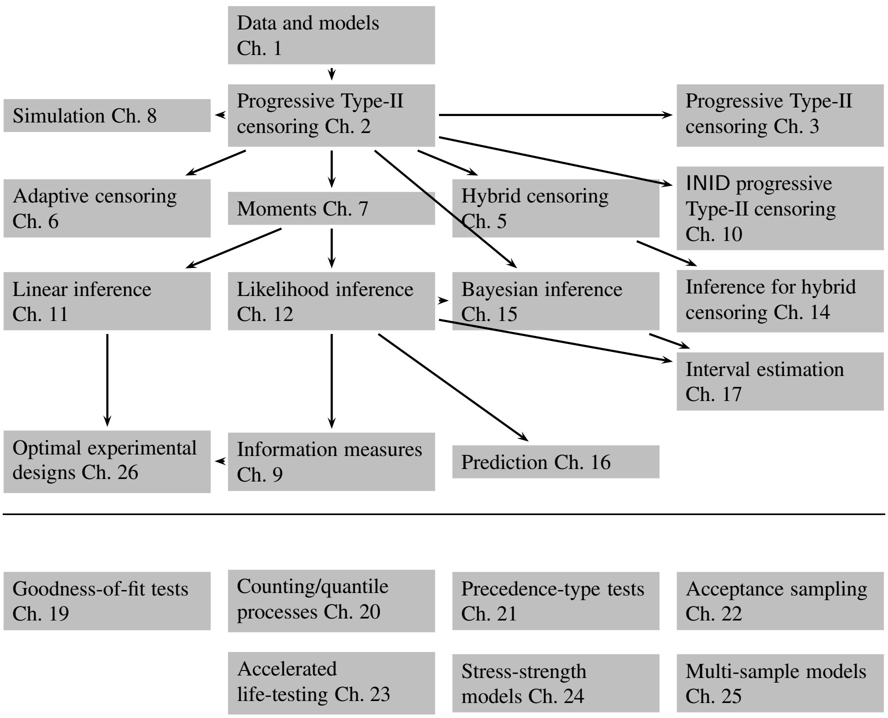

Figure 4 Journey through the book: The Type-II path. The chapters in the lower part of the chart may be chosen as a continuation.

smoothly. This book has been a true labor of love for us and it is our sincere hope and wish that it will serve as a valuable guide for researchers in the years to come and stimulate much more research activity in this interesting and useful area of research!

Section and Equation Numbering and Referencing

Throughout this book, chapters, sections, and subsections are labeled consecutively by Arabic numbers. Theorems, definitions, remarks, examples, etc., are jointly labeled as chapter no.section no.X, where X is restarted by a new section (for instance, Theorem 1.2.3). Equations are referenced by (chapter no. X) where the counter X is restarted by a new chapter (for example, (3.5) refers to the fifth numbered equation in Chap. 3).

Further, theorems, definitions, remarks, examples, etc., are set in sans serif font. The end of a proof is marked by □.

References are organized in the bibliography in alphabetical order by the first author and consecutively numbered. They are referred to in an author–number style like Cramer and Balakrishnan [292]. For three and more authors, the reference has the form Balakrishnan et al. [129].

Hamilton, ON, Canada N. Balakrishnan
Aachen, Germany Erhard Cramer
February 2014

Contents

Part I
Distribution Theory and Models

Chapter 1
Progressive Censoring: Data and Models

Given numbers $(R_1, \ldots, R_r)$ (the censoring scheme or censoring plan) and n units put simultaneously on a life test, the general idea of progressive censoring is to withdraw some units from the test within the test duration. As depicted in Fig. 1.1, units on test are removed from the experiment at times $\tau_1 < \cdots < \tau_r$ provided that enough units are available at each particular time point τ_j.

Although various models of progressively censored data have been discussed in the literature,[1] most of these models can be traced back to either

(1) progressive Type-II censoring, or
(2) progressive Type-I censoring,

with possibly some variations on them. In progressive Type-II censoring, the intervention times τ_j correspond to observed failure times in the sense that τ_j is the next failure time after withdrawing some units from the experiment. Due to the design of the experiment, the intervention times $\tau_1 < \cdots < \tau_r$ are random, but both the number of observations and the number of removals at τ_j are fixed. This mechanism is a generalization of Type-II right censoring (cf. Arnold et al. [58], David and Nagaraja [327]). The procedure is introduced in detail in Sect. 1.1.1.

Progressive Type-I right censoring is based on prefixed time points $\tau_1 < \cdots < \tau_r$. Failure times are successively observed until the final time point τ_r at which the experiment is terminated. Therefore, the intervention times $\tau_1 < \cdots < \tau_r$ are all fixed, but the sample size as well as the effectively employed censoring plan is random (and may differ from the originally planned censoring scheme due to the unavailability of enough surviving units to carry out the required censoring at some stage). Furthermore, the test duration is bounded by τ_r, while it is random

[1] As pointed out in Balakrishnan and Aggarwala [86, p. 2], the term progressive censoring has also been used as an alternate term for sequential testing (see, e.g., Chatterjee and Sen [246], Majumdar and Sen [631], and Sinha and Sen [805, 806]). Furthermore, the term multi-censored sample is used (see, e.g., Herd [440]).

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_1,

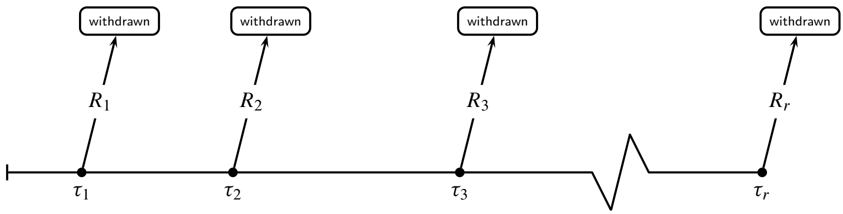

Fig. 1.1 Progressive censoring of a life test with intervention times $\tau_1 < \cdots < \tau_r$ and censoring scheme $(R_1, \ldots, R_r)$

in the case of progressive Type-II censoring. It can be seen as a generalization of Type-I right censoring used quite commonly, e.g., in survival analysis (cf. Klein and Moeschberger [536]). The construction process of progressive Type-I censoring is presented in Sect. 1.1.2.

In this chapter, we introduce the basic ideas of the censoring procedures sketched above and illustrate the generation processes. Furthermore, we provide some sample data which are used throughout this book for illustrative purposes. As an example of a variation of these models, we present two kinds of

(3) progressive hybrid censoring

in Sect. 1.1.3. Further variations of progressively censored data are introduced in the following chapters.

1.1 Progressively Censored Data

The following notations are used throughout this book.

Notation 1.1.1. In progressive censoring, the following notations are used:

(1) $n, m, R_1, R_2, \ldots \in \mathbb{N}_0$ are all integers;
(2) m is the sample size (which may be random in some models);
(3) n is the total number of units in the experiment;
(4) R_j is the number of (effectively employed) removals at the jth censoring time;
(5) $\mathscr{R} = (R_1, \ldots, R_r)$ denotes the censoring scheme, where r denotes the number of censoring times.

It is worth mentioning that the following procedures are based on the random variables (or their realizations) only. Distributional assumptions are not involved in the construction process of progressively censored order statistics. Therefore, we introduce the basic models from a data perspective. The procedures are illustrated by the following data set often used in the progressive censoring framework.

Data 1.1.2 (Nelson [676], p. 105). The following 19 measurements are failure times (in minutes) for an insulating fluid between two electrodes subject to a voltage of 34 kV:

0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67	4.85	6.50
7.35	8.01	8.27	12.06	31.75	32.52	33.91	36.71	72.89	

1.1.1 Progressive Type-II Censoring

In the progressive Type-II censoring approach, the removals are carried out at observed failure times. The prefixed number of units is immediately withdrawn from the surviving units upon observing a failure. Therefore, the number of observations is fixed in advance, while the duration of the experiment is random. The generation of progressively Type-II censored order statistics can be carried out according to the following procedure.

Procedure 1.1.3 (Generation of progressively Type-II censored order statistics). Let $(\Omega, \mathfrak{A}, P)$ be a probability space and $X_1, \ldots, X_n$ be random variables on $(\Omega, \mathfrak{A}, P)$. Let $\mathscr{R} = (R_1, \ldots, R_m)$ be a censoring scheme.

For $\omega \in \Omega$, the progressively Type-II censored sample

$$X_{1:m:n}^{\mathscr{R}}(\omega), \ldots, X_{m:m:n}^{\mathscr{R}}(\omega),$$

based on $X_1(\omega), \ldots, X_n(\omega)$, is generated as follows:

① Calculate the order statistics $X_{1:n}(\omega) \le \cdots \le X_{n:n}(\omega)$;
② Let $\mathcal{N}_1 = \{1, \ldots, n\}$, $i = 1$;
③ Let $k_i = \min \mathcal{N}_i$ and put $X_{i:m:n}^{\mathscr{R}}(\omega) = X_{k_i:n}(\omega)$;
④ Choose randomly a without-replacement sample $\mathcal{R}_i \subseteq \mathcal{N}_i \setminus \{k_i\}$ with $|\mathcal{R}_i| = R_i$;
⑤ If $i < m$, set $\mathcal{N}_{i+1} = \mathcal{N}_i \setminus (\{k_i\} \cup \mathcal{R}_i)$ and go to ③, or else stop.

Thus,

$$X_{1:m:n}^{\mathscr{R}}(\omega), \ldots, X_{m:m:n}^{\mathscr{R}}(\omega) = (X_{k_1:n}(\omega), \ldots, X_{k_m:n}(\omega)).$$

Figure 1.2 depicts the generation procedure of progressively Type-II censored order statistics.

Given n, m and a censoring scheme $\mathscr{R}$, we define the set of admissible (Type-II) censoring schemes as

$$\mathscr{C}_{m,n}^{m} = \Big\{(r_1, \ldots, r_m) \in \mathbb{N}_0^m : \sum_{i=1}^{m} r_i = n - m\Big\}. \tag{1.1}$$

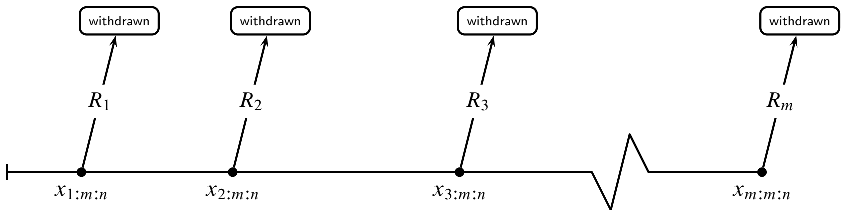

Fig. 1.2 Generation process of progressively Type-II censored order statistics

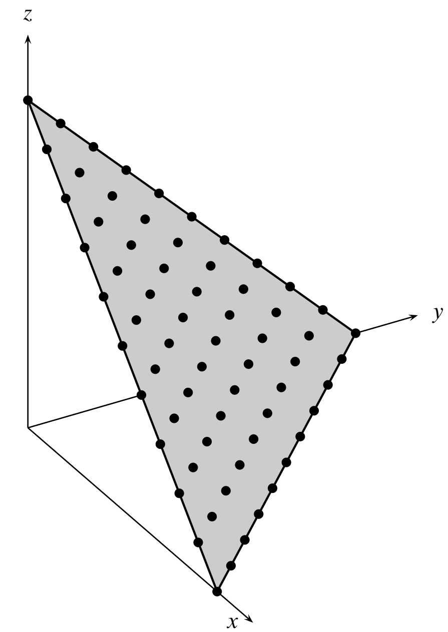

Fig. 1.3 Set $\mathscr{C}^3_{3,13}$ of admissible censoring schemes and its convex hull

The set $\mathscr{C}^3_{3,13}$ is displayed in Fig. 1.3. The admissible censoring schemes are represented by a dot. They are located in the shaded area which is the convex hull of $\mathscr{C}^3_{3,13}$.

For convenience, we introduce the notation 0^{*k} for k successive zeros. Thus, the scheme $(0, 0, 3, 0, 3, 0, 0, 5)$ is written in short as $(0^{*2}, 3, 0, 3, 0^{*2}, 5)$, or generally

$$(a_1, 0^{*n_1}, a_2, a_3, 0^{*n_2}, a_4, 0^{*n_3}) = (a_1, \underbrace{0, \ldots, 0}_{n_1 \text{ times}}, a_2, a_3, \underbrace{0, \ldots, 0}_{n_2 \text{ times}}, a_4, \underbrace{0, \ldots, 0}_{n_3 \text{ times}}).$$

On this note, (1^{*m}) means the censoring scheme $(1, \ldots, 1) \in \mathscr{C}^m_{m,n}$ etc. Table 1.1 presents some particular schemes which will be important in the subsequent analysis.

Scheme $\mathscr{R} = (R_1, \ldots, R_m)$	Meaning
$\mathscr{O}_m = (0^{*m-1}, n-m)$	Right censoring, i.e., first m order statistics in a sample of size n
(0^{*m})	Complete (ordered) sample of size $m = n$
$\mathscr{O}_1 = (n-m, 0^{*m-1})$	First-step censoring plan (FSP), i.e., removal takes place just after the first failure
$\mathscr{O}_k = (0^{*k-1}, n-m, 0^{*m-k})$	One-step censoring plan (OSP), i.e., removal takes place just after the kth failure, $2 \le k \le m-1$

Table 1.1 Particular censoring schemes

For one-step censoring plans, progressive censoring is carried out only at one failure time so that the censoring procedure is rather simple in this case. For instance, the kth one-step censoring plan is defined as

$$R_i = \begin{cases} n-m, & i = k, \\ 0, & \text{otherwise.} \end{cases}$$

It is denoted by $\mathscr{O}_k = (0^{*k-1}, n-m, 0^{*m-k})$, $k = 1, \ldots, m$. These particular schemes will especially be of interest in the area of experimental design (see Chap. 26). Up to this point, we note only that these censoring schemes are the vertices of the convex hull of admissible schemes $\mathscr{C}^m_{m,n}$ which forms a simplex. Of course, the vertices are also admissible schemes. Moreover, these particular censoring plans yield both simple distributions and a quite simple probabilistic structure in the IID case which will be very useful later on.

For progressive Type-II censoring, we introduce the numbers

$$\gamma_k = \sum_{j=k}^{m} (R_j + 1), \tag{1.2}$$

which represent the number of surviving objects before the kth failure, $k = 1, \ldots, m$. These numbers will be very useful in this area because there is a one-to-one (linear) relationship to the censoring scheme $\mathscr{R}$: $R_j = \gamma_j - \gamma_{j+1} - 1$, $j = 1, \ldots, m-1$, and $R_m = \gamma_m - 1$. In order to emphasize this connection, we use the notation $\gamma_j(\mathscr{R})$ if the dependence on the censoring plan is important.

In particular, it is easy to see that, for $\mathscr{R} \in \mathscr{C}^m_{m,n}$,

$$n = \gamma_1 > \gamma_2 > \cdots > \gamma_m \ge 1. \tag{1.3}$$

This illustrates that we have $m-1$ free parameters. Moreover, the set of admissible $(\gamma_2, \ldots, \gamma_m)$ is given by

$$\mathscr{G}_{m,n} = \{(\gamma_2, \ldots, \gamma_m) \in \mathbb{N}^{m-1} : n-1 \ge \gamma_2 > \cdots > \gamma_m \ge 1\}. \tag{1.4}$$

The set $\mathscr{G}_{3,13}$ is illustrated in Fig. 1.4.

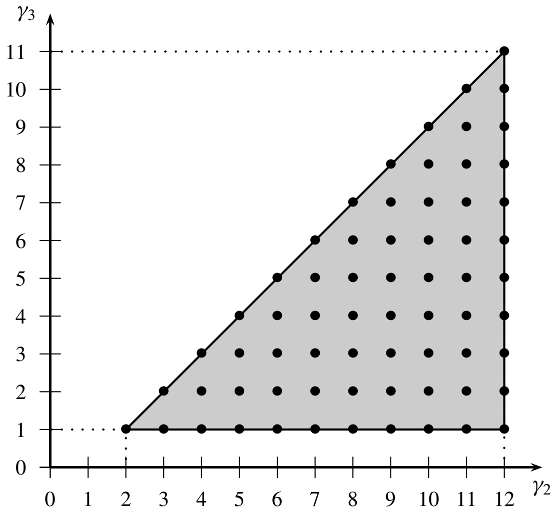

Fig. 1.4 Set $\mathscr{G}_{3,13}$ of admissible (γ_2, γ_3) and its convex hull ($\gamma_1 = 13$)

$\mathscr{R} \in \mathscr{C}^m_{m,n}$	$\gamma_1, \ldots, \gamma_m$
$(0, \ldots, 0) = (0^{*m})$	$\gamma_j = m - j + 1$
$(R, \ldots, R) = (R^{*m}), R \in \mathbb{N}_0$	$\gamma_j = (m - j + 1)(R + 1)$
$\mathscr{O}_m = (0^{*m-1}, n - m)$	$\gamma_j = n - j + 1$
$\mathscr{O}_k$ (OSP)	$\gamma_j = \begin{cases} n - j + 1, & 1 \le j \le k \\ m - j + 1, & k + 1 \le j \le m \end{cases}$

Table 1.2 Censoring schemes and corresponding γ_j's

The one-to-one relationship between $\mathscr{C}^m_{m,n}$ and $G_{m,n}$ shows that the number of admissible Type-II progressive censoring schemes is given by

$$|\mathscr{C}^m_{m,n}| = |\mathscr{G}_{m,n}| = \binom{n-1}{m-1}. \tag{1.5}$$

Some censoring schemes and the corresponding γ_j's are given in Table 1.2.

Example 1.1.4. Using the Data 1.1.2, we illustrate the generation of progressively Type-II censored order statistics by the example given in Viveros and Balakrishnan [875] (see also Balakrishnan and Aggarwala [86, p. 95]). We consider the censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$. The observed failure times are given by

0.19	0.78	0.96	1.31	2.78	4.85	6.50	7.35

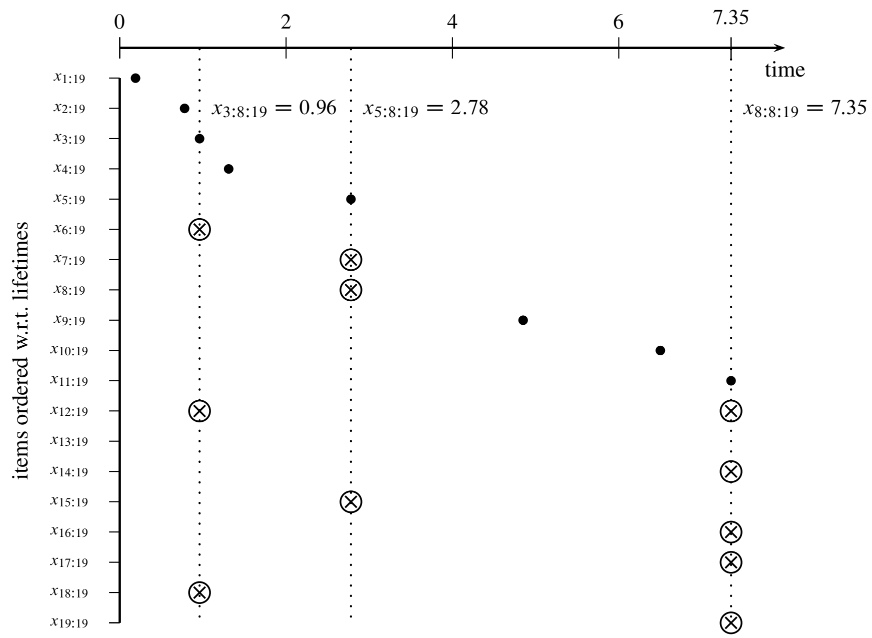

Fig. 1.5 Generation process of a progressively Type-II censored data set from Data 1.1.2 (*filled circle* denotes an observed failure time, *open crossed circle* denotes a censored value)

The items belonging to the following lifetimes are progressively censored:

Censoring time	Censored data				
$x_{3:8:19}$	3.16	8.01	36.71		
$x_{5:8:19}$	4.15	4.67	31.75		
$x_{8:8:19}$	8.27	12.06	32.52	33.91	72.89

Figure 1.5 depicts the generation procedure of progressively Type-II censored order statistics in this particular setting for Data 1.1.2. As a result, we get Data 1.1.5.

Data 1.1.5 (Nelson's progressively Type-II censored data). Progressively Type-II censored Data 1.1.2 with censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$ are as follows:

i	1	2	3	4	5	6	7	8
$x_{i:8:19}$	0.19	0.78	0.96	1.31	2.78	4.85	6.50	7.35
R_i	0	0	3	0	3	0	0	5

Remark 1.1.6. From the construction process of progressively Type-II censored order statistics, it is clear that the first r progressively Type-II censored order statistics $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{r:m:n}$ (and thus their joint distribution) depend only on the censoring numbers $R_1, \ldots, R_{r-1}$ and the sample size $\gamma_1 = n$. Therefore, we can alternatively use the right truncated censoring scheme

$$\mathscr{R}_{\rhd r-1} = (R_1, \ldots, R_{r-1}) \tag{1.6}$$

and the notation $X^{\mathscr{R}_{\rhd r-1}}_{1:m:n}, \ldots, X^{\mathscr{R}_{\rhd r-1}}_{r:m:n}$ for the Type-II right censored sample.

General Progressive Type-II Censoring

The notion of general progressively Type-II censored order statistics was introduced in Balakrishnan and Sandhu [123] (see also Balakrishnan and Aggarwala [86]). In addition to progressive Type-II censoring, left censoring of the data is also introduced in this case. In particular, it is assumed that the first r failure times are not observed. Then, starting with the $(r+1)$th failure, the progressive Type-II censoring procedure described above is applied. Introducing the censoring scheme

$$\mathscr{R} = (0^{*r}, R_{r+1}, \ldots, R_m) \in \mathscr{C}^m_{m,n}, \tag{1.7}$$

general progressively Type-II censored order statistics can be seen as a left censored sample of the progressively Type-II censored order statistics. Namely, we have the $m - r$ observations

$$X^{\mathscr{R}}_{r+1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n},$$

from the complete progressively Type-II censored sample $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$. In the sense of general progressive Type-II censoring, these random variables are denoted by

$$X^{\mathscr{R}_{\lhd r}}_{r+1:m:n}, \ldots, X^{\mathscr{R}_{\lhd r}}_{m:m:n}$$

with the left truncated censoring scheme $\mathscr{R}_{\lhd r} = (R_{r+1}, \ldots, R_m) \in \mathscr{C}^{m-r}_{m-r,n-r}$.

1.1.2 Progressive Type-I Censoring

In the progressive Type-I censoring approach, the removals are carried out at prefixed censoring times

$$T_1 < \cdots < T_k,$$

where the largest value T_k denotes the maximum experimental time. Therefore, the lifetime experiment is terminated at T_k provided that some units have survived until this time. Notice that the number of observations m is random with a value in $\{0, \ldots, n\}$. Moreover, it is worth mentioning that a progressive censoring scheme is only prefixed for time T_j, $1 \le j \le k-1$, because the experiment is terminated at T_k. Thus, all remaining objects are removed at this time. In the literature, the censoring scheme is also prefixed for the final censoring time T_k. However, this is only to have censoring plans with dimension k. Since the experiment is terminated at T_k, all surviving units are removed from the life test. Therefore, the censoring number R_k depends always on the history of the experiment and is of random nature, and so the initially planned censoring scheme will be of dimension $k-1$ in progressive Type-I censoring. It is denoted by $\mathscr{R}^0 = (R_1^0, \ldots, R_{k-1}^0)$. The corresponding set of admissible censoring schemes is denoted by

$$\mathscr{C}_{\ell,n}^{k-1} = \Big\{(r_1, \ldots, r_{k-1}) \in \mathbb{N}_0^{k-1} : \sum_{i=1}^{k-1} r_i \le n - \ell\Big\},$$

where $\ell \in \{0, \ldots, n\}$ may be a prefixed integer. For $\ell = n$, the model corresponds to common Type-I censoring. The parameter ℓ denotes the number of units that are not progressively censored (but possibly Type-I censored at T_k). Notice that the effectively applied censoring scheme is included in $\mathscr{C}_{k,n}^{k}$ as defined in (1.1).

Procedure 1.1.7 (Generation of progressively Type-I censored order statistics). Let $(\Omega, \mathfrak{A}, P)$ be a probability space and $X_1, \ldots, X_n$ be random variables on $(\Omega, \mathfrak{A}, P)$. Let $\mathscr{R}^0 = (R_1^0, \ldots, R_{k-1}^0)$ be the initially planned censoring scheme and let $T_1 < \cdots < T_k$ be ordered real numbers with $T_0 = -\infty$, $T_{k+1} = \infty$.

For $\omega \in \Omega$, the progressively Type-I censored sample

$$X_{1:m:n}^{\mathscr{R},T}(\omega), \ldots, X_{m:m:n}^{\mathscr{R},T}(\omega),$$

based on $X_1(\omega), \ldots, X_n(\omega)$, is generated as follows:

① Calculate the order statistics $X_{1:n}(\omega) \le \cdots \le X_{n:n}(\omega)$;
② Define $\mathcal{P}_j = \{\alpha \in \{1, \ldots, n\} : T_{j-1} \le X_{\alpha:n}(\omega) < T_j\}$, $1 \le j \le k+1$;
③ Let $\mathcal{N}_1 = \{1, \ldots, n\}$; $\ell = 0$;
④ Increase ℓ by 1 and let $\mathcal{Q}_\ell = \mathcal{P}_\ell \cap \mathcal{N}_\ell = \{\nu_{\ell,1}, \ldots, \nu_{\ell,s_\ell}\}$ with $\nu_{\ell,1} < \cdots < \nu_{\ell,s_\ell}$;
⑤ If $|\mathcal{N}_\ell \cap \bigcup_{\alpha=\ell+1}^{k+1} \mathcal{P}_\alpha| > R_\ell^0$, then

choose randomly a without-replacement sample $\mathcal{R}_\ell \subseteq \mathcal{N}_\ell \cap \bigcup_{\alpha=\ell+1}^{k+1} \mathcal{P}_\alpha$ with $|\mathcal{R}_\ell| = R_\ell^0$;

else

let $R_\ell = |\mathcal{N}_\ell \cap \bigcup_{\alpha=\ell+1}^{k+1} \mathcal{P}_\alpha|$, $R_\alpha = 0$, $\alpha = \ell+1, \ldots, k$, set $\mathcal{Q}_\alpha = \emptyset$, $\alpha = \ell, \ldots, k$, and go to ⑦;

⑥ Set $\mathcal{N}_{\ell+1} = \mathcal{N}_\ell \setminus (\mathcal{R}_\ell \cup \mathcal{P}_\ell)$ and go to ④;

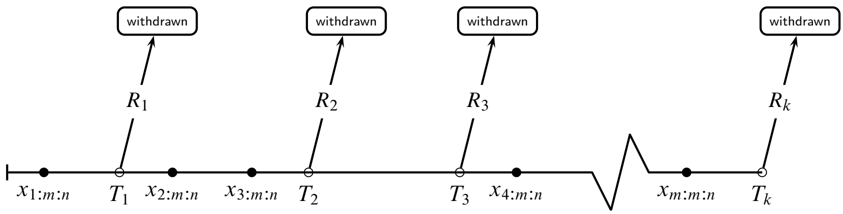

Fig. 1.6 Generation process of progressively Type-I censored order statistics

⑦ $\mathcal{M} = \bigcup_{\ell=1}^{k} \mathcal{Q}_\ell$; If $m = M(\omega) = |\mathcal{M}| > 0$, then

$$(X_{j:m:n}^{\mathscr{R},T}(\omega))_{j=1,\dots,m} = (X_{j:n}(\omega))_{j\in\mathcal{M}}.$$

$(X_{j:M:n}^{\mathscr{R},T})_{j=1,\dots,M}$ denotes the sample of progressively Type-I censored order statistics. Notice that $\mathcal{M} = \bigcup_{\ell=1}^{k} \mathcal{Q}_\ell$ represents the observed order statistics which may be an empty set. In this case, all random variables have been censored during the experiment. The random variable $M = |\mathcal{M}|$ denotes the number of observations, whereas the effectively applied censoring scheme is given by $\mathscr{R} = (R_1, \dots, R_k)$. By definition, $R_j \le R_j^0$, $j = 1, \dots, k-1$. Notice that $\sum_{j=1}^{k-1} R_j < \sum_{j=1}^{k-1} R_j^0$ is possible which means that the total number of withdrawn units may be smaller than the initially planned number of removals. For convenience, we denote the number of observations in the intervals $(-\infty, T_1)$ and $[T_{j-1}, T_j)$, $j \in \{2, \dots, k\}$, by the random variables D_1 and D_j, $j \in \{2, \dots, k\}$. Notice that $M = \sum_{j=1}^{k} D_j$.

Figure 1.6 depicts the generation procedure of progressively Type-I censored order statistics from an ordered sample $x_{1:n}, \dots, x_{n:n}$.

Example 1.1.8. Using Nelson's data 1.1.2, we illustrate the generation of progressively Type-I censored order statistics. First, let $m = 3$, $T_1 = 3$, $T_2 = 9$, $T_3 = 18$, and $\mathscr{R}^0 = (2^{*2})$. The observed failure times are given by

0.19	0.78	0.96	1.31	2.78	3.16	4.15	4.67	4.85	6.50
7.35	8.27	12.06							

The items belonging to the following lifetimes are progressively censored:

Censoring time	Censored data	
$T_1 = 3$	8.01	72.89
$T_2 = 9$	31.75	36.71
$T_3 = 18$	32.52	33.91

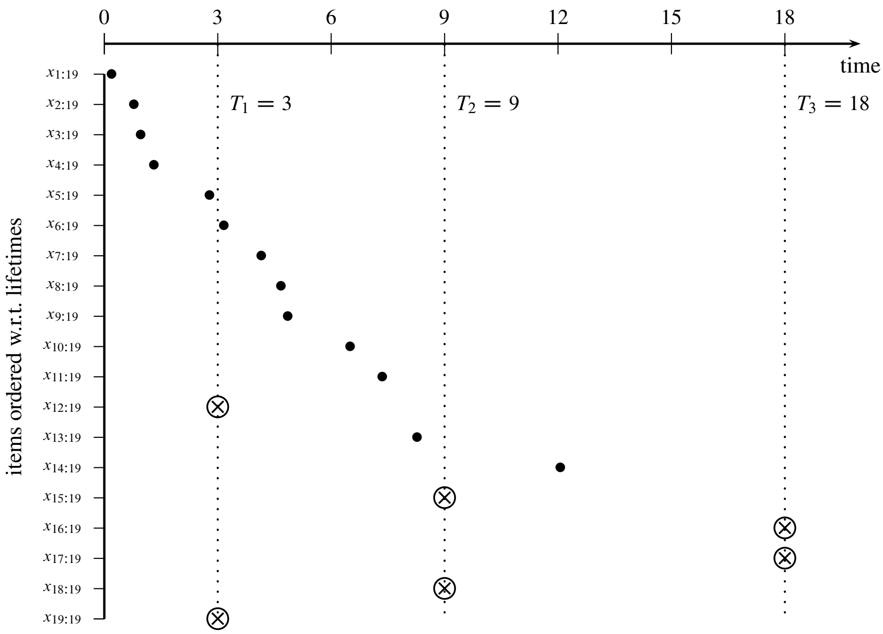

Fig. 1.7 Generation process of a progressively Type-I censored data set from Data 1.1.2 (*filled circle* denotes an observed failure time, *open crossed circle* denotes a censored value)

Thus, two units are left at the final censoring time T_3 so that the total number of progressively censored units is 6. The effectively applied censoring plan is $\mathscr{R} = (2^{*3})$. Figure 1.7 depicts the generation procedure of progressively Type-I censored order statistics in this particular setting for Data 1.1.2. As a result, we get Data 1.1.9.

Data 1.1.9 (Nelson's progressively Type-I censored data). Progressively Type-I censored Data 1.1.2 with initial censoring scheme $\mathscr{R}^0 = (2^{*2})$:

0.19	0.78	0.96	1.31	2.78	$^{\star}3$	$^{\star}3$	3.16	4.15	4.67
4.85	6.50	7.35	8.27	$^{\star}9$	$^{\star}9$	12.06	$^{\star}18$	$^{\star}18$	

$\star$ indicates (progressive) Type-I censoring

Remark 1.1.10. Progressive Type-I censoring ensures that the maximum experimental time is bounded by T_k. However, the number of observations is random and may be zero. This scheme generalizes the usual Type-I censoring in the sense that, for $k = 1$, both procedures coincide.

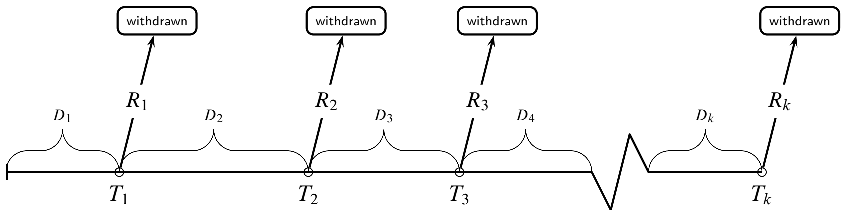

Fig. 1.8 Generation process of progressively Type-I interval censored data

Progressive Type-I Interval Censoring

Figure 1.8 depicts the generation procedure of progressively Type-I interval censored data for an ordered sample $x_{1:n}, \dots, x_{n:n}$ with censoring times $T_1 < \dots < T_k$ ($T_0 = -\infty$). In comparison to progressive Type-I censoring, only the number of observations D_i in an interval $(T_{i-1}, T_i]$ is given, i.e.,

$$D_i = \sum_{j=1}^{n} \mathbb{1}_{(T_{i-1},T_i]}(X_{j:n}), \quad 1 \le i \le k. \tag{1.8}$$

The observed values of the progressively Type-I censored order statistics are not available.

Aggarwala [11] pointed out that the censoring numbers $R_1, \dots, R_k$ can also result from a sample-dependent approach. Given percentages $\pi_1, \dots, \pi_k$ with $\pi_k = 1$, the effectively applied censoring plan is defined relative to the number of surviving units. In particular, the censoring numbers are iteratively specified via

$$R_i = \left\lfloor \left(n - \sum_{j=1}^{i} D_j - \sum_{j=1}^{i-1} R_j\right)\pi_i \right\rfloor = \left\lfloor \left(n - D_{\bullet i} - R_{\bullet i-1}\right)\pi_i \right\rfloor, \quad 1 \le i \le k. \tag{1.9}$$

1.1.3 Progressive Hybrid Censoring

Childs et al. [260] and Kundu and Joarder [561] proposed two progressive hybrid censoring procedures by introducing a stopping time T^* to a progressively Type-II censored experiment with progressively Type-II censored order statistics $X_{1:m:n} \le \dots \le X_{m:m:n}$. The termination times are defined by a given (fixed) threshold time T and the following definitions:

(i) $T_1^* = \min\{X_{m:m:n}, T\}$. This procedure is called Type-I progressive hybrid censoring scheme;
(ii) $T_2^* = \max\{X_{m:m:n}, T\}$. This procedure is called Type-II progressive hybrid censoring scheme.

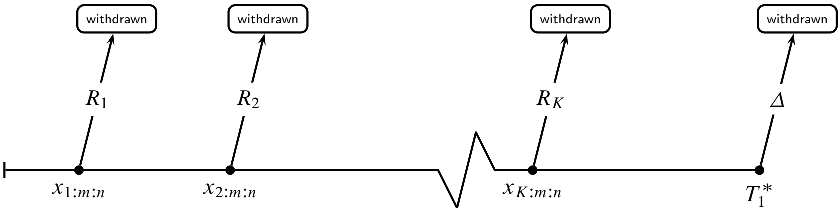

Fig. 1.9 Generation process of Type-I progressive hybrid censored order statistics.

For Type-II censored data, the first stopping time has been proposed by Epstein [352] and the second one by Childs et al. [259]. According to the above construction, the number of observations is random. In particular, it is possible to have less than m observations in Type-I progressive hybrid censoring, while in Type-II progressive hybrid censoring, we will have at least m observations.

Type-I Progressive Hybrid Censoring

In this setup, the life-testing experiment is stopped when either m failures have been observed or the threshold time T has been exceeded. Figure 1.9 depicts the generation procedure of Type-I progressive hybrid censored order statistics. The random variable K is defined by the property $X_{K:m:n} < T_1^* \leq X_{K+1:m:n}$, where $X_{0:m:n} = -\infty$. The random variable Δ describes the removals at the termination time. It is given by

$$\Delta = \begin{cases} \gamma_m - 1 = R_m, & X_{m:m:n} \leq T, \\ \gamma_K, & X_{m:m:n} > T. \end{cases}$$

It has to be mentioned that the number of observations may be zero, i.e., for the case when $X_{1:m:n} > T$. We illustrate this procedure with Data 1.1.5.

Example 1.1.11. For the Data 1.1.5, we introduce a maximum experimental time $T = 7$. Then, $T_1^* = \min\{x_{8:8:19}, 7\} = 7$ is the termination time of the life-testing experiment. Therefore, the resulting sample is given by

0.19	0.78	0.96	1.31	2.78	4.85	6.50

The units belonging to the following lifetimes are progressively censored:

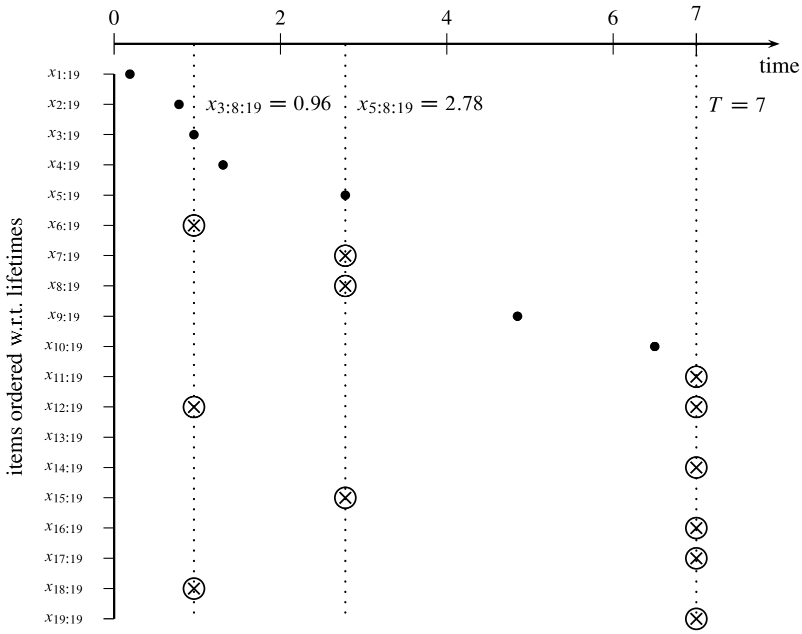

Fig. 1.10 Generation process of a Type-I progressively hybrid censored data set from Data 1.1.2 (*filled circle* denotes an observed failure time, *open crossed circle* denotes a censored value)

Censoring time	Censored data					
$x_{3:8:19}$	3.16	8.01	36.71			
$x_{5:8:19}$	4.15	4.67	31.75			
$T = 7$	7.35	8.27	12.06	32.52	33.91	72.89

Figure 1.10 depicts the generation procedure of Type-I hybrid progressively censored order statistics in this particular setting for Data 1.1.2.

Data 1.1.12 (Nelson's Type-I progressively hybrid censored data). Progressively Type-II censored Data 1.1.2 with initial censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$ and threshold $T = 7$:

i	1	2	3	4	5	6	7	8
x_i	0.19	0.78	0.96	1.31	2.78	4.85	6.50	*7
R_i	0	0	3	0	3	0	0	6

$\star$ indicates Type-I censoring

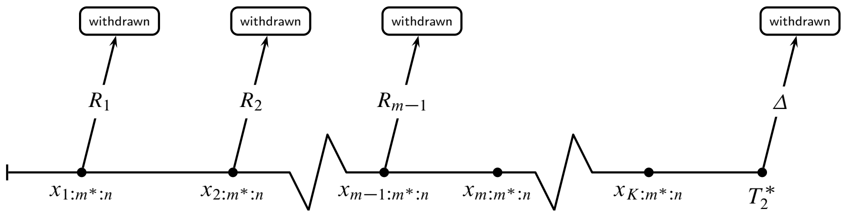

Fig. 1.11 Generation process of Type-II progressive hybrid censored order statistics.

Type-II Progressive Hybrid Censoring

As mentioned before, the number of observations is at least m, i.e., the progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{m:m:n}$ form the first m observations in the sample. To be more precise, the number of observations is between m and $R_m + m$. The idea of this hybrid procedure is to guarantee a minimum number of m observations as well as to come as close as possible to a "minimum" test duration specified by T. If $X_{m:m:n} \geq T$, the experiment terminates at the mth failure so that the progressive censoring procedure is carried out as initially planned. For $X_{m:m:n} < T$, we want to come as close as possible from below to the threshold T. This means that after the mth failure all occurring failures are observed until the threshold T is exceeded. Therefore, the censoring scheme is modified as follows:

$$\mathscr{R}^* = (R_1, \ldots, R_{m-1}, 0^{*R_m+1}) \in \mathscr{C}^{R_m+m}_{R_m+m,n}.$$

The resulting sample is given by a right censored sample

$$X^{\mathscr{R}^*}_{1:R_m+m:n}, \ldots, X^{\mathscr{R}^*}_{K:R_m+m:n},$$

where K is defined by the inequality $X^{\mathscr{R}^*}_{K:R_m+m:n} \leq T < X^{\mathscr{R}^*}_{K+1:R_m+m:n}$ ($X^{\mathscr{R}^*}_{R_m+m+1:R_m+m:n} = \infty$). Figure 1.11 depicts the generation procedure of Type-II progressive hybrid censored order statistics, where $m^* = R_m + m$. Δ is defined as $n - K - \sum_{j=1}^{m-1} R_j$.

Example 1.1.13. For the Data 1.1.5, we introduce a "minimum" experimental time $T = 18$. Then, $T_2^* = \max\{x_{8:8:19}, 18\} = 18$ is the termination time of the life-testing experiment. Therefore, the observed failure times are given by

0.19	0.78	0.96	1.31	2.78	4.85	6.50	7.35	8.27	12.06

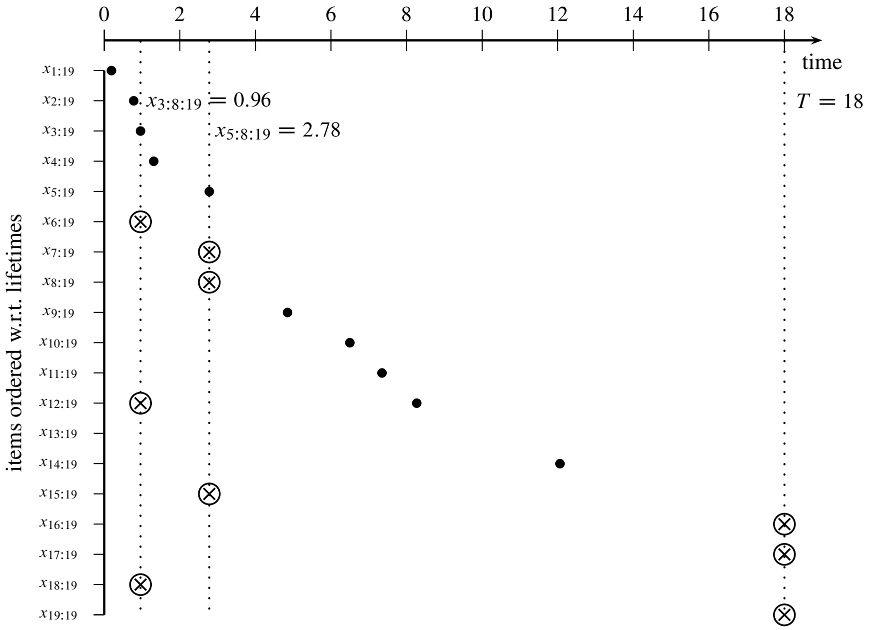

Fig. 1.12 Generation process of a Type-II progressively hybrid censored data set from Data 1.1.2 (*filled circle* denotes an observed failure time, *open crossed circle* denotes a censored value)

The units belonging to the following lifetimes are (progressively) censored:

Censoring time	Censored data		
$x_{3:8:19}$	3.16	8.01	36.71
$x_{5:8:19}$	4.15	4.67	31.75
$T = 18$	32.52	33.91	72.89

Figure 1.12 depicts the generation procedure of Type-II progressively hybrid censored order statistics in this particular setting for Data 1.1.2.

Data 1.1.14 (Nelson's Type-II progressively hybrid censored data). Type-II progressively hybrid censored Data 1.1.2 with initial censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$ and threshold $T = 18$:

i	1	2	3	4	5	6	7	8	9	10	11
x_i	0.19	0.78	0.96	1.31	2.78	4.85	6.50	7.35	8.27	12.06	18
R_i	0	0	3	0	3	0	0	0	0	0	$^{\star}3$

$\star$ indicates Type-I censoring

1.2 Probabilistic Models in Progressive Censoring

The progressively censored order statistics introduced in the preceding section are directly constructed from a sample $X_1, \ldots, X_n$ of random variables. The generation procedures are based only on the random variables or their realizations. Distributional properties of $X_1, \ldots, X_n$ are not involved in this process. In the following chapters, we will consider several distributional assumptions for the underlying joint cumulative distribution function $F^{X_1,\ldots,X_n}$. For brevity, we introduce the most important cases here.

Most papers on progressive censoring deal with the IID model.

Model 1.2.1 (IID model). Let F be a cumulative distribution function.

In the IID model, progressively censored samples are based on the following distributional assumption:

$X_1, \ldots, X_n$ are IID random variables with common cumulative distribution function F.

Progressively censored experiments have been considered for a wide range of lifetime distributions such as normal, exponential, gamma, Rayleigh, Weibull, extreme value, log-normal, inverse Gaussian, logistic, Laplace, and Pareto distributions, and properties of these distributions are used many times in the derivations. For details and results, we refer to, e.g., Johnson et al. [483, 484] and Marshall and Olkin [640].

The assumption of identical distribution is dropped in the INID model.

Model 1.2.2 (INID model). Let $F_1, \ldots, F_n$ be cumulative distribution functions.

In the INID model, progressively censored samples are based on the following distributional assumption:

$X_1, \ldots, X_n$ are independent random variables with $X_i \sim F_i$, $1 \leq i \leq n$.

Model 1.2.3 (Single outlier model). Let F and G be cumulative distribution functions.

In the single outlier model, progressively censored samples are based on the following distributional assumption:

$X_1, \ldots, X_n$ are independent random variables with $X_i \sim F$, $1 \leq i \leq n-1$, and $X_n \sim G$.

Model 1.2.4 (p-outlier model). Let F and G be cumulative distribution functions.

In the p-outlier model, progressively censored samples are based on the following distributional assumption:

$X_1, \ldots, X_n$ are independent random variables with $X_i \sim F$, $1 \leq i \leq n-p$, and $X_i \sim G$, $n-p+1 \leq i \leq n$.

In the preceding models, independence or distributional assumptions are made. In the general model, we do not impose any restriction.

Model 1.2.5 (General model). In the general model, progressively censored samples are based on the following distributional assumption:

$X_1, \dots, X_n$ are random variables with joint cumulative distribution function $F^{X_1,\dots,X_n}$.

Cramer and Lenz [303] have formulated the probabilistic assumptions in the generation process of progressively Type-II censored order statistics in detail. These postulates reflect the conditions for random removals commonly accepted in progressive Type-II censoring. The notations are taken from Procedure 1.1.3. Moreover, define the random variables $K_1, \dots, K_m$ via the relation $K_j(\omega) = k_j$, $\omega \in \Omega$, $1 \leq j \leq m$. Hence, $X_{K_1:n}, \dots, X_{K_m:n}$ represent the order statistics in the sample $X_1, \dots, X_n$ which form the progressively Type-II censored sample $X_{1:m:n}, \dots, X_{m:m:n}$.

Assumption 1.2.6. Let $i \in \{2, \dots, m\}$. Then,

(i) given k_i and the set $N_i \setminus \{k_i\}$, the random set $\mathcal{R}_{i,k_i} \subseteq N_i \setminus \{k_i\}$ with $|\mathcal{R}_{i,k_i}| = R_i$ is independent of

$$(K_1, \mathcal{R}_{1,K_1}, \dots, K_{i-1}, \mathcal{R}_{i-1,K_{i-1}});$$

(ii) given $N_i \setminus \{k_i\}$, the without-replacement sample $\mathcal{R}_{i,k_i}$ of size R_i is drawn from $N_i \setminus \{k_i\}$ according to a uniform distribution;

(iii) the lifetimes $X_1, \dots, X_n$ and the progressive censoring assignment $(K_1, \mathcal{R}_{1,K_1}, \dots, K_m, \mathcal{R}_{m,K_m})$ are independent.

Chapter 2
Progressive Type-II Censoring: Distribution Theory

2.1 Joint Distribution

The general quantile representation for progressively Type-II censored order statistics due to Balakrishnan and Dembińska [96] (see also Balakrishnan and Dembińska [97] and Cramer and Kamps [301]) provides a powerful tool in the derivation of distributional results. Many identities can be obtained first for uniform distributions and then transferred to any particular distribution of interest.

Theorem 2.1.1. Suppose $X_{1:m:n}, \dots, X_{m:m:n}$ and $U_{1:m:n}, \dots, U_{m:m:n}$ are progressively Type-II censored order statistics based on a cumulative distribution function F and a uniform distribution, respectively. Then,

$$(X_{j:m:n})_{1 \le j \le m} \stackrel{\mathrm{d}}{=} \big(F^{\leftarrow}(U_{j:m:n})\big)_{1 \le j \le m}.$$

Proof. Let $X_1, \dots, X_n$ and $U_1, \dots, U_n$ be IID samples from F and a uniform distribution on a probability space $(\Omega, \mathfrak{A}, P)$, respectively. Then,

$$(X_1, \dots, X_n) \stackrel{\mathrm{d}}{=} (F^{\leftarrow}(U_1), \dots, F^{\leftarrow}(U_n))$$

and we conclude that the vector $(X_{1:m:n}, \dots, X_{m:m:n})$ has the same distribution as progressively Type-II censored order statistics based on the sample $F^{\leftarrow}(U_1), \dots, F^{\leftarrow}(U_n)$. Therefore, it is sufficient to prove that these progressively Type-II censored order statistics have the same values as the random variables $\big(F^{\leftarrow}(U_{j:m:n})\big)_{1 \le j \le m}$ for any fixed $\omega \in \Omega$ in the underlying probability space. For brevity, let $u_j = U_j(\omega)$, $1 \le j \le n$, and $u_i^* = U_{i:m:n}(\omega)$, $1 \le i \le m$. Notice that $F^{\leftarrow}$ is an increasing function and that for given numbers $x_1, \dots, x_r$,

$$\min_{1 \le k \le r} F^{\leftarrow}(x_k) = F^{\leftarrow}\Big(\min_{1 \le k \le r} x_k\Big). \tag{2.1}$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_2,

From the generation process 1.1.3, we find that u_j^* is defined by the minimum of a selection M_j of numbers. Thus,

$$u_j^* = \min_{i \in M_j} u_i, \quad 1 \le j \le m,$$

and we obtain from (2.1) that

$$F^{\leftarrow}(u_j^*) = \min_{i \in M_j} F^{\leftarrow}(u_i), \quad 1 \le j \le m.$$

This yields the desired quantile representation. □

The result can alternatively be proved by using the mixture representation in Theorem 10.1.1 due to Fischer et al. [371]. It is worth mentioning that the quantile representation in Theorem 2.1.1 shows that

$$F^{\mathbf{X}^{\mathscr{R}}} = F^{\mathbf{U}^{\mathscr{R}}} \circ (F^{*m}). \tag{2.2}$$

In particular, any marginal cumulative distribution function can be written in this way. For instance, $F^{X_{r:m:n}} = F^{U_{r:m:n}} \circ F$, $F^{X_{r:m:n}, X_{s:m:n}} = F^{U_{r:m:n}, U_{s:m:n}} \circ (F, F)$, etc.

An important tool in the analysis of progressively Type-II censored order statistics is the joint density function of uniform progressively Type-II censored order statistics which is given in the following theorem. A formal proof is provided in Sect. 10.2 for the more general INID situation.

Theorem 2.1.2. The joint density function of uniform progressively Type-II censored order statistics $U_{1:m:n}, \ldots, U_{m:m:n}$ is given by

$$f^{U_{1:m:n},\ldots,U_{m:m:n}}(\mathbf{u}_m) = \prod_{j=1}^{m} \left[\gamma_j (1-u_j)^{R_j}\right], \ 0 \le u_1 \le \cdots \le u_m \le 1. \tag{2.3}$$

If F is absolutely continuous, the joint density function of progressively Type-II censored order statistics is given in the following corollary (cf. Cohen [267], Herd [440], and Balakrishnan and Aggarwala [86]). It follows directly from Theorems 2.1.1 and 2.1.2 [see also (2.2)].

Corollary 2.1.3. The joint density function of progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{m:m:n}$ based on a cumulative distribution function F with density function f is given by

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}_m) = \prod_{j=1}^{m} \left[\gamma_j f(x_j)(1-F(x_j))^{R_j}\right], \quad x_1 \le \cdots \le x_m. \tag{2.4}$$

Example 2.1.4.

(i) For order statistics, i.e., $m = n$, $\mathscr{R} = (0^{*m})$ and $\gamma_j = n - j + 1$, $1 \le j \le n$, the joint density function is given by

$$f^{X_{1:n},\dots,X_{n:n}}(\mathbf{x}_n) = n! \prod_{j=1}^{n} f(x_j), \quad x_1 \le \dots \le x_n \tag{2.5}$$

(cf. Arnold et al. [58] and David and Nagaraja [327]).

(ii) The censoring plan $\mathscr{R} = (R^{*m})$ with equal removal number $R \in \mathbb{N}_0$ is called equi-balanced censoring scheme. Progressively Type-II censored order statistics with such a censoring scheme possess the joint density function

$$\begin{aligned} f^{X_{1:m:n},\dots,X_{m:m:n}}(\mathbf{x}_m) &= \prod_{j=1}^{m} \left[\gamma_j f(x_j)(1 - F(x_j))^R\right] \\ &= m! \prod_{j=1}^{m} \left[(R+1) f(x_j)(1 - F(x_j))^R\right], \quad x_1 \le \dots \le x_m. \end{aligned} \tag{2.6}$$

Notice that $\gamma_j = (m - j + 1)(R + 1)$, $1 \le j \le m$ (see also Table 1.2). Defining g by $g(t) = (R + 1) f(t)(1 - F(t))^R$, we find that the density function in (2.6) equals the joint density function of order statistics from a sample of size m and with density function g. Hence, this particular scheme does not lead to a new model. It can be seen simply as an order statistic model from a different distribution. Notice that this distribution is the same as that of the minimum of $R+1$ IID random variables from f. This comment applies also to the models with non-absolutely continuous distribution.

(iii) In the OSP-case with censoring scheme $\mathscr{O}_k$, $k \in \{1, \dots, m\}$, the joint density function is given by

$$\begin{aligned} f^{\mathbf{X}^{\mathscr{O}_k}}(\mathbf{x}_m) &= \Big[\prod_{j=1}^{m} \left[\gamma_j f(x_j)\right]\Big](1 - F(x_k))^{n-m} \\ &= \frac{n!(m-k)!}{(n-k)!}\Big[\prod_{j=1}^{m} f(x_j)\Big](1 - F(x_k))^{n-m}, \quad x_1 \le \dots \le x_m. \end{aligned}$$

2.2 On the Connection to Generalized Order Statistics and Sequential Order Statistics

It is obvious from the joint density function of uniform progressively Type-II censored order statistics presented in Theorem 2.1.2 that uniform progressively Type-II censored order statistics can be seen as particular uniform generalized order statistics introduced by Kamps [498, 499] (see also Cramer [285, 288], Cramer and Kamps [300, 301], and Kamps [502]). Commonly, generalized order statistics are parametrized by one of the following sets of parameters which are very similar to those for progressively Type-II censored order statistics given on page 7:

(i) $k, m_1, \dots, m_{n-1}$,
(ii) $\gamma_1, \dots, \gamma_n > 0$.

The density function of uniform generalized order statistics is usually given as (see Kamps [498, p. 49])

$$f^{U(1,n,\mathbf{m},k),\dots,U(n,n,\mathbf{m},k)}(\mathbf{u}_n) = k\Big(\prod_{j=1}^{n-1}\gamma_j\Big)\Big[\prod_{j=1}^{n-1}(1-u_j)^{m_j}\Big](1-u_n)^{k-1},$$
$$0 \le u_1 \le \dots \le u_n \le 1, \qquad (2.7)$$

where $\mathbf{m} = (m_1, \dots, m_{n-1})$. Generalized order statistics $X(1, n, \mathbf{m}, k), \dots, X(n, n, \mathbf{m}, k)$ based on an arbitrary cumulative distribution function F are defined via the quantile transformation

$$X(j, n, \mathbf{m}, k) = F^{\leftarrow}(U(j, n, \mathbf{m}, k)), \quad 1 \le j \le n,$$

so that the same comment applies to progressively Type-II censored order statistics from an arbitrary cumulative distribution function using the representation in terms of the quantile function (see, e.g., the density function given in Corollary 2.1.3).

Hence, the joint density function has a similar form as (2.7) (see (2.4) for progressively Type-II censored order statistics). Sometimes, the parameters are suppressed in the notation and (uniform) generalized order statistics are also denoted by $U_{*,1}, \dots, U_{*,n}$. Notice that m and n are differently used in both models. However, we have the correspondences $R_j = m_j$ and k equals the last γ_j. To be more specific, we consider uniform progressively Type-II censored order statistics with censoring scheme $\mathscr{R}$. Then, they can be seen as uniform generalized order statistics

$$U(1, m, \mathscr{R}_{\rhd m-1}, R_m + 1), \dots, U(m, m, \mathscr{R}_{\rhd m-1}, R_m + 1),$$

where $\mathscr{R}_{\rhd m-1} = (R_1, \dots, R_{m-1})$ denotes a right truncated censoring scheme [see (1.6)]. Thus, progressively Type-II censored order statistics are generalized order statistics in distribution wherein some restrictions have to be imposed on the parameters $\gamma_1, \ldots, \gamma_m > 0$ of generalized order statistics as introduced in Cramer and Kamps [301] [see also (1.3)]:

(i) $\gamma_j \in \mathbb{N}$, $j = 1, \ldots, m$,
(ii) $n = \gamma_1 > \cdots > \gamma_m \geq 1$.

Therefore, in the model of progressively Type-II censored order statistics, the parameters $\gamma_1, \ldots, \gamma_m$ are strictly decreasingly ordered positive integers. Although this difference seems to be minor, the picture becomes simpler for progressively Type-II censored order statistics in many cases. In particular, calculations become easier and representations get simpler. An example may be the representation of the marginal density functions (see (2.28) for progressively Type-II censored order statistics and Cramer and Kamps [301] for the density functions of generalized order statistics in terms of Meijer's G-functions). Moreover, it turns out that many results obtained for generalized order statistics are valid for progressively Type-II censored order statistics without imposing further restrictions on the parameters.

Notice that the connection is only of distributional nature, but it can be used in many areas. For instance, results for moments can be directly applied to progressively Type-II censored order statistics. Similar comments apply to characterizations, stochastic orders, reliability properties, inferential results, etc., which are available for generalized order statistics with arbitrary parameters. Therefore, many results can be directly taken from properties of generalized order statistics. We utilize this connection in the following by reformulating the results in terms of progressively Type-II censored order statistics. On the other hand, extensions to generalized order statistics are also possible in many settings.

However, one has to be careful using this connection because many results for generalized order statistics are often obtained only for the so-called m-generalized order statistics. In this case, the parameters satisfy the condition $m_1 = \cdots = m_{n-1}$. For progressively Type-II censored order statistics, this corresponds to the case of an equi-balanced censoring scheme $\mathscr{R} = (R_1, \ldots, R_m) = (R^{*m})$ with $R \in \mathbb{N}_0$.

Moreover, it has to be mentioned that some results are also available in terms of sequential order statistics from some cumulative distribution functions $F_1, \ldots, F_m$. This model has been introduced in Kamps [498] in order to extend the model of k-out-of-m systems (see also Burkschat [230], Cramer [288], and Cramer and Kamps [300]). According to Cramer and Kamps [301], the distribution of sequential order statistics $X_*^{(1)}, \ldots, X_*^{(m)}$ (based on $F_1, \ldots, F_n$) can be represented via quantile-type transformations

$$X_*^{(r)} = F_r^{\leftarrow}(X^{(r)}) \quad \text{with } X^{(r)} = 1 - V_r \overline{F}_r(X_*^{(r-1)}), \quad 1 \leq r \leq m,$$

where $X_*^{(0)} = -\infty$, $F_1, \ldots, F_m$ are cumulative distribution functions with $F_1^{\leftarrow}(1) \leq \cdots \leq F_m^{\leftarrow}(1)$, and $V_1, \ldots, V_m$ are independent random variables with $V_r \sim \mathsf{Beta}(m - r + 1, 1)$, $1 \leq r \leq m$.

As pointed out in Cramer and Kamps [300], sequential order statistics can be seen as generalized order statistics based on F if the cumulative distribution functions

$F_1, \ldots, F_m$ satisfy the proportional hazards relation $\overline{F}_j = \overline{F}^{\alpha_j}$, $1 \le j \le m$, for some continuous cumulative distribution function F and $\alpha_1, \ldots, \alpha_m > 0$. Using this connection, results for sequential order statistics can also be applied to progressively Type-II censored order statistics.

Finally, the distribution of exponential progressively Type-II censored order statistics is connected to the distribution of order statistics from a Weinman multivariate exponential distribution which is an extension of Freund's bivariate exponential distribution (see Block [206], Freund [383], and Weinman [896]). As pointed out by Cramer and Kamps [297] and Cramer and Kamps [300], this connection can also be utilized in the framework of progressively Type-II censored order statistics.

2.3 Results for Particular Population Distributions

2.3.1 Exponential Distributions

In this section, progressively Type-II censored order statistics are based on a two-parameter exponential distribution $\mathsf{Exp}(\mu, \vartheta)$, $\mu \in \mathbb{R}, \vartheta > 0$. From Corollary 2.1.3, we find directly the respective representation of the joint density function in the exponential case.

Corollary 2.3.1. The joint density function of exponential progressively Type-II censored order statistics $Z_{1:m:n}, \ldots, Z_{m:m:n}$ from an $\mathsf{Exp}(\mu, \vartheta)$-distribution is given by

$$f^{\mathbf{Z}^{\mathscr{R}}}(\mathbf{x}_m) = \Big(\prod_{j=1}^{m} \gamma_j\Big) \exp\Big\{-\frac{1}{\vartheta}\sum_{j=1}^{m}(R_j+1)(x_j-\mu)\Big\}, \quad \mu \le x_1 \le \cdots \le x_m. \tag{2.8}$$

The joint density function given in (2.8) yields directly the fundamental result that the normalized spacings of exponential progressively Type-II censored order statistics are IID exponential random variables. This observation is due to Thomas and Wilson [843] (see also Viveros and Balakrishnan [875]). Let

$$S_r^{\mathscr{R}} = \gamma_r (Z_{r:m:n}^{\mathscr{R}} - Z_{r-1:m:n}^{\mathscr{R}}), \quad r = 1, \ldots, m, \tag{2.9}$$

be the (normalized) spacings of $Z_{1:m:n}^{\mathscr{R}}, \ldots, Z_{m:m:n}^{\mathscr{R}}$, where $Z_{0:m:n}^{\mathscr{R}} = \mu$. Moreover, let $\mathbf{S}^{\mathscr{R}} = (S_1^{\mathscr{R}}, \ldots, S_m^{\mathscr{R}})'$ and $\mathbf{Z}^{\mathscr{R}} = (Z_{1:m:n}^{\mathscr{R}}, \ldots, Z_{m:m:n}^{\mathscr{R}})'$. Then,

$$\mathbf{S}^{\mathscr{R}} = T(\mathbf{Z}^{\mathscr{R}} - \mu\mathbb{1}) \tag{2.10}$$

with

$$T = \begin{pmatrix} \gamma_1 & 0 & \cdots & \cdots & 0 \\ -\gamma_2 & \gamma_2 & 0 & \cdots & 0 \\ 0 & -\gamma_3 & \gamma_3 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\gamma_m & \gamma_m \end{pmatrix}.$$

Theorem 2.3.2. The spacings $S_1^{\mathscr{R}}, \dots, S_m^{\mathscr{R}}$ are independently and identically distributed with $S_r^{\mathscr{R}} \sim \text{Exp}(\vartheta)$, $r = 1, \dots, m$.

Proof. Since $\gamma_j > 0$, $1 \le j \le m$, T is a regular matrix with

$$T^{-1} = \begin{pmatrix} 1/\gamma_1 & 0 & \cdots\cdots & & 0 \\ 1/\gamma_1 & 1/\gamma_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & \vdots \\ 1/\gamma_1 & 1/\gamma_2 & \cdots & \cdots & 1/\gamma_m \end{pmatrix} \quad \text{and} \quad \det T = \prod_{j=1}^{m} \gamma_j .$$

Now, the density transformation theorem yields the density function

$$f^{\mathbf{S}^{\mathscr{R}}}(\mathbf{t}) = \frac{1}{|\det T|} \cdot f^{\mathbf{Z}^{\mathscr{R}}}(T^{-1}\mathbf{t} + \mu\mathbb{1}), \quad \mathbf{t} = (t_1, \dots, t_m). \tag{2.11}$$

Noticing that $\gamma_j - \gamma_{j+1} = R_j + 1$, $1 \le j \le m-1$, and $\gamma_m = R_m + 1$, we find

$$\begin{aligned} \sum_{j=1}^{m} (R_j + 1)[(T^{-1}\mathbf{t} + \mu\mathbb{1})_j - \mu] &= \underbrace{(\gamma_1 - \gamma_2, \dots, \gamma_{m-1} - \gamma_m, \gamma_m)}_{=\mathbb{1}'T} T^{-1}\mathbf{t} \\ &= \mathbb{1}'\mathbf{t} = \sum_{j=1}^{m} t_j . \end{aligned}$$

Thus, (2.11) in combination with (2.8) yields the density function $f^{\mathbf{S}^{\mathscr{R}}}(\mathbf{t}) = \exp\big\{-\frac{1}{\vartheta}\sum_{j=1}^{m} t_j\big\}$, $t_1, \dots, t_m \ge 0$. This proves the desired result. □

Theorem 2.3.2 yields the following well-known result for spacings of order statistics due to Sukhatme [826]. It follows from Theorem 2.3.2 by choosing the censoring scheme $\mathscr{R} = (0^{*m})$.

Corollary 2.3.3 (Sukhatme [826]). The spacings $S_{1,n}, \dots, S_{n,n}$ of exponential order statistics $Z_{1:n}, \dots, Z_{n:n}$ from an $\text{Exp}(\mu, \vartheta)$-distribution are independently and identically distributed with $S_{r,n} \sim \text{Exp}(\vartheta)$, $r = 1, \dots, n$.

The preceding result can be extended easily to one-parameter exponential families.

Remark 2.3.4. Suppose the cumulative distribution function F_θ of a one-parameter exponential family is given by

$$F_\theta(x) = 1 - \mathrm{e}^{-\eta(\theta)d(x)}, \quad x \in (\alpha, \omega), \tag{2.12}$$

with $-\infty \le \alpha < \omega \le \infty$, $d(\alpha+) = \lim_{x\to\alpha+} d(x) = 0$, $d(\omega-) = \lim_{x\to\omega-} d(x) = \infty$, where d is nondecreasing and differentiable and η is positive and twice differentiable. Then, the random variables $\gamma_j(d(X_{j:m:n}) - d(X_{j-1:m:n}))$, $1 \le j \le m$, are IID exponential random variables with mean $1/\eta(\theta)$ (see, for example, Cramer and Kamps [300]). This family is also discussed in the context of Fisher information in Sect. 9.1.3.

The exponential family defined via (2.12) can be characterized by the property of hazard rate factorization, i.e., by $\lambda_\theta(x) = \eta(\theta)d'(x)$. It includes, for instance, the exponential distribution (scale parameter), the extreme value distribution (location parameter), the Weibull distribution (scale parameter), and the Pareto distribution (shape parameter). Characterizations of distribution in terms of the Fisher information are given by, e.g., Hofmann et al. [445], Zheng [941], and Gertsbakh and Kagan [396].

Remark 2.3.5. Bairamov and Eryılmaz [78] discussed minimal and maximal (non-normalized) spacings for exponential progressively Type-II censored order statistics, i.e.,

$$S_1^{\star\mathscr{R}} = \frac{1}{\gamma_1} S_1^{\mathscr{R}} = Z_{1:m:n}^{\mathscr{R}},$$

$$S_j^{\star\mathscr{R}} = \frac{1}{\gamma_j} S_j^{\mathscr{R}} = Z_{j:m:n}^{\mathscr{R}} - Z_{j-1:m:n}^{\mathscr{R}}, \quad j = 2, \ldots, m.$$

In particular, they were interested in the random indicators η and ν with

$$S_\nu^{\star\mathscr{R}} = \min_{1\le j\le m} S_j^{\star\mathscr{R}}, \quad S_\eta^{\star\mathscr{R}} = \max_{1\le j\le m} S_j^{\star\mathscr{R}}.$$

Clearly, Theorem 2.3.2 implies $S_j^{\star\mathscr{R}} \sim \mathrm{Exp}(\vartheta/\gamma_j)$, $1 \le j \le m$. Bairamov and Eryılmaz [78] obtained expressions for the joint probability mass function as well as for the marginal probability mass functions of ν and η. For instance, for $k = 1, \ldots, m$ and an underlying Exp(ϑ)-distribution, $S_j^{\star\mathscr{R}}$, $1 \le j \le m$, are independent random variables. This directly leads to the expressions

$$P(\nu = k) = \frac{\gamma_k}{\sum_{j=1}^m \gamma_j},$$

$$P(\eta = k) = \gamma_k \int_0^\infty \prod_{j=1, j\neq k}^m \left(1 - \mathrm{e}^{-\gamma_j t}\right) \mathrm{e}^{-\gamma_k t}\, dt,$$

which are independent of the scale parameter ϑ. Moreover, the joint and marginal cumulative distribution functions of the maximal spacing can be obtained. For $0 < x < y$, we get

$$P(S_\nu^{\star\mathscr{R}} \le x, S_\eta^{\star\mathscr{R}} \le y) = \prod_{j=1}^{m} \big(1 - \mathrm{e}^{-\gamma_j y/\vartheta}\big) - \prod_{j=1}^{m} \big(\mathrm{e}^{-\gamma_j x/\vartheta} - \mathrm{e}^{-\gamma_j y/\vartheta}\big),$$

$$P(S_\nu^{\star\mathscr{R}} \le x) = 1 - \exp\Big\{-\Big(\sum_{j=1}^{m}\gamma_j\Big)x/\vartheta\Big\},$$

$$P(S_\eta^{\star\mathscr{R}} \le y) = \prod_{j=1}^{m} \big(1 - \mathrm{e}^{-\gamma_j y/\vartheta}\big).$$

From (2.10), we find the following representation of exponential progressively Type-II censored order statistics in terms of the spacings:

$$\mathbf{Z}^{\mathscr{R}} = T^{-1}\mathbf{S}^{\mathscr{R}} + \mu\mathbb{1} \quad \text{or} \quad Z_{r:m:n}^{\mathscr{R}} = \mu + \sum_{j=1}^{r} \frac{1}{\gamma_j} S_j^{\mathscr{R}}, 1 \le r \le m. \tag{2.13}$$

Thus, we can write exponential progressively Type-II censored order statistics as a weighted sum of independent exponential random variables. This expression will be very useful in deriving marginal distributions, moments, recurrence relations, etc.

Moreover, (2.13) yields an interesting representation of progressively Type-II censored order statistics. In particular, we have from Theorem 2.1.1 in the exponential case

$$Z_{r:m:n} \stackrel{\mathsf{d}}{=} \mu - \vartheta \log(1 - U_{r:m:n}), \quad 1 \le r \le m,$$

or, equivalently, with $F_{\exp}(t) = 1 - \mathrm{e}^{-(t-\mu)/\vartheta}$, $t \ge \mu$,

$$F_{\exp}(Z_{r:m:n}) \stackrel{\mathsf{d}}{=} U_{r:m:n}, \quad 1 \le r \le m.$$

From (2.13), we note that

$$F_{\exp}(Z_{r:m:n}) = 1 - \prod_{j=1}^{r} \Big(\mathrm{e}^{-S_j^{\mathscr{R}}/\vartheta}\Big)^{1/\gamma_j}$$

with $S_j^{\mathscr{R}}/\vartheta \sim \mathsf{Exp}(1)$. Thus, $U_j = \mathrm{e}^{-S_j^{\mathscr{R}}/\vartheta}$, $1 \le j \le m$, are independent uniformly distributed random variables. This yields the representation

$$U_{r:m:n} \stackrel{\mathsf{d}}{=} 1 - \prod_{j=1}^{r} U_j^{1/\gamma_j}, \quad 1 \le r \le m,$$

of $U_{r:m:n}$ as a product of independent random variables. Combining this expression with the quantile representation from Theorem 2.1.1, we arrive at the following theorem (see also Cramer and Kamps [301]).

Theorem 2.3.6. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from an arbitrary cumulative distribution function F and $U_1, \ldots U_m \stackrel{\text{iid}}{\sim} \mathsf{U}(0,1)$. Then,

$$X_{r:m:n} \stackrel{\mathsf{d}}{=} F^{\leftarrow}\Big(1 - \prod_{j=1}^{r} U_j^{1/\gamma_j}\Big), \quad 1 \le r \le m. \tag{2.14}$$

Sometimes, the following representation in terms of exponential progressively Type-II censored order statistics is useful which is immediate from Theorem 2.1.1 and the above theorem.

Corollary 2.3.7. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from an arbitrary cumulative distribution function F and $Z_{1:m:n}, \ldots,$ $Z_{m:m:n}$ be progressively Type-II censored order statistics with the same censoring scheme. Moreover, let $Z_1, \ldots, Z_m \stackrel{\text{iid}}{\sim} \mathsf{Exp}(1)$ and $\Psi(x) = F^{\leftarrow}(1 - \mathsf{e}^{-x})$, $x \ge 0$. Then,

$$X_{r:m:n} \stackrel{\mathsf{d}}{=} F^{\leftarrow}\left(1 - \mathsf{e}^{-Z_{r:m:n}}\right) \stackrel{\mathsf{d}}{=} \Psi\Big(\sum_{j=1}^{r} \frac{1}{\gamma_j} Z_j\Big), \quad 1 \le r \le m.$$

The above representation can be simplified for order statistics. In this particular setup, we find the following result which shows that uniform order statistics are beta distributed. Therefore, order statistics have been called transformed beta variables (see, e.g., Blom [208]).

Corollary 2.3.8. For uniform order statistics $U_{1:n}, \ldots, U_{n:n}$, we have $U_{r:n} \sim \mathsf{Beta}(r, n-r+1)$, $1 \le r \le n$.

Proof. By definition, we have $\gamma_j = n - j + 1$, $1 \le j \le n$. Thus, we obtain for uniform order statistics

$$1 - U_{r:n} \stackrel{\mathsf{d}}{=} \prod_{j=1}^{r} U_j^{1/\gamma_j}, \quad U_j^{1/\gamma_j} \sim \mathrm{Beta}(n-j+1, 1).$$

Using a result of Rao [738] (see also Jambunathan [477], Kotlarski [545], Fan [359], and Johnson et al. [484, p. 257]) we get

$$\prod_{j=1}^{r} U_j^{1/\gamma_j} \sim \mathsf{Beta}(n-r+1, r).$$

Hence, $U_{r:n}$ has a $\mathsf{Beta}(r, n-r+1)$-distribution. □

A simple representation also holds for one-step censoring plans.

Corollary 2.3.9. Let $\mathscr{O}_k$, $1 \le k \le m$, be a one-step censoring plan. Then,

$$U_{r:m:n} \stackrel{\mathsf{d}}{=} U_{r:n}, \quad 1 \le r \le k,$$

$$U_{r:m:n} \stackrel{\mathsf{d}}{=} 1 - (1 - U_{k:n}) \cdot (1 - \widetilde{U}_{r-k:m-k}), \quad k+1 \le r \le m,$$

where $\widetilde{U}_{r-k:m-k}$ denotes the $(r-k)$th order statistic in a sample $\widetilde{U}_1, \ldots, \widetilde{U}_{m-k}$ from a uniform distribution and independent of $U_1, \ldots, U_n$. Thus, the distribution of $1 - U_{r:m:n}$ is given by the distribution of a product of independent beta random variables with parameters $(n-k+1, k)$ and $(m-r+1, r-k)$, respectively.

Proof. From Table 1.2, we have $\gamma_j = n-j+1$, $1 \le j \le k$, and $\gamma_j = m-j+1$, $k+1 \le j \le m$. Thus, for $1 \le r \le k$, we obtain from Corollary 2.3.8 that $U_{r:m:n} \stackrel{\mathsf{d}}{=} U_{r:n} \sim \mathsf{Beta}(r, n-r+1)$. Let $r > k$. Then, the product $\prod_{j=1}^{r} U_j^{1/\gamma_j}$ equals

$$\prod_{j=1}^{k} U_j^{1/(n-j+1)} \prod_{j=k+1}^{r} U_j^{1/(m-j+1)} = \prod_{j=1}^{k} U_j^{1/(n-j+1)} \prod_{j=1}^{r-k} U_j^{1/(m-k-j+1)}.$$

The first product has a $\mathsf{Beta}(n-k+1, k)$ distribution, while the second one has a $\mathsf{Beta}(m-r+1, r-k)$ distribution. By the independence of the factors, we obtain the desired result. □

Since $1 - U_{j:\nu} \stackrel{\mathsf{d}}{=} U_{\nu-j+1:\nu}$, the result of Corollary 2.3.9 can be expressed as

$$\begin{aligned} U_{r:m:n} &\stackrel{\mathsf{d}}{=} U_{r:n}, \quad 1 \le r \le k, \\ U_{r:m:n} &\stackrel{\mathsf{d}}{=} 1 - U_{n-k+1:k} \cdot \widetilde{U}_{m-r+1:m-k}, \quad k+1 \le r \le m. \end{aligned} \tag{2.15}$$

This representation can be used for simulation purposes (see Algorithm 8.1.8).

2.3.2 Reflected Power Distribution and Uniform Distribution

Theorem 2.3.6 has some interesting implications to generalized Pareto distributions (see Definition A.1.11). In particular, we find for reflected power distributions with $F^{\leftarrow}(t) = 1 - (1-t)^{1/\beta}$, $t \in (0,1)$, the following identity.

Corollary 2.3.10. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a reflected power function distribution RPower(β), $\beta > 0$. Then,

$$X_{r:m:n} \stackrel{\mathsf{d}}{=} 1 - \prod_{j=1}^{r} U_j^{1/(\beta\gamma_j)}, \quad 1 \le r \le m.$$

For $\beta = 1$, this yields the representation for the uniform distribution, i.e.,

$$U_{r:m:n} \stackrel{\mathsf{d}}{=} 1 - \prod_{j=1}^{r} U_j^{1/\gamma_j}, \quad 1 \le r \le m. \tag{2.16}$$

Using this representation, we can easily derive the following result. For $\beta = 1$, it can be found in Balakrishnan and Aggarwala [86].

Corollary 2.3.11. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a reflected power function distribution RPower(β). Then, with $X_{0:m:n} = 0$,

$$V_j = \left(\frac{1 - X_{j:m:n}}{1 - X_{j-1:m:n}}\right)^{\beta}, \quad 1 \le j \le m,$$

are independent random variables with $V_j \stackrel{\mathsf{d}}{=} U_j^{1/\gamma_j} \sim$ Beta$(\gamma_j, 1)$, $1 \le j \le m$.

For $\beta = 1$, we have, with $U_{0:m:n} = 0$,

$$V_j = \frac{1 - U_{j:m:n}}{1 - U_{j-1:m:n}}, \quad 1 \le j \le m,$$

to be independent random variables with $V_j \stackrel{\mathsf{d}}{=} U_j^{1/\gamma_j} \sim$ Beta$(\gamma_j, 1)$, $1 \le j \le m$.

In the uniform case, this yields the following well-known result of Malmquist [633].

Corollary 2.3.12. Let $U_{1:n}, \dots, U_{n:n}$ be order statistics from a uniform distribution. Then, with $U_{0:n} = 0$,

$$V_j = \frac{1 - U_{j:n}}{1 - U_{j-1:n}}, \quad 1 \le j \le n,$$

are independent random variables with $V_j \sim$ Beta$(n - j + 1, 1)$, $1 \le j \le n$.

2.3.3 Pareto Distributions

For Pareto distributions Pareto(α), we find with $F^{\leftarrow}(t) = (1-t)^{-1/\alpha}$, $t \in (0,1)$, the following result.

Corollary 2.3.13. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a Pareto distribution Pareto(α), $\alpha > 0$. Then,

$$X_{r:m:n} \stackrel{\mathrm{d}}{=} \prod_{j=1}^{r} U_j^{-1/(\alpha\gamma_j)}, \quad 1 \le r \le m.$$

Using this representation we can easily derive the following result (see also Balakrishnan and Aggarwala [86, p. 24]).

Corollary 2.3.14. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a Pareto distribution Pareto(α). Then, with $X_{0:m:n} = 1$,

$$W_j = \frac{X_{j:m:n}^{\alpha}}{X_{j-1:m:n}^{\alpha}}, \quad 1 \le j \le m,$$

are independent random variables with $W_j \stackrel{\mathrm{d}}{=} U_j^{-1/\gamma_j} \sim$ Pareto(γ_j), $1 \le j \le m$.

In the case of order statistics, this yields a well-known result for Pareto distributions which was first mentioned by Malik [632]. Further references are Huang [458], Arnold [49], and Johnson et al. [483].

2.3.4 Progressive Withdrawal and Dual Generalized Order Statistics

It is a well-known property of order statistics, $X_{j:n}$, $1 \le j \le n$, from a symmetric distribution (symmetric about 0), that

$$X_{j:n} \stackrel{\mathrm{d}}{=} -X_{n-j+1:n}, \quad 1 \le j \le n, \tag{2.17}$$

or that, jointly,

$$(X_{1:n}, \dots, X_{n:n}) \stackrel{\mathrm{d}}{=} (-X_{n:n}, \dots, -X_{1:n}); \tag{2.18}$$

see, for example, David and Nagaraja [327] and Arnold et al. [58]. Thus, the negatives of the order statistics are once again distributed as order statistics from the same symmetric distribution. It will therefore be natural to see whether a similar connection holds for progressively Type-II censored order statistics from a symmetric distribution as it will facilitate the handling of these random variables (see, e.g., Sect. 7.4). In the case of order statistics, the result in (2.18) is easily observed by considering the joint density function of order statistics given in Example 2.1.4. Using the representation in (2.5) and the fact that $f(x) = f(-x)$, $x \in \mathbb{R}$, the joint density function can be rewritten in the desired form. A similar argument for the identity (2.17) using the marginal density function has been employed in Balakrishnan and Aggarwala [86]. Alternatively, we may use the quantile representation of order statistics given in Theorem 2.1.1. Using the identity $F(x) = 1 - F(-x)$, $x \in \mathbb{R}$, for the cumulative distribution function for symmetric distributions, we get $F^{\leftarrow}(t) = -F^{\leftarrow}(1-t)$, $t \in (0,1)$. This implies for $1 \le r \le n$

$$-X_{r:n} \stackrel{\mathsf{d}}{=} -F^{\leftarrow}(U_{r:n}) = F^{\leftarrow}(1-U_{r:n}) \stackrel{\mathsf{d}}{=} F^{\leftarrow}(U_{n-r+1:n}) \stackrel{\mathsf{d}}{=} X_{n-r+1:n},$$

where we have used Corollary 2.3.8 and that $1 - X \sim \mathsf{Beta}(\beta, \alpha)$ holds for a $\mathsf{Beta}(\alpha, \beta)$-distributed random variable X.

However, in working with the progressively Type-II censored order statistics, we begin with the joint distribution of all m progressively Type-II censored order statistics. As before, let $X_{1:m:n}^{\mathscr{R}}, \ldots, X_{m:m:n}^{\mathscr{R}}$ denote the sample of progressively Type-II censored order statistics of size m obtained from a random sample of size n with censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$ from a symmetric distribution. Multiplying each random variable by -1 we get the decreasingly ordered sample

$$Y_1 = -X_{m:m:n}^{\mathscr{R}}, \ldots, Y_m = -X_{1:m:n}^{\mathscr{R}}.$$

It follows from the quantile representation in Theorem 2.3.6 and the quantile function

$$F_{-X}^{\leftarrow}(t) = -F^{\leftarrow}(1-t), \quad t \in (0,1),$$

that

$$Y_r \stackrel{\mathsf{d}}{=} -F^{\leftarrow}\Big(1 - \prod_{j=1}^{m-r+1} U_j^{1/\gamma_j}\Big) = F_{-X}^{\leftarrow}\Big(\prod_{j=1}^{m-r+1} U_j^{1/\gamma_j}\Big), \quad 1 \le r \le m.$$

This representation tells us that $Y_1, \ldots, Y_m$ are connected to the so-called dual generalized order statistics introduced by Burkschat et al. [234]. Moreover, we get the following expression for the joint density function:

$$f^{Y_1,\ldots,Y_m}(\mathbf{t}_m) = \prod_{j=1}^{m} \left[\gamma_j f(t_j) F^{R_j}(t_j)\right], \quad t_1 \le \cdots \le t_m.$$

As a result, the negatives of the progressively Type-II censored order statistics are generally not jointly distributed as progressively Type-II censored order statistics. Further results and applications can be found in Balakrishnan and Aggarwala [86, p. 71–81] and Burkschat et al. [234].

2.4 Marginal Distributions

Using the results of the preceding sections, explicit representations for the marginal distributions of progressively Type-II censored order statistics can be established. First, we notice that a right censored progressively Type-II censored sample can be seen as progressively Type-II censored order statistics from the same distribution with a modified censoring scheme. Thus, right censored progressively censored samples can always be seen as a complete progressively censored sample with a modified censoring scheme. In particular, we have the following result (see Balakrishnan and Aggarwala [86]).

Theorem 2.4.1. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a cumulative distribution function F with censoring scheme $\mathscr{R} = (R_1, \dots, R_m)$.

Then, for $1 \le r \le m$, the right censored sample $X_{1:m:n}^{\mathscr{R}}, \dots, X_{r:m:n}^{\mathscr{R}}$ can be seen as a complete sample of progressively Type-II censored order statistics $X_{1:r:n}^{\mathscr{R}_r}, \dots, X_{r:r:n}^{\mathscr{R}_r}$ from the same population with censoring scheme $\mathscr{R}_r = (R_1, \dots, R_{r-1}, \gamma_r - 1)$.

Proof. The iterative construction of progressively Type-II censored order statistics presented in Procedure 1.1.3 yields directly the above property. The only property that has to be shown is the particular structure of the censoring scheme. But, according to the construction process 1.1.3, γ_r denotes the number of items in the experiment before the rth failure. Thus, stopping the experiment after the rth failure is equivalent to removing the remaining $\gamma_r - 1$ units.

Alternatively, the iterative construction in Theorem 2.3.6 can be used for this purpose. □

Thus, right censoring of progressively Type-II censored samples results in the same model with a reduced number of observations and a modified censoring scheme. In particular, we can apply the preceding results and obtain, for example, the joint marginal density function of $X_{1:m:n}, \dots, X_{r:m:n}$ as

$$f^{X_{1:m:n},\dots,X_{r:m:n}}(\mathbf{x}_r) = \prod_{j=1}^{r-1} \left[\gamma_j f(x_j)(1-F(x_j))^{R_j}\right] \gamma_r f(x_r)(1-F(x_r))^{\gamma_r - 1},$$

$$x_1 \le \dots \le x_r.$$

In particular, Theorem 2.4.1 illustrates that any progressively Type-II censored order statistic can be seen as a maximal progressively Type-II censored order statistic with an appropriately chosen censoring scheme.

In order to calculate the marginal distributions, we consider the exponential case first. The presentation of the following results uses the notation proposed in Kamps and Cramer [503]. An alternative but equivalent representation has been established in Balakrishnan et al. [132] using the integral identity (2.31) (see also Balakrishnan [84] and Nagaraja [667]). This representation has also been exploited in many papers.

2.4.1 Exponential Distribution

The marginal distributions of exponential progressively Type-II censored order statistics can be derived using the sum representation (2.13). This expression shows that we are interested in finding distributions of sums of independent but not necessarily identically distributed exponential random variables. An important point in the following derivations is that the γ's cannot be equal. Such problems have been considered earlier by, e.g., Likeš [597] and Kamps [497] (see also Johnson et al. [483, p. 552]). This type of distribution is called hyperexponential distribution or generalized Erlang distribution (see Johnson and Kotz [482, p. 222]). A review on this topic including various applications of hyperexponential distributions is provided by Botta et al. [217]. This yields directly the following result.

Theorem 2.4.2 (Kamps and Cramer [503]). Let $Z_{1:m:n}, \dots, Z_{m:m:n}$ be standard exponential progressively Type-II censored order statistics. Then,

$$F^{Z_{r:m:n}}(t) = 1 - \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} \frac{1}{\gamma_j} a_{j,r} \mathrm{e}^{-\gamma_j t}, \quad t > 0, \tag{2.19}$$

where $a_{j,r} = \prod_{\substack{i=1\\ i\neq j}}^{r} \frac{1}{\gamma_i - \gamma_j}$, $1 \le j \le r \le n$. The density function of $Z_{r:m:n}$ is given by

$$f^{Z_{r:m:n}}(t) = \Big(\prod_{j=1}^{r} \gamma_j\Big) \sum_{j=1}^{r} a_{j,r} \mathrm{e}^{-\gamma_j t}, \quad t > 0. \tag{2.20}$$

Schenk [782] considered multiply censored samples of progressively Type-II censored order statistics (see also Cramer [287]). He derived expressions for the corresponding density functions. His derivations are based on the Markov property of the thinned sample $Z_{k_1:m:n}, \dots, Z_{k_\ell:m:n}$ with $1 \le k_1 < k_2 < \dots < k_\ell \le m$ (see Sect. 2.5.1). Using the sum representation (2.13), we find

$$Z_{k_2:m:n} = \sum_{j=1}^{k_2} \frac{1}{\gamma_j} S_j^{\mathscr{R}} = Z_{k_1:m:n} + \sum_{j=k_1+1}^{k_2} \frac{1}{\gamma_j} S_j^{\mathscr{R}}.$$

Thus, the cumulative distribution function of $Z_{j_2:m:n}$, given $Z_{j_1:m:n} = s$, is given by $P\big(\sum_{j=k_1+1}^{k_2} \frac{1}{\gamma_j} S_j^{\mathscr{R}} \le t - s\big)$, $t \ge s$. Hence, the density function follows from (2.20) as

$$f^{Z_{k_2:m:n}|Z_{k_1:m:n}}(t|s) = \Big(\prod_{j=k_1+1}^{k_2} \gamma_j\Big) \sum_{j=k_1+1}^{k_2} a_{j,k_2}^{(k_1)} \mathrm{e}^{-\gamma_j(t-s)}, \quad t > s > 0,$$

where $a_{j,k_2}^{(k_1)} = \prod_{\substack{\nu=k_1+1\\ \nu\neq j}}^{k_2} \frac{1}{\gamma_\nu - \gamma_j}$. Combining these expressions, we arrive at the joint density function of two exponential progressively Type-II censored order statistics given in Kamps and Cramer [503] ($t > s > 0$):

$$\begin{aligned} f^{Z_{k_1:m:n},Z_{k_2:m:n}}(s,t) &= f^{Z_{k_1:m:n}}(s) f^{Z_{k_2:m:n}|Z_{k_1:m:n}}(t|s) \\ &= \Big(\prod_{j=1}^{k_2} \gamma_j\Big) \sum_{i=1}^{k_1} \sum_{j=k_1+1}^{k_2} a_{i,k_1} a_{j,k_2}^{(k_1)} \mathrm{e}^{-\gamma_j(t-s)} \mathrm{e}^{-\gamma_i s}. \end{aligned}$$

A repeated application of the preceding result yields the joint density function of the multiply censored sample:

$$\begin{aligned} & f^{Z_{k_1:m:n},\ldots,Z_{k_\ell:m:n}}(x_{k_1},\ldots,x_{k_\ell}) \\ &\quad = \prod_{i=1}^{\ell} \Big[\Big(\prod_{j=k_{i-1}+1}^{k_i} \gamma_j\Big) \sum_{j=k_{i-1}+1}^{k_i} a_{j,k_i}^{(k_{i-1})} \mathrm{e}^{-\gamma_j(x_{k_i} - x_{k_{i-1}})}\Big], \end{aligned} \tag{2.21}$$

where $k_0 = 0$, $0 = x_0 \le x_{k_1} \le \cdots \le x_{k_\ell}$, and

$$a_{j,k_i}^{(k_{i-1})} = \prod_{\substack{\nu=k_{i-1}+1\\ \nu\neq j}}^{k_i} \frac{1}{\gamma_\nu - \gamma_j}, \quad k_{i-1} + 1 \le j \le k_i, 1 \le i \le \ell. \tag{2.22}$$

This result can be directly applied to a general progressively Type-II censored sample $X_{r+1:m:n}^{\mathscr{R}_{\lhd r}}, \ldots, X_{m:m:n}^{\mathscr{R}_{\lhd r}}$ with the left truncated censoring scheme $\mathscr{R}_{\lhd r} = (R_{r+1}, \ldots, R_m) \in \mathscr{C}_{m-r,n-r}^{m-r}$. The corresponding density function is given by

$$f^{Z_{r+1:m:n},\dots,Z_{m:m:n}}(x_{r+1},\dots,x_m)$$
$$= \binom{n}{r}\Big(\prod_{j=r+1}^{m}\gamma_j\Big)(1-\mathrm{e}^{-x_{r+1}})^r \exp\Big\{-\sum_{j=r+1}^{m}(R_j+1)x_j\Big\},$$
$$0 \le x_{r+1} \le \cdots \le x_m. \tag{2.23}$$

2.4.2 Uniform Distribution

Due to its importance, it is useful to present the representations of the marginal density functions and cumulative distribution functions for uniform progressively Type-II censored order statistics. They can be taken directly from Theorem 2.4.2 using a quantile transformation.

Corollary 2.4.3. Let $U_{1:m:n},\dots,U_{m:m:n}$ be uniform progressively Type-II censored order statistics. Then, for $1 \le r \le m$,

$$F^{U_{r:m:n}}(t) = 1 - \Big(\prod_{i=1}^{r}\gamma_i\Big)\sum_{j=1}^{r}\frac{1}{\gamma_j}a_{j,r}(1-t)^{\gamma_j}, \quad t \in [0,1].$$

The density function of $U_{r:m:n}$ is given by

$$f^{U_{r:m:n}}(t) = \Big(\prod_{j=1}^{r}\gamma_j\Big)\sum_{j=1}^{r}a_{j,r}(1-t)^{\gamma_j-1}, \quad t \in [0,1]. \tag{2.24}$$

2.4.3 General Distributions

Using the quantile transformation result 2.1.1, the preceding results can be directly applied to arbitrary distributions. For brevity, we present only the expressions in the univariate and bivariate case. From Theorem 2.4.2, we obtain the following result.

Corollary 2.4.4. Let $X_{1:m:n},\dots,X_{m:m:n}$ be progressively Type-II censored order statistics from a cumulative distribution function F. Then, for $1 \le r \le m$,

$$F^{X_{r:m:n}}(t) = 1 - \Big(\prod_{i=1}^{r}\gamma_i\Big)\sum_{j=1}^{r}\frac{1}{\gamma_j}a_{j,r}(1-F(t))^{\gamma_j}, \quad t \in \mathbb{R}. \tag{2.25}$$

From (2.25), we find with $t \to -\infty$ the identity

$$1 = \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} \frac{1}{\gamma_j} a_{j,r}. \tag{2.26}$$

Applying this identity and writing $F_{1:\gamma_j} = 1 - (1-F)^{\gamma_j}$, we get a representation of the cumulative distribution function in terms of distributions of minima as

$$F^{X_{r:m:n}}(t) = \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} \frac{1}{\gamma_j} a_{j,r} F_{1:\gamma_j}(t). \tag{2.27}$$

Noticing that for order statistics the identity

$$\Big(\prod_{i=1}^{r} \gamma_i\Big) \frac{1}{\gamma_j} a_{j,r} = (-1)^{r-j} \binom{n}{j-1} \binom{n-j}{r-j}$$

holds, we find

$$F_{r:n}(t) = \sum_{j=n-r+1}^{n} (-1)^{j-n+r-1} \binom{j-1}{n-r} \binom{n}{j} F_{1:j}(t).$$

This identity is given, for instance, in David and Nagaraja [327, p. 46] and Arnold et al. [58, p. 113] in terms of moments. For moments of order statistics, this result is due to Srikantan [822].

For absolutely continuous distributions, we have the following representation of the density function.

Corollary 2.4.5. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from an absolutely continuous cumulative distribution function F with density function f. Then, for $1 \le r \le m$,

$$f^{X_{r:m:n}}(t) = f(t) \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} a_{j,r} (1 - F(t))^{\gamma_j - 1}, \quad t \in \mathbb{R}. \tag{2.28}$$

For $1 \le k_1 < k_2 \le m$ and $t > s$, the bivariate density function is given by

$$f^{X_{k_1:m:n}, X_{k_2:m:n}}(s,t) = f(s) f(t) \Big(\prod_{j=1}^{k_2} \gamma_j\Big) \sum_{i=1}^{k_1} \sum_{j=k_1+1}^{k_2} a_{i,k_1} a_{j,k_2}^{(k_1)} \Big[\frac{1-F(t)}{1-F(s)}\Big]^{\gamma_j - 1} (1 - F(s))^{\gamma_i - 2}. \tag{2.29}$$

Remark 2.4.6. For order statistics, the representations given above simplify to the well-known expressions

$$F_{r:n}(t) = \sum_{j=r}^{n} \binom{n}{j} F^{j}(t)(1-F(t))^{n-j}, \quad t \in \mathbb{R},$$
$$f_{r:n}(t) = r\binom{n}{r} F^{r-1}(t)(1-F(t))^{n-r} f(t), \quad t \in \mathbb{R} \tag{2.30}$$

(see, e.g., Arnold et al. [58], David and Nagaraja [327]).

For the cumulative distribution function and $x_1 \le x_2$, the expression (2.29) yields by integration the following representation which holds for any baseline cumulative distribution function F:

$$F^{X_{k_1:m:n},X_{k_2:m:n}}(x_1,x_2) = F^{X_{k_1:m:n}}(x_1) - \Big(\prod_{j=1}^{k_2}\gamma_j\Big)\sum_{i=1}^{k_1}\sum_{j=k_1+1}^{k_2}\frac{a_{i,k_1}a_{j,k_2}^{(k_1)}}{\gamma_j(\gamma_i-\gamma_j)}\overline{F}^{\gamma_j}(x_2)\Big[1-\overline{F}^{\gamma_i-\gamma_j}(x_1)\Big].$$

The result can be established by using the relation

$$F^{X_{k_1:m:n},X_{k_2:m:n}}(x_1,x_2) = F^{X_{k_1:m:n}}(x_1) - P(X_{k_1:m:n} \le x_1, X_{k_2:m:n} > x_2).$$

Assuming uniform progressively Type-II censored order statistics and using (2.29), the probability on the right-hand side reads

$$c_{k_2-1}\sum_{i=1}^{k_1}\sum_{j=k_1+1}^{k_2} a_{i,k_1}a_{j,k_2}^{(k_1)}\int_0^{x_1}\int_{x_2}^{1}(1-t)^{\gamma_i-\gamma_j-1}(1-s)^{\gamma_j-1}ds\,dt$$
$$= c_{k_2-1}\sum_{i=1}^{k_1}\sum_{j=k_1+1}^{k_2}\frac{a_{i,k_1}a_{j,k_2}^{(k_1)}}{\gamma_j(\gamma_i-\gamma_j)}(1-x_2)^{\gamma_j}\Big[1-(1-x_1)^{\gamma_i-\gamma_j}\Big].$$

Using the quantile transformation and (2.2), we arrive at the desired representation. Notice that, for $x_2 < x_1$, one has $F^{X_{k_1:m:n},X_{k_2:m:n}}(x_1,x_2) = F^{X_{k_2:m:n}}(x_2)$.

Multiply Censored Progressively Type-II Censored Order Statistics

Similar results can be established for multiply censored samples (see (2.21) and, for generalized order statistics, Cramer [287]). The joint density function of progressively Type-II censored order statistics $X_{k_1:m:n},\ldots,X_{k_\ell:m:n}$ is given by

$$f^{X_{k_1:m:n},\dots,X_{k_\ell:m:n}}(x_{k_1},\dots,x_{k_\ell})$$
$$= \prod_{i=1}^{\ell}\Big[\frac{f(x_{k_i})}{1-F(x_{k_i})}\Big(\prod_{j=k_{i-1}+1}^{k_i}\gamma_j\Big)\sum_{j=k_{i-1}+1}^{k_i} a_{j,k_i}^{(k_{i-1})}\Big(\frac{1-F(x_{k_i})}{1-F(x_{k_{i-1}})}\Big)^{\gamma_j}\Big],$$

where $k_0 = 0$, $x_0 = -\infty$, $x_{k_1} \le \cdots \le x_{k_\ell}$. For order statistics, the expression simplifies to that given in Kong [543].

In the special case of a general progressively Type-II censored sample $X_{r+1:m:n}^{\mathscr{R}_{\lhd r}},\dots,X_{m:m:n}^{\mathscr{R}_{\lhd r}}$ with the left truncated censoring scheme $\mathscr{R}_{\lhd r} = (R_{r+1},\dots,R_m) \in \mathscr{C}_{m-r,n-r}^{m-r}$, the corresponding joint density function is given by ([see 86, p. 10] and (2.21) for the exponential distribution)

$$f(x_{r+1},\dots,x_m) = \binom{n}{r} F^r(x_{r+1})\Big(\prod_{j=r+1}^{m}\gamma_j f(x_j)(1-F(x_j))^{R_j}\Big),$$
$$x_{r+1} \le \cdots \le x_m.$$

An Important Recurrence Relation

The preceding results yield the following connection between cumulative distribution functions and density functions.

Corollary 2.4.7. For $r \in \{1,\dots,m-1\}$,

$$F^{X_{r:m:n}}(t) - F^{X_{r+1:m:n}}(t) = \frac{1}{\gamma_{r+1}}(1-F(t))f^{U_{r+1:m:n}}(F(t)), \quad t \in \mathbb{R}.$$

Proof. From (2.25), we obtain

$$F^{X_{r:m:n}}(t) = 1 - \Big(\prod_{i=1}^{r}\gamma_i\Big)\sum_{j=1}^{r}\frac{1}{\gamma_j}a_{j,r}(1-F(t))^{\gamma_j}.$$

Since $(\gamma_{r+1}-\gamma_j)a_{j,r+1} = a_{j,r}$, $1 \le j \le r$, we get

$$F^{X_{r:m:n}}(t) - F^{X_{r+1:m:n}}(t)$$
$$= \Big(\prod_{i=1}^{r}\gamma_i\Big)\Big(a_{r+1,r+1}(1-F(t))^{\gamma_{r+1}} + \sum_{j=1}^{r}a_{j,r+1}(1-F(t))^{\gamma_j}\Big)$$
$$= \frac{1}{\gamma_{r+1}}(1-F(t))f^{U_{r+1:m:n}}(F(t)), \quad t \in \mathbb{R},$$

which yields the desired result. □

For order statistics, the above relation simplifies to

$$F^{X_{r:n}}(t) = F^{X_{r+1:n}}(t) + \frac{1}{n-r+1}(1-F(t))f^{U_{r+1:n}}(F(t))$$
$$= F^{X_{r+1:n}}(t) + \binom{n}{r}F^r(t)(1-F(t))^{n-r},$$

which can be found in David and Shu [328].

An Alternative Approach to Derive the Marginals

Balakrishnan et al. [132] presented an alternative approach to derive the marginals of progressively Type-II censored order statistics. They tackled the problem through an explicit evaluation of the resulting integrals as shown in Lemma 1 in Balakrishnan et al. [132]. Using the notation introduced in (2.22), a version of this lemma adapted to the present setting is as follows.

Lemma 2.4.8. Let $\mathscr{R} = (R_1, \ldots, R_{r+1})$ be a censoring scheme, $r \geq k+1 \geq 1$. Then, for a cumulative distribution function F with density function f and $t \leq y$, the following identity holds:

$$\int_t^y \int_t^{x_r} \cdots \int_t^{x_{k+2}} \prod_{i=k+1}^{r} \Big[f(x_i)\overline{F}^{R_i}(x_i)\Big] dx_{k+1} \ldots dx_r$$
$$= \sum_{j=k+1}^{r+1} a_{j,r+1}^{(k)} \overline{F}(t)^{\gamma_{k+1}-\gamma_j} \overline{F}(y)^{\gamma_j-\gamma_{r+1}}. \quad (2.31)$$

For $y \to \infty$, we get

$$\int_t^\infty \int_t^{x_r} \cdots \int_t^{x_{k+2}} \prod_{i=k+1}^{r} \Big[f(x_i)\overline{F}^{R_i}(x_i)\Big] dx_{k+1} \ldots dx_r$$
$$= a_{r+1,r+1}^{(k)} \overline{F}(t)^{\gamma_{k+1}-\gamma_{r+1}}.$$

This integral representation will be helpful in many settings. For instance, it will be used later to derive the power function of precedence-type tests under Lehmann alternative [see (21.9)].

Connection of Marginals to Interpolation Polynomials

Cramer [289] established a connection of one-dimensional marginal density functions and cumulative distribution functions to divided differences and Lagrangian interpolation polynomials (see, e.g., Neumaier [679]). For $t \geq 0$, let

$$h_t : [0,\infty) \longrightarrow [0,\infty) \quad \text{be defined by } h_t(x) = t^x, \quad x \geq 0. \tag{2.32}$$

Then, the divided differences $h_t[x_j,\dots,x_\nu]$ of order $\nu - j$ at $x_1 > \cdots > x_m$ are defined by

$$\begin{aligned} h_t[x_j] &= h_t(x_j),\\ h_t[x_j,\dots,x_\nu] &= \frac{h_t[x_{j+1},\dots,x_\nu] - h_t[x_j,\dots,x_{\nu-1}]}{x_\nu - x_j}, \end{aligned}$$

for $1 \leq j < \nu \leq m$. Then,

$$f^{U_{r:m:n}}(t) = (-1)^{r-1} h_{1-t}[\gamma_1 - 1,\dots,\gamma_r - 1], \quad t \in [0,1].$$

Cramer [289] showed that the survival function of $X_{r:m:n}$ can be written as a specific Lagrangian interpolation polynomial $\mathcal{P}_r^{F(t)}$ evaluated at the point zero

$$\begin{aligned} \mathcal{P}_r^{F(t)}(0) &= 1 - F^{X_{r:m:n}}(t) = \sum_{j=1}^{r} \frac{\prod\limits_{\nu=1,\nu\neq j}^{r} (0-\gamma_\nu)}{\prod\limits_{\nu=1,\nu\neq j}^{r} (\gamma_j - \gamma_\nu)} \overline{F}^{\gamma_j}(t)\\ &= \sum_{j=1}^{r} \frac{\prod\limits_{\nu=1,\nu\neq j}^{r} (0-\gamma_\nu)}{\prod\limits_{\nu=1,\nu\neq j}^{r} (\gamma_j - \gamma_\nu)} h_{1-F(t)}(\gamma_j). \end{aligned}$$

Thus, $\mathcal{P}_r^{F(t)}$ interpolates the function $h_{1-F(t)}$ given in (2.32) at the points $\gamma_1,\dots,\gamma_r$. Precisely, the evaluation of the polynomial at zero is an extrapolation since zero does not belong to the range of the γ_j's. In particular, this shows that the cumulative distribution function of a progressively Type-II censored order statistic can be understood as a Lagrangian interpolation polynomial.

2.5 Conditional Distributions

2.5.1 *Markov Property*

From the joint density function of progressively Type-II censored order statistics given in (2.4), the Markov property can be easily derived for an absolutely continuous cumulative distribution function F. However, this property holds even for continuous cumulative distribution function.

Theorem 2.5.1. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F.

Then, $X_{1:m:n}, \dots, X_{m:m:n}$ form a Markov chain with transition probabilities $(2 \le r \le m)$

$$P(X_{r:m:n} \le t | X_{r-1:m:n} = s) = 1 - \left(\frac{1-F(t)}{1-F(s)}\right)^{\gamma_r}, \quad s \le t \text{ with } F(s) < 1.$$

Proof. First, we consider the uniform distribution. According to representation (2.16), we find

$$U_{r:m:n} = 1 - U_r^{1/\gamma_r}(1 - U_{r-1:m:n}), \quad 2 \le r \le m. \tag{2.33}$$

Using the independence of U_j, $1 \le j \le m$, the progressively Type-II censored order statistics $U_{1:m:n}, \dots, U_{r:m:n}$ form a Markov chain with $(s \le t < 1)$

$$P(U_{r:m:n} \le t | U_{r-1:m:n} = s) = P\left(U_r^{1/\gamma_r} \ge \frac{1-t}{1-s}\right) = 1 - \left(\frac{1-t}{1-s}\right)^{\gamma_r}.$$

Using the properties of the quantile function in Lemma A.2.2 and Theorem 2.3.6, we obtain from the continuity of F and (2.33) that

$$\begin{aligned} X_{r:m:n} &= F^{\leftarrow}(U_{r:m:n}) = F^{\leftarrow}(1 - U_r^{1/\gamma_r}[1 - U_{r-1:m:n}]) \\ &= F^{\leftarrow}(1 - U_r^{1/\gamma_r}[1 - F(F^{\leftarrow}(U_{r-1:m:n}))]) \\ &= F^{\leftarrow}(1 - U_r^{1/\gamma_r}[1 - F(X_{r-1:m:n})]). \end{aligned} \tag{2.34}$$

Therefore, the independence of U_r and $X_{1:m:n}, \dots, X_{r-1:m:n}$ yields for $s_1 \le \dots \le s_{r-1} \le t$ with $F(s_{r-1}) < 1$, the conditional cumulative distribution function

$$\begin{aligned} &P(X_{r:m:n} \le t | X_{j:m:n} = s_j, j = 1, \dots, r-1) \\ &\quad = P(F^{\leftarrow}(1 - U_r^{1/\gamma_r}[1 - F(s_{r-1})]) \le t) = P\left(U_r^{1/\gamma_r} \ge \frac{1-F(t)}{1-F(s_{r-1})}\right). \end{aligned}$$

This proves the desired result. □

As proved by Balakrishnan and Dembińska [96], the above result does not hold for noncontinuous distributions (see also Tran [854]). More details on the noncontinuous case are provided in Sect. 2.8. For order statistics, the Markov property is a well-known property (see Arnold et al. [58] and David and Nagaraja [327]). Obviously, Theorem 2.5.1 can be extended to the following result (see Balakrishnan and Aggarwala [86, p. 15]).

Theorem 2.5.2. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F. Then, conditional on $X_{r-1:m:n} = x_{r-1}$, the random variables $X_{r:m:n}, \dots, X_{m:m:n}$ are progressively Type-II censored order statistics from a left truncated cumulative distribution function

$$G_{x_{r-1}}(y) = \frac{F(y) - F(x_r)}{1 - F(x_{r-1})}, \quad x_{r-1} \le y, F(x_{r-1}) < 1, \tag{2.35}$$

and left (truncated) censoring scheme $\mathscr{R}_{\lhd r-1} = (R_r, \dots, R_m)$.

Proof. Applying (2.34) to the random variables $X_{r:m:n}, \dots, X_{m:m:n}$, we find

$$\begin{aligned} X_{r:m:n} &= F^{\leftarrow}(1 - U_r^{1/\gamma_r}[1 - F(X_{r-1:m:n})]), \\ &\vdots \\ X_{m:m:n} &= F^{\leftarrow}\Big(1 - \prod_{j=r}^{m} U_j^{1/\gamma_j}[1 - F(X_{r-1:m:n})]\Big). \end{aligned}$$

Since the quantile function of the truncated cumulative distribution function in (2.35) is given by

$$G^{\leftarrow}_{x_{r-1}}(t) = F^{\leftarrow}(1 - (1-t)(1 - F(x_{r-1}))), \quad t \in (0,1),$$

we find for $\ell = r, \dots, m$,

$$X_{\ell:m:n} = F^{\leftarrow}\Big(1 - \prod_{j=r}^{\ell} U_j^{1/\gamma_j}[1 - F(X_{r-1:m:n})]\Big) = G^{\leftarrow}_{X_{r-1:m:n}}\Big(1 - \prod_{j=r}^{\ell} U_j^{1/\gamma_j}\Big).$$

Therefore, conditional on $X_{r-1:m:n} = x_{r-1}$, Theorem 2.3.6 yields the assertion. Notice that the parameters $\gamma_r, \dots, \gamma_m$ yield the left truncated censoring scheme $\mathscr{R}_{\lhd r-1} = (R_r, \dots, R_m)$. □

For order statistics, the corresponding result is due to Scheffé and Tukey [781] (see also Arnold et al. [58, p. 25–26] and David and Nagaraja [327]). Horn and Schlipf [450] utilized this representation to develop an efficient algorithm for the generation of Type-II doubly censored data.

The Markov property yields the following factorization of the density function of $\mathbf{X}^{\mathscr{R}}$. For $2 \leq r \leq m$, the conditional density function $f_{r|r-1:m:n}(\cdot|s)$ of $X_{r:m:n}$, given $X_{r-1:m:n} = s$, is defined by

$$f_{r|r-1:m:n}(t|s) = \begin{cases} \gamma_r \frac{f(t)}{1-F(s)} \cdot \left(\frac{1-F(t)}{1-F(s)}\right)^{\gamma_r - 1}, & s \leq t \text{ with } F(s) < 1, \\ f_{r:m:n}(t), & \text{otherwise.} \end{cases}$$

Of course, $f_{1|0:m:n} = f_{1:m:n}$ is the marginal density function of $X_{1:m:n}$.

Corollary 2.5.3. The density function of a progressively Type-II censored sample $\mathbf{X}^{\mathscr{R}}$ factorizes as follows:

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}_m) = \prod_{r=1}^{m} f_{r|r-1:m:n}(x_r|x_{r-1}), \quad x_1 \leq \cdots \leq x_m. \tag{2.36}$$

According to Cramer [287], a conditional P^F-density function exists provided that the population cumulative distribution function F is continuous. It is given by

$$f_{r|r-1:m:n}(t|s) = \frac{\gamma_r}{1-F(s)} \cdot \left(\frac{1-F(t)}{1-F(s)}\right)^{\gamma_r - 1} \quad P^F \text{a.e.} \tag{2.37}$$

2.5.2 Distributions of Generalized Spacings

The preceding results can be used to calculate the density functions of generalized spacings (so-called subranges or contrasts), i.e., of the (r, s)-spacing

$$S_{r,s}^{\star\mathscr{R}} = X_{r:m:n} - X_{s:m:n}, \quad 1 \leq s < r \leq m. \tag{2.38}$$

From Lemma 3 of Kamps and Cramer [503], we get the following expressions:

$$\begin{aligned} f^{S_{r,s}^{\star\mathscr{R}}}(w) &= \prod_{j=1}^{s} \gamma_j \int_{\mathbb{R}} \Big(\sum_{i=r+1}^{s} a_{i,s}^{(r)} \Big(\frac{1-F(v+w)}{1-F(v)}\Big)^{\gamma_i} \Big) \\ &\quad \times \Big(\sum_{i=1}^{r} a_{i,r} (1-F(v))^{\gamma_i} \Big) \frac{f(v)}{1-F(v)} \, \frac{f(v+w)}{1-F(v+w)} \, dv, \\ F^{S_{r,s}^{\star\mathscr{R}}}(w) &= 1 - \int_{\mathbb{R}} H_{r,s}\Big(\frac{1-F(v+w)}{1-F(v)}\Big) dF^{X_{r:m:n}}(v), \end{aligned} \tag{2.39}$$

where the function H is defined by $H_{r,s}(z) = \left(\prod_{j=r+1}^{s} \gamma_j\right) \sum_{i=r+1}^{s} a_{i,s}^{(r)} \frac{1}{\gamma_i} z^{\gamma_i}$, $z \in [0, 1]$ [see also (2.29)].

2.5.3 Block Independence of Progressively Type-II Censored Order Statistics

In this section, we study the distribution of a progressively Type-II censored sample randomly divided into two blocks by a threshold $T \in \mathbb{R}$.

Lemma 2.5.4. For a fixed time T, let D denote the number of progressively Type-II censored order statistics that do not exceed T, i.e.,

$$D = \sum_{j=1}^{m} \mathbb{1}_{(-\infty,T]}(X_{j:m:n}). \tag{2.40}$$

Then,

$$\begin{aligned}
P(D = 0) &= (1 - F(T))^n, \\
P(D = d) &= \Big(\prod_{i=1}^{d} \gamma_i\Big) \sum_{j=1}^{d+1} a_{j,d+1}(1 - F(T))^{\gamma_j}, \quad d = 1, \ldots, m-1, \\
P(D = m) &= F^{X_{m:m:n}}(T) = 1 - \Big(\prod_{i=1}^{m} \gamma_i\Big) \sum_{j=1}^{m} \frac{1}{\gamma_j} a_{j,m}(1 - F(T))^{\gamma_j}.
\end{aligned} \tag{2.41}$$

Proof. For $d = 0$, we have $P(D = 0) = P(X_{1:m:n} > T) = (1 - F(T))^n$.
Let $d \in \{1, \ldots, m-1\}$. From the definition of D, we have

$$\begin{aligned}
P(D = d) =& P(X_{d:m:n} \le T < X_{d+1:m:n}) \\
=& P(X_{d:m:n} \le T) - P(X_{d+1:m:n} \le T) = F^{X_{d:m:n}}(T) - F^{X_{d+1:m:n}}(T).
\end{aligned}$$

From Corollary 2.4.7, we conclude for $d \in \{1, \ldots, m-1\}$,

$$F^{X_{d:m:n}}(T) - F^{X_{d+1:m:n}}(T) = \frac{1}{\gamma_{d+1}}(1 - F(T)) f^{U_{d+1:m:n}} F(T)). \tag{2.42}$$

An application of (2.24), i.e.,

$$f^{U_{d+1:m:n}}(t) = \Big(\prod_{i=1}^{d+1} \gamma_i\Big) \sum_{j=1}^{d+1} a_{j,d+1}(1 - t)^{\gamma_j - 1} \quad t \in (0, 1),$$

proves the desired result. The case $d = m$ follows from $P(D = m) = P(X_{m:m:n} \le T) = F^{X_{m:m:n}}(T)$ and (2.25). □

In the case of order statistics, the distribution of D simplifies considerably. In particular, we have (see Iliopoulos and Balakrishnan [469]) that D has a binomial distribution with parameters n and $F(T)$, i.e.,

$$P(D = d) = \binom{n}{d} F^d(T)(1 - F(T))^{n-d}, \quad d = 0, \dots, n.$$

This follows directly from (2.42) in the above proof.

The following conditional independence result for progressively Type-II censored order statistics is due to Iliopoulos and Balakrishnan [469].

Theorem 2.5.5. Let $d \in \{1, \dots, m-1\}$, F be a cumulative distribution function, $\mathscr{R} = (R_1, \dots, R_m)$ be a censoring scheme, and $\mathbf{K}_d = (K_1, \dots, K_d)$ be a discrete random vector on the Cartesian product $\times_{j=1}^d \{0, \dots, R_j\}$ with probability mass function

$$p^{\mathbf{K}_d}(\mathbf{k}_d) = \frac{1}{P(D = d)} F^{\eta_1(d)}(T)(1 - F(T))^{n-\eta_1(d)} \prod_{i=1}^d \frac{\gamma_i}{\eta_i(d)} \binom{R_i}{k_i},$$

where $\eta_i(d) = \sum_{j=i}^d (k_j + 1)$, $1 \le i \le d$.

Conditional on $D = d$, the two random vectors $(X_{1:m:n}^{\mathscr{R}}, \dots, X_{d:m:n}^{\mathscr{R}})$ and $(X_{d+1:m:n}^{\mathscr{R}}, \dots, X_{m:m:n}^{\mathscr{R}})$ are independent with

$$(X_{1:m:n}^{\mathscr{R}}, \dots, X_{d:m:n}^{\mathscr{R}}) \stackrel{\mathsf{d}}{=} \left(V_{1:d:\kappa_d}^{\mathscr{K}}, \dots, V_{d:d:\kappa_d}^{\mathscr{K}}\right)$$

$$(X_{d+1:m:n}^{\mathscr{R}}, \dots, X_{m:m:n}^{\mathscr{R}}) \stackrel{\mathsf{d}}{=} \left(W_{1:m-d:\gamma_d}, \dots, W_{m-d:m-d:\gamma_d}\right)$$

where $\mathscr{K} = \mathbf{K}_d$, $\kappa_d = \sum_{j=1}^d (1 + K_j)$, and

(i) $V_1, \dots, V_n$ are IID random variables with the right truncated cumulative distribution function F_T given by

$$F_T(t) = \frac{F(t)}{F(T)}, \quad t \le T,$$

(ii) $W_1, \dots, W_{\gamma_d}$ are IID random variables with left truncated cumulative distribution function G_T

$$G_T(t) = 1 - \frac{1 - F(t)}{1 - F(T)}, \quad t \ge T.$$

Remark 2.5.6. Notice that the sample size κ_d for the progressively Type-II censored order statistics $V_{1:d:\kappa_d}^{\mathscr{K}}, \dots, V_{d:d:\kappa_d}^{\mathscr{K}}$ is a random variable. This representation means that the distribution of $(X_{1:m:n}^{\mathscr{R}}, \dots, X_{d:m:n}^{\mathscr{R}})$, given $D = d$, is a mixture of distributions of progressively Type-II censored order statistics with mixing distribution $p^{\mathscr{K}}$. It is well known that the right truncation of progressively

Type-II censored order statistics does not result in progressively Type-II censored order statistics from the corresponding right truncated distribution (see, for example, Balakrishnan and Aggarwala [86]). This is due to the fact that those observations (progressively) censored before T could have values larger than T.

Proof of Theorem 2.5.5. First, notice that we can restrict ourselves to the uniform distribution replacing T by $F(T)$. Then, notice that the conditional density function of $\mathbf{U}^{\mathscr{R}} = (U_{1:m:n}, \ldots, U_{m:m:n})$, given $D = d$, is given by

$$f^{\mathbf{U}^{\mathscr{R}}|D=d}(\mathbf{x}_m) = \frac{1}{P(D=d)} f^{\mathbf{U}^{\mathscr{R}}}(\mathbf{x}_m) \mathbb{1}_{[x_d, x_{d+1})}(F(T)).$$

Decomposing the joint density function of $U_{1:m:n}, \ldots, U_{m:m:n}$, we obtain

$$\begin{aligned} f^{\mathbf{U}^{\mathscr{R}}}(\mathbf{x}_m) &= \prod_{i=1}^m \gamma_i \prod_{j=1}^m (1-x_i)^{R_i} \\ &= \left[\prod_{i=1}^d \gamma_i \prod_{i=1}^d (1-x_i)^{R_i}\right] \times \left[\prod_{i=d+1}^m \gamma_i \prod_{i=d+1}^m (1-x_i)^{R_i}\right] \\ &= \left(\prod_{i=1}^d \gamma_i\right) h_1(\mathbf{x}_d) \cdot h_2(x_{d+1}, \ldots, x_m). \end{aligned}$$

Now, we consider h_1. Let $f_{\mathbf{k}_d}$ denote the joint density function of progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{d:m:n}$ from a right truncated uniform distribution (at $F(T)$) and censoring scheme $\mathbf{k}_d = (k_1, \ldots, k_d)$. Then, by using the binomial theorem, we find for $0 < x_1 < \cdots < x_d \le F(T)$,

$$\begin{aligned} h_1(\mathbf{x}_d) &= \prod_{i=1}^d (1-x_i)^{R_i} = \prod_{i=1}^d (1 - F(T) + F(T) - x_i)^{R_i} \\ &= \prod_{i=1}^d \left\{ \sum_{k_i=0}^{R_i} \binom{R_i}{k_i} (F(T)-x_i)^{k_i} (1-F(T))^{R_i-k_i} \right\} \\ &= \prod_{i=1}^d \left\{ \sum_{k_i=0}^{R_i} \binom{R_i}{k_i} F^{k_i+1}(T)(1-F(T))^{R_i-k_i} \frac{1}{F(T)} \left(1 - \frac{x_i}{F(T)}\right)^{k_i} \right\} \\ &= \sum_{k_1=0}^{R_1} \cdots \sum_{k_d=0}^{R_d} f_{\mathbf{k}_d}(\mathbf{x}_d) F^{\eta_1(d)}(T)(1-F(T))^{n-\gamma_{d+1}-\eta_1(d)} \left\{ \prod_{i=1}^d \frac{1}{\eta_i(d)} \binom{R_i}{k_i} \right\} \end{aligned}$$

$$= \frac{P(D=d)}{(1-F(T))^{\gamma_{d+1}} \prod_{i=1}^{d} \gamma_i} \sum_{k_1=0}^{R_1} \cdots \sum_{k_d=0}^{R_d} f_{\mathbf{k}_d}(\mathbf{x}_d) p^{\mathbf{K}_d}(\mathbf{k}_d).$$

On the other hand, from $\gamma_{d+1} = \sum_{j=d+1}^{m}(R_i + 1)$, we obtain for $F(T) < x_{d+1} < \cdots < x_m$,

$$\begin{aligned} h_2(x_{d+1}, \ldots, x_m) &= (1-F(T))^{\gamma_{d+1}} \prod_{i=d+1}^{m} \gamma_i \prod_{i=d+1}^{m} \frac{1}{1-F(T)} \left(\frac{1-x_i}{1-F(T)}\right)^{R_i} \\ &= (1-F(T))^{\gamma_{d+1}} g_{R_{d+1},\ldots,R_m}(x_{d+1}, \ldots, x_m), \end{aligned}$$

where $g_{R_{d+1},\ldots,R_m}$ denotes the joint density function of progressively Type-II censored order statistics from the left truncated uniform distribution (at $1 - F(T)$) with left truncated censoring scheme $\mathscr{R}_{\lhd d} = (R_{d+1}, \ldots, R_m)$. Combining all the results, we find

$$\begin{aligned} & f^{U_{1:m:n},\ldots,U_{m:m:n}|D=d}(\mathbf{x}_m) \\ &= \left\{ \sum_{k_1=0}^{R_1} \cdots \sum_{k_d=0}^{R_d} f_{\mathbf{k}_d}(\mathbf{x}_d) p^{\mathbf{K}_d}(\mathbf{k}_d) \right\} g_{R_{d+1},\ldots,R_m}(x_{d+1}, \ldots, x_m). \end{aligned}$$

The factorization of the density function yields the independence result. Finally, the joint density functions $f_{\mathbf{k}_d}$ and $g_{R_{d+1},\ldots,R_m}$ yield the claimed distributions. □

In the case of order statistics, the above theorem simplifies. In particular, we find from $R_1 = \cdots = R_d = 0$ that $p^{\mathbf{K}_d}$ is a one-point distribution in (0^{*d}). Thus, the corresponding result due to Iliopoulos and Balakrishnan [469] is given in the following corollary.

Corollary 2.5.7. Let $d \in \{1, \ldots, n-1\}$ and F be a cumulative distribution function. Conditional on $D = d$, the random vectors $(X_{1:n}, \ldots, X_{d:n})$ and $(X_{d+1:n}, \ldots, X_{n:n})$ are mutually independent with

$$\begin{aligned} (X_{1:n}, \ldots, X_{d:n}) &\stackrel{\mathsf{d}}{=} (V_{1:d}, \ldots, V_{d:d}), \\ (X_{d+1:n}, \ldots, X_{n:n}) &\stackrel{\mathsf{d}}{=} (W_{1:n-d}, \ldots, V_{n-d:n-d}). \end{aligned}$$

The distributions of $V_1, \ldots, V_d$ and $W_1, \ldots, W_{n-d}$ are as given in Theorem 2.5.5.

Finally, it has to be mentioned that the result of Theorem 2.5.5 can be extended to multiple cut-points $-\infty \equiv T_0 < T_1 < \cdots < T_k$. Instead of D, the random vector $(D_1, \ldots, D_k)$ is considered, where D_j counts the number of progressively

Type-II censored order statistics in the interval $(T_{j-1}, T_j]$. Further details are given in Iliopoulos and Balakrishnan [469].

2.5.4 Dependence Structure of Progressively Type-II Censored Order Statistics

In this section, we focus on the notion of multivariate total positivity. This property of a density function is important in many areas including reliability theory, since it implies association of the components of the corresponding random vector. This notion of dependence was introduced by Esary et al. [354].

Definition 2.5.8. An n-dimensional real valued random vector $\mathbf{X}$ is associated if

$$\mathrm{Cov}(g(\mathbf{X}), h(\mathbf{X})) \geq 0$$

for every pair of increasing functions $g, h : \mathbb{R}^n \longrightarrow \mathbb{R}$.

This definition has some interesting implications (cf. Szekli [829, Chap. 3]). An important feature due to Esary et al. [354] is that any subset of associated random variables is associated as well. For instance, this implies that two associated random variables are positively correlated. Furthermore, association of a random vector $(X_1, \ldots, X_n)'$ implies the inequality

$$P(X_1 > x_1, \ldots, X_n > x_n) \geq \prod_{i=1}^{n} P(X_i > x_i), \quad x_1, \ldots, x_n \in \mathbb{R}. \tag{2.43}$$

Although association is a desirable feature of random variables, it is often difficult to verify. A more restrictive property which implies association, but is often easy to verify, is multidimensional total positivity.

Definition 2.5.9 (Karlin and Rinott [510]). A density function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is MTP_2 (multidimensional totally positive) if

$$f(\mathbf{x}) f(\mathbf{y}) \leq f(\mathbf{x} \wedge \mathbf{y}) f(\mathbf{x} \vee \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

A random vector $\mathbf{X}$ is said to be MTP_2 if its density function is MTP_2.

The basic properties of multidimensional total positivity are derived in Karlin and Rinott [510]. In our setup, it is important that the indicator function $\mathbb{1}_{\mathbb{R}^n_{\leq}}(\cdot)$ is MTP_2 and that a product of the form

$$\left(\prod_{i=1}^{n} f_i(x_i)\right) g(\mathbf{x}_n)$$

has this property provided that f_i are nonnegative and that g is MTP_2. Since the joint density function of progressively Type-II censored order statistics has this structure, it immediately yields the following result well known for order statistics from absolutely continuous distributions (cf. Karlin and Rinott [510]). The MTP_2 property for discrete order statistics was established by Rüschendorf [761].

Theorem 2.5.10 (Cramer [287]). Let F be a continuous cumulative distribution function. Then, progressively Type-II censored order statistics $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ based on F with censoring scheme $\mathscr{R}$ are MTP_2. In particular, covariances are nonnegative, i.e., $\mathrm{Cov}(X^{\mathscr{R}}_{j:m:n}, X^{\mathscr{R}}_{r:m:n}) \geq 0$ for all $1 \leq j, r \leq n$.

Obviously, the MTP_2 property implies nonnegative covariances. For order statistics, nonnegative correlation was first claimed by Bickel [201]. Further, it should be noted that the MTP_2 property implies association of the progressively Type-II censored order statistics (see, e.g., Cohen and Sackrowitz [275]). However, as pointed out in Cramer and Lenz [303], the MTP_2 property does not hold in general if the assumption of identical distribution is dropped for the data $X_1, \ldots, X_n$. Nevertheless, association still holds (see Sect. 10.3).

From Theorem 2.5.10, if follows for progressively Type-II censored order statistics that any marginal distribution of at least two (different) progressively Type-II censored order statistics has the MTP_2 property (cf. Karlin and Rinott [510, Proposition 3.2 for the general result on associated random variables]). This feature of progressively Type-II censored order statistics has many interesting implications concerning the dependence structure of progressively Type-II censored order statistics. As mentioned above, it implies association of progressively Type-II censored order statistics which means that all the covariances are nonnegative. It implies inequality (2.43) giving a lower bound for the multivariate survival function in terms of the univariate survival functions.

Burkschat [229] has studied the dependence structure of spacings of generalized order statistics. His results can be directly applied to spacings of progressively Type-II censored order statistics [cf. (2.9)]

$$S^{\star\mathscr{R}}_j = X^{\mathscr{R}}_{j:m:n} - X^{\mathscr{R}}_{j-1:m:n}, j = 2, \ldots, m, \quad S^{\star\mathscr{R}}_1 = X^{\mathscr{R}}_{1:m:n}. \tag{2.44}$$

First, the notion of conditionally increasing in sequence is discussed which is defined as follows.

Definition 2.5.11. A random vector $\mathbf{X} = (X_1, \ldots, X_m)$ is said to be conditionally increasing/decreasing in sequence (CIS/DIS) if

$$P(X_j > t_j | \mathbf{X}_{j-1} = \mathbf{x}_{j-1})$$

are increasing/decreasing in $\mathbf{x}_{j-1} = (x_1, \ldots, x_{j-1})$ for any $j \in \{2, \ldots, m\}$.

Burkschat [229] proved the following result.

Theorem 2.5.12. Let F be IFR (DFR). Then, the vector of spacings $\mathbf{S}^{\star\mathscr{R}}$ is DIS (CIS).

Furthermore, he showed that spacings of progressively Type-II censored order statistics have the MTP_2 property when the baseline distribution satisfies some additional conditions. A similar result is available for the MMR_2 property (see Burkschat [229, Theorem 2.9]).

Theorem 2.5.13. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from an absolutely continuous cumulative distribution function F with censoring scheme $\mathscr{R}$ and hazard rate λ_F. Moreover, let the density function f be positive on the support (α, ∞) of F. Then, provided that

(i) f is log-convex on (α, ∞) or
(ii) F is DFR and λ_F is log-convex on (α, ∞),

the vector of spacings $\mathbf{S}^{\star\mathscr{R}}$ is MTP_2.

2.6 Basic Recurrence Relations

The expression for the marginal cumulative distribution function given in (2.25) can be used to establish a generalization of the triangle rule for cumulative distribution functions of order statistics due to Cole [278] in the continuous case (see also David and Joshi [325] and David and Nagaraja [327, p. 44]):

$$F_{r:n-1} = \frac{r}{n} F_{r+1:n} + \left(1 - \frac{r}{n}\right) F_{r:n}, \quad 1 \le r \le n-1. \tag{2.45}$$

In order to prove the required result, we use the following lemma.

Lemma 2.6.1. Let $Z_1, \ldots, Z_{r+1}$ be IID exponential random variables and $\gamma_1 > \cdots > \gamma_{r+1} > 0$. Then, for $t \in \mathbb{R}$,

$$P\Big(\sum_{j=2}^{r+1} \frac{1}{\gamma_j} Z_j \le t\Big) = \Big(1 - \frac{\gamma_{r+1}}{\gamma_1}\Big) P\Big(\sum_{j=1}^{r+1} \frac{1}{\gamma_j} Z_j \le t\Big) + \frac{\gamma_{r+1}}{\gamma_1} P\Big(\sum_{j=1}^{r} \frac{1}{\gamma_j} Z_j \le t\Big).$$

Proof. To prove the above recurrence relation, we make use of the Laplace transform of the exponential random variables Z_j/γ_j:

$$\mathscr{L}_{Z_j/\gamma_j}(t) = E(\mathrm{e}^{-tZ_j/\gamma_j}) = \frac{\gamma_j}{t+\gamma_j}, \quad t > -\gamma_j, 1 \le j \le r+1.$$

Noticing that

$$\frac{\gamma_1 - \gamma_{r+1}}{(\gamma_1 + t)(\gamma_{r+1} + t)} = \frac{1}{\gamma_{r+1} + t} - \frac{1}{\gamma_1 + t},$$

we find for the Laplace transform $\mathscr{L}_{r+1}$ of $\sum_{j=1}^{r+1} \frac{1}{\gamma_j} Z_j$, for $t > -\gamma_{r+1}$,

$$\begin{aligned}(\gamma_1 - \gamma_{r+1})\mathscr{L}_{r+1}(t) &= (\gamma_1 - \gamma_{r+1}) \prod_{j=1}^{r+1} \frac{\gamma_j}{t + \gamma_j} \\ &= \gamma_1 \prod_{j=2}^{r+1} \frac{\gamma_j}{t + \gamma_j} - \gamma_{r+1} \prod_{j=1}^{r} \frac{\gamma_j}{t + \gamma_j} \\ &= \gamma_1 \tilde{\mathscr{L}}_r(t) - \gamma_{r+1}\mathscr{L}_r(t),\end{aligned}$$

where $\mathscr{L}_r$ and $\tilde{\mathscr{L}}_r$ are the Laplace transforms of $\sum_{j=1}^{r} \frac{1}{\gamma_j} Z_j$ and $\sum_{j=2}^{r+1} \frac{1}{\gamma_j} Z_j$, respectively. A simple rearrangement yields the identity

$$\tilde{\mathscr{L}}_r = \Big(1 - \frac{\gamma_{r+1}}{\gamma_1}\Big)\mathscr{L}_{r+1} + \frac{\gamma_{r+1}}{\gamma_1}\mathscr{L}_r,$$

which proves the result. □

Theorem 2.6.2. Marginal cumulative distribution functions of progressively Type-II censored order statistics from an arbitrary cumulative distribution function F and with censoring scheme $\mathscr{R}$ satisfy the recurrence relation

$$F^{(R_2,\dots,R_m)}_{r:m-1:n-R_1-1} = \Big(1 - \frac{\gamma_{r+1}}{n}\Big) F^{\mathscr{R}}_{r+1:m:n} + \frac{\gamma_{r+1}}{n} F^{\mathscr{R}}_{r:m:n}, \quad 1 \le r \le m-1. \quad (2.46)$$

Proof. From the representation in Corollary 2.3.7, we conclude that it is sufficient to consider exponential progressively Type-II censored order statistics. But, the corresponding identity for exponential progressively Type-II censored order statistics follows directly from Lemma 2.6.1 and (2.13) (see also Corollary 2.3.7). □

Remark 2.6.3.

(i) Relation (2.46) was first established by Kamps and Cramer [503] using density representations;
(ii) $(R_2, \dots, R_m) = \mathscr{R}_{\lhd 1}$ is a left truncated censoring scheme. For order statistics, this yields directly the classical triangle rule (2.45);
(iii) It is easy to see that the right-hand side of (2.46) is a convex combination of cumulative distribution functions with probabilities $1 - \frac{\gamma_{r+1}}{n}$ and $\frac{\gamma_{r+1}}{n}$. Notice that $\frac{\gamma_{r+1}}{n}$ is the probability that a particular choice of random variables remains in the experiment after the rth censoring step;

(iv) The above identities hold for density functions and moments as well (provided they exist):
(v) Obviously, the representation in Lemma 2.6.1 holds also for other choices of γ_1 and γ_{r+1}. Thus, we can obtain other identities by selecting other γ's.

A similar relation has been established by Balakrishnan et al. [137] for bivariate marginals. For $1 \le r < s < m-1$, they obtained

$$F^{\mathscr{R}_{\lhd 1}}_{r,s:m-1:n-R_1-1} = \left(1-\frac{\gamma_{r+1}}{n}\right)F^{\mathscr{R}}_{r+1,s+1:m:n} + \frac{\gamma_{r+1}-\gamma_{s+1}}{n}F^{\mathscr{R}}_{r,s+1:m:n} + \frac{\gamma_{s+1}}{n}F^{\mathscr{R}}_{r,s:m:n}. \quad (2.47)$$

This shows that the bivariate cumulative distribution function $F^{\mathscr{R}_{\lhd 1}}_{r,s:m-1:n-R_1-1}$ can be written as a convex combination of three bivariate cumulative distribution functions from a progressively censored sample with censoring plan $\mathscr{R}$. This quadruple rule extends a result for order statistics due to Srikantan [822] (see also David and Joshi [325]):

$$F_{r,s:n-1} = \frac{r}{n}F_{r+1,s+1:n} + \frac{s-r}{n}F_{r,s+1:n} + \left(1-\frac{s}{n}\right)F_{r,s:n}.$$

Related results are given by Govindarajulu [409] and Balasubramanian and Beg [164].

2.7 Shape of Density Functions

We now present unimodality properties of progressively Type-II censored order statistics established in Cramer [286]. A cumulative distribution function F is said to be unimodal with a mode η if F is convex on $(-\infty, \eta)$ and concave on (η, ∞). In particular, we consider the stronger concept of log-concavity. This approach extends well-known results for order statistics which are summarized in Dharmadhikari and Joag-dev [339]. For instance, it is proved that progressively Type-II censored order statistics based on a strongly unimodal cumulative distribution function F are strongly unimodal (cf. Barlow and Proschan [167], Huang and Ghosh [460]).

2.7.1 Log-Concavity of Uniform Progressively Type-II Censored Order Statistics

Cramer et al. [313] proved by an induction argument similar to that given in Balakrishnan et al. [129, Lemma 2.7] that uniform generalized order statistics are unimodal. Since uniform progressively Type-II censored order statistics are a particular model of generalized order statistics, this implies the desired result. Considering the notion of strong unimodality, the respective result is strengthened and the proof is simplified.

A cumulative distribution function F is said to be strongly unimodal if the convolution of F with any unimodal cumulative distribution function G is unimodal (cf. Ibragimov [468] and Hájek and Šidák [429]). It can be seen (cf. Dharmadhikari and Joag-dev [339]) that the set of strongly unimodal cumulative distribution functions is closed under convolutions and weak limits and that any degenerate cumulative distribution function is strongly unimodal. A fundamental result of Ibragimov [468] says that nondegenerate strongly unimodal distributions are absolutely continuous with a log-concave density function (cf. Dharmadhikari and Joag-dev [339, Theorem 1.9, Lemma 1.4]).

Definition 2.7.1. Let $m \in \mathbb{N}$. A nonnegative function $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ is said to be log-concave if $\log g$ is concave, i.e., $\log g(\lambda t + (1-\lambda)z) \geq \lambda \log g(t) + (1-\lambda)\log g(z)$ for all $t, z \in \mathbb{R}^m$ and $\lambda \in [0,1]$.

The result of Ibragimov [468] links this property of the density function with strong unimodality of the corresponding cumulative distribution function (cf. Dharmadhikari and Joag-dev [339, Theorem 1.10]). If F is a nondegenerate cumulative distribution function, then F is strongly unimodal iff F is absolutely continuous and its density function f is log-concave. At this point, it has to be mentioned that log-concavity of the density f is equivalent to the property that f is a Pólya frequency function of order 2 (for brevity, we write f PF_2). Pólya frequency functions of order 2 are functions such that $K(x,y) = f(x-y)$ is totally positive of order 2 (cf. Karlin [509]). Barlow and Proschan [168, p. 76] proved that the PF_2 property of a cumulative distribution function F is equivalent to its IFR property. In Lemma 5.8, they established that a strongly unimodal cumulative distribution function has the IFR property.

In order to prove that the cumulative distribution function of a uniform progressively Type-II censored order statistic is unimodal, it is shown that the associated density function is log-concave. The following theorem shows that the joint density of uniform progressively Type-II censored order statistics is log-concave.

Theorem 2.7.2. The joint density function $f^{\mathbf{U}^{\mathscr{R}}}$ of uniform progressively Type-II censored order statistics is log-concave.

Proof. The joint density function of $U_{1:m:n}, \ldots, U_{m:m:n}$ can be expressed as $f^{\mathbf{U}^{\mathscr{R}}}(\mathbf{u}_m) = c\prod_{j=1}^m g_j(u_j)$, $0 \leq u_1 \leq \cdots \leq u_m < 1$, where $g_j(t) = (1-t)^{R_j}$,

$t \in [0, 1]$, are log-concave functions, $1 \leq j \leq m$. Since a product of log-concave functions is log-concave, $f^{\mathbf{U}^{\mathscr{R}}}$ is log-concave. □

The preceding theorem is now applied to the marginal distributions of uniform progressively Type-II censored order statistics. We make use of the following lemma which was established independently by Prékopa [729] and Brascamp and Lieb [219] (see also Eaton [345]).

Lemma 2.7.3. Let $m, n \in \mathbb{N}$ and $f : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}$ be a log-concave density function. Let $g : \mathbb{R}^m \longrightarrow \mathbb{R}$ be defined by $g(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{x}, \mathbf{y})\, d\mathbf{y}$, $\mathbf{x} \in \mathbb{R}^m$. Then, g is log-concave on $\mathbb{R}^m$.

Corollary 2.7.4. Suppose $U_{1:m:n}, \ldots, U_{m:m:n}$ are uniform progressively Type-II censored order statistics. Then, any marginal density is log-concave. In particular, $F^{U_{r:m:n}}$ is strongly unimodal and, thus, unimodal, $1 \leq r \leq m$.

2.7.2 The Shape of Densities of Uniform Progressively Type-II Censored Order Statistics

In addition to the unimodality and log-concavity properties, the shape of the density functions of uniform progressively Type-II censored order statistics can be classified. The following result is taken from Bieniek [203] who established it in the more general case of generalized order statistics.

Theorem 2.7.5. The density functions of uniform progressively Type-II censored order statistics have the following shapes:

(i) $f^{U_{1:m:n}}$ is constant for $n = 1$, linear decreasing for $n = 2$, and convex decreasing for $n \geq 3$;
(ii) Let $m \geq 2$. $f^{U_{2:m:n}}$ is
 (a) linear increasing for $\gamma_2 = 1$ and $\gamma_1 = n = 2$;
 (b) concave increasing for $\gamma_2 = 1$ and $\gamma_1 = n \geq 3$;
 (c) concave increasing–decreasing for $\gamma_2 = 2$;
 (d) concave increasing, concave decreasing, and convex decreasing for $\gamma_2 \geq 3$;
(iii) For $m \geq r \geq 3$, $f^{U_{r:m:n}}$ is
 (a) convex increasing for $\gamma_r = 1$ and $\gamma_{r-1} = 2$;
 (b) convex–concave increasing for $\gamma_r = 1$ and $\gamma_{r-1} \geq 3$;
 (c) convex increasing, concave increasing, and concave decreasing for $\gamma_r = 2$;
 (d) convex increasing, concave increasing, concave decreasing, and convex decreasing for $\gamma_r \geq 3$.

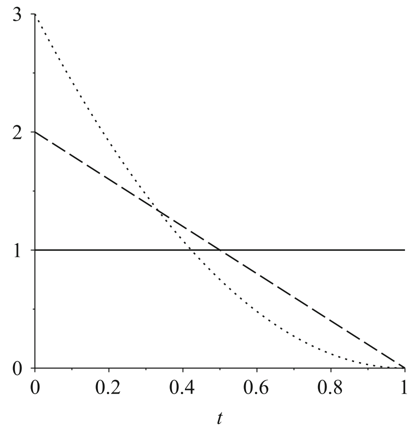

Fig. 2.1 Plots of $f^{U_{1:m:n}}$ for $n = 1$ (*solid line*), $n = 2$ (*dashed line*), and $n = 3$ (*dotted line*)

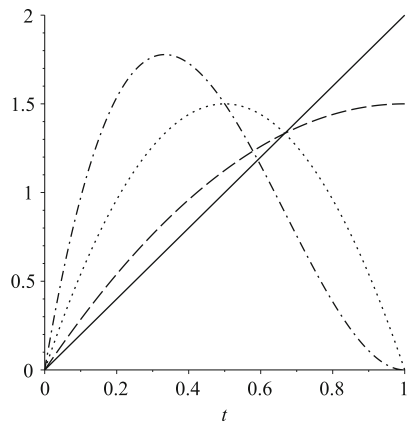

Fig. 2.2 Plots of $f^{U_{2:m:n}}$ for $\gamma_2 = 1, n = 2$ (*solid line*) $\gamma_2 = 1, n = 3$ (*dashed line*), $\gamma_2 = 2, n = 3$ (*dotted line*), and $\gamma_2 = 3, n = 4$ (*dashed–dotted line*)

Figures 2.1, 2.2, and 2.3 illustrate the shapes of density functions of uniform progressively Type-II censored order statistics given in (2.24).

In order to prove these shapes, Bieniek [203] established a variation diminishing property of density functions of uniform progressively Type-II censored order statistics. This property is well known for density functions of uniform order statistics, i.e., Bernstein polynomials (see Schoenberg [786]). In particular, let $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{R}^m \setminus \{0\}$. Then, he proved that the number of zeros of any linear combination

$$H_{\mathbf{a}} = \sum_{j=1}^{m} a_j f^{U_{j:m:n}} \tag{2.48}$$

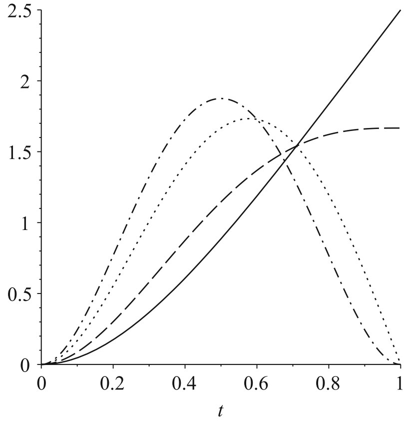

Fig. 2.3 Plots of $f^{U_{3:m:n}}$ for $\gamma_3 = 1, \gamma_2 = 2, n = 5$ (*solid line*) $\gamma_3 = 1, \gamma_2 = 4, n = 5$ (*dashed line*), $\gamma_3 = 2, \gamma_2 = 4, n = 5$ (*dotted line*), and $\gamma_3 = 3, \gamma_2 = 4, n = 5$ (*dashed–dotted line*)

in the unit interval (0, 1) does not exceed the number of sign changes $S^-(\mathbf{a})$ in the sequence $(a_1, \ldots, a_r)$ (after deleting the zeroes in **a**). Denoting by $Z(f)$ the number of zeroes in (0, 1), the result is given as follows.

Theorem 2.7.6. For any censoring scheme $\mathscr{R}$ and $\mathbf{a} = (a_1, \ldots, a_m) \in \mathbb{R}^m \setminus \{0\}$,

$$Z(H_{\mathbf{a}}) \leq S^-(\mathbf{a}) \leq m - 1.$$

Moreover, Bieniek [203] obtained for progressively Type-II censored order statistics that the sign of H_a close to zero (1) is determined by the sign of the first (last) nonzero element of **a**. The result for order statistics has been established by Gajek and Rychlik [386].

2.7.3 *Unimodality and Log-Concavity of Progressively Type-II Censored Order Statistics Based on F*

For exponential distribution, strong unimodality is obvious.

Theorem 2.7.7. Any marginal density function of progressively Type-II censored order statistics based on an exponential distribution is log-concave. Moreover, the one-dimensional cumulative distribution functions are strongly unimodal.

Huang and Ghosh [460] presented a proof that the cumulative distribution function of an order statistic is strongly unimodal provided that the underlying cumulative distribution function F is strongly unimodal. The same property was obtained earlier by Barlow and Proschan [167, Theorem 7.2] in terms of PF_2

functions (which is an equivalent formulation of log-concavity of the density function). This result has been extended to progressively Type-II censored order statistics by Cramer [286]. Chen et al. [254, Theorem 2.1] present a more general result in terms of generalized order statistics that covers the log-concavity property as a special case. For a vector $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_{m-1})$ with $0 \le \tau_j \le j$, let

$$\mathbf{V}(\boldsymbol{\tau}) = (X_{1:m:n}, X_{2:m:n} - X_{\tau_1:m:n}, \ldots, X_{m:m:n} - X_{\tau_{m-1}:m:n}),$$

where $X_{0:m:n} = 0$.

Theorem 2.7.8 (Chen et al. [254]). Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics based on a cumulative distribution function F with log-concave density function f. Then, $\mathbf{V}(\boldsymbol{\tau})$, and, thus, each subvector of $\mathbf{V}(\boldsymbol{\tau})$, has a log-concave density function.

Choosing $\boldsymbol{\tau} = (0^{*m-1})$, the following result due to Cramer [286] is included as a special case.

Corollary 2.7.9. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics based on a strongly unimodal cumulative distribution function F. Then, any marginal density function is log-concave and $F^{X_{r:m:n}}$ is strongly unimodal, $1 \le r \le m$.

Remark 2.7.10. As a consequence of Corollary 2.7.9, progressively Type-II censored order statistics based on strongly unimodal cumulative distribution functions are strongly unimodal and, therefore, unimodal. Examples for strongly unimodal cumulative distribution functions include the following distributions: exponential, normal, truncated normal, Laplace, and particular Weibull and gamma distributions. Further examples are presented in Hájek and Šidák [429, Table 1, p. 16] and Barlow and Proschan [168, p. 79].

Theorem 2.7.8 includes also results for generalized p-spacings of progressively Type-II censored order statistics $X_{p+j:m:n} - X_{j:m:n}$. For instance, it extends a result of Misra and van der Meulen [651] for p-spacings of order statistics. For completeness, we present the result for spacings.

Corollary 2.7.11. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics based on a strongly unimodal cumulative distribution function F. Then, the vector $\mathbf{S}^{\star\mathscr{R}}$ of spacings has a log-concave density function.

Alam [31] proves that order statistics based on an absolutely continuous cumulative distribution function F with density function f are unimodal if the reciprocal function $1/f$ is convex. Note that concavity of $\log f$ implies convexity of $1/f$. As pointed out by Huang and Ghosh [460], the Cauchy distribution has the above property, but it is not strongly unimodal. In the next theorem, Alam's [31] result is extended to progressively Type-II censored order statistics. A generalization to generalized order statistics is available in Cramer [285] and Alimohammadi and Alamatsaz [39].

Theorem 2.7.12 (Cramer [286]). Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics based on an absolutely continuous cumulative distribution function F with density function f.

Then, convexity of $1/f$ implies unimodality of $F^{X_{r:m:n}}$, $1 \le r \le m$.

Remark 2.7.13. Theorem 2.7.12 provides an alternative proof for the unimodality of the cumulative distribution function of a uniform progressively Type-II censored order statistic. Since the density of the standard uniform distribution is constant on $(0,1)$ and, therefore, its reciprocal is trivially convex, the theorem leads directly to the unimodality of the respective cumulative distribution function. Further examples are normal, exponential, gamma, and Cauchy distributions.

2.8 Discrete Progressively Type-II Censored Order Statistics

Progressively Type-II censored order statistics from noncontinuous distributions have been studied in Balakrishnan and Dembińska [95,96,97]. The results are based on the quantile representation of progressively Type-II censored order statistics which has been established in Theorem 2.1.1 (see Balakrishnan and Dembińska [96, 97]). For discrete distribution, the quantile representation yields directly the following probability mass function. An extensive discussion of discrete order statistics can be found in Arnold et al. [58, Chap. 3].

Theorem 2.8.1. The joint probability mass function of discrete progressively Type-II censored order statistics $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ from a discrete cumulative distribution function F with support $\mathbb{D}$ is given by

$$P(X^{\mathscr{R}}_{j:m:n} = x_j, 1 \le j \le m) = \int_{\mathcal{A}} f^{\mathbf{U}^{\mathscr{R}}}(\mathbf{u}_m) d\mathbf{u}_m, \quad \mathbf{x}_m \in \mathbb{D}^m, \tag{2.49}$$

where

$$\mathcal{A} = \{\mathbf{u}_m \mid u_1 \le \cdots \le u_m, F(x_j-) < u_j \le F(x_j), j = 1, \ldots, m\}.$$

For discrete order statistics, the corresponding integral representation of the joint probability mass function can be found in Arnold et al. [58, p. 46]. Obviously, a similar representation holds for marginal probability mass functions by replacing $f^{\mathbf{U}^{\mathscr{R}}}$ by the corresponding marginal density function of uniform progressively Type-II censored order statistics. In particular, it follows that the one-dimensional marginal cumulative distribution functions given in (2.25) hold also for discrete parents.

Balakrishnan and Dembińska [96] pointed out that discrete progressively Type-II censored order statistics do not form a Markov chain when the support contains at least three points. The same results has been established independently in the more general setting of generalized order statistics in Tran [854] (see also Cramer and Tran [307]). This work contains also expressions of density function w.r.t. the product measure $\bigotimes_{i=1}^{m} P^F$ for arbitrary discontinuous cumulative distribution functions F.

As in Gan and Bain [391], the concept of tie-runs is applied to obtain a simple expression for the joint probability mass function in the discrete case.

Definition 2.8.2. Let $x_1, \dots, x_r \in \mathbb{R}$ with $x_1 \le \dots \le x_r$. Then, $x_1, \dots, x_r$ is said to have k tie-runs with lengths $\tau_1, \dots, \tau_k$ if

$$x_1 = \dots = x_{\tau_1} < x_{\tau_1+1} = \dots = x_{\tau_1+\tau_2} < \cdots$$
$$\cdots < x_{\tau_1+\dots+\tau_{k-1}+1} = \dots = x_{\tau_1+\dots+\tau_k}$$

with $\sum_{j=1}^{k} \tau_j = r$.

Furthermore, we introduce the lexicographic distribution function to F as given in Arnold et al. [57] and Reiss [750, p. 34].

Definition 2.8.3. Let $F : \mathbb{R} \longrightarrow [0,1]$ be a cumulative distribution function. The function

$$L_F : \begin{cases} \mathbb{R} \times [0,1] \longrightarrow [0,1] \\ (x,u) \qquad \longmapsto F(x-) + u[F(x) - F(x-)] \end{cases}$$

is said to be lexicographic distribution function of F.

This yields an alternative representation of the density function of progressively Type-II censored order statistics.

Theorem 2.8.4 (Tran [854], Cramer and Tran [307]). Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from an arbitrary cumulative distribution function F, and let A_F denote the set of points of discontinuity of F, i.e.,

$$A_F = \{x \in \mathbb{R} : F(x) - F(x-) > 0\}.$$

For $x_1 \le \dots \le x_r$, let $\tau_1, \dots, \tau_k \in \mathbb{N}$ denote the lengths of tie-runs in this sequence.

Then, the $\bigotimes_{j=1}^{r} P^F$-joint density function of the first r progressively Type-II censored order statistics is given by

$$f^{\mathbf{X}_r^{\mathscr{R}}}(\mathbf{x}_r) = \prod_{j=1}^{r} \gamma_j \int_{[0,1]^r} \mathbb{1}_{[0,1]^r_{\le}}(L_F(x_1,u_1), \dots, L_F(x_r,u_r))$$
$$\times \Big[\prod_{j=1}^{r} (1 - L_F(x_j,u_j))^{R_j} \Big] d\lambda^r(\mathbf{u}_r) \qquad (2.50)$$

$$= \prod_{j=1}^{r} \gamma_j \mathbb{1}_{\mathbb{R}^r_{\leq}}(\mathbf{x}_r) \Big[\prod_{j \in I_k^c} \frac{(1 - F(x_{l_j}))^{\gamma_{l_{j-1}+1} - \gamma_{l_j+1} - \tau_j}}{\tau_j!} \Big]$$
$$\times \Big[\prod_{j \in I_k} (F(x_{l_j}) - F(x_{l_j}-))^{-\tau_j}$$
$$\times \int_{\mathcal{B}_j(x_{l_j})} \prod_{l=l_{j-1}+1}^{l_j} (1 - z_l)^{R_l} dz_{l_{j-1}+1} \cdots dz_{l_j} \Big],$$

where $l_j = \sum_{i=1}^{j} \tau_i$, $j = 1, \ldots, k$, $l_0 = 0$, $\gamma_{r+1} = 0$, and $I_k = \{j \in \{1, \ldots, k\} \mid x_{l_j} \in A_F\}$, $I_k^c = \{1, \ldots, k\} \setminus I_k$, and, for $t \in \mathbb{R}$,

$$\mathcal{B}_j(t) = \{\mathbf{v}_{\tau_j} : F(t-) \leq v_1 \leq \cdots \leq v_{\tau_j} \leq F(t)\} \subseteq [0,1]^{\tau_j}_{\leq}, \quad j \in I_k.$$

For order statistics, (2.50) yields a representation due to Arnold et al. [57, Lemma 2.1], i.e.,

$$f^{X_{1:n},\ldots,X_{r:n}}(\mathbf{x}_r)$$
$$= \frac{n!}{(n-r)!} \int_{[0,1]^r} \mathbb{1}_{[0,1]^r_{\leq}}(L_F(x_1, u_1), \ldots, L_F(x_r, u_r)) \overline{L}_F^{n-r}(x_r, u_r) d\lambda^r(\mathbf{u}_r).$$

Remark 2.8.5. The representations in Theorem 2.8.4 simplify when the values $x_1, \ldots, x_r$ are restricted to either continuity or discontinuity points.

For $x_j \notin A_F$, $j = 1, \ldots, r$, and $x_1 < \cdots < x_r$, then $I_k = \emptyset$ and $I_k^c = \{1, \ldots, k\}$. Then, a representation of the $\bigotimes_{j=1}^{r} P^F$-density function results which is similar to that known in the continuous case [see (2.4)]:

$$f^{X_{1:m:n},\ldots,X_{r:m:n}}(\mathbf{x}_r) = \prod_{j=1}^{r-1} \big[\gamma_j (1 - F(x_j))^{R_j}\big] \gamma_r (1 - F(x_r))^{\gamma_r - 1}.$$

For $x_j \in A_F$, $j = 1, \ldots, r$, and $x_1 < \cdots < x_r$, then

$$f^{X_{1:m:n},\ldots,X_{r:m:n}}(\mathbf{x}_r) = \Big[\prod_{j=1}^{r} \Big(F(x_j) - F(x_j-) \Big)^{-1} \Big] \int_{B_r} f^{\mathbf{U}_r^{\mathscr{R}}}(\mathbf{u}_r) d\lambda^r(\mathbf{u}_r),$$

where $B_r = \bigtimes_{j=1}^{r} [F(x_j-), F(x_j)]$. This expression is seen to coincide with (2.49) for $r = m$ noticing that $f^{X_{1:m:n},\ldots,X_{r:m:n}}$ is given w.r.t. the product measure $\bigotimes_{j=1}^r P^F$. Multiplying $f^{X_{1:m:n},\dots,X_{r:m:n}}$ with $\prod_{j=1}^r P(X_j = x_j) = \prod_{j=1}^r \big(F(x_j) - F(x_j-)\big)$ leads to (2.49). Since B_r is a Cartesian product, one gets

$$f^{X_{1:m:n},\dots,X_{r:m:n}}(\mathbf{x}_r) = \prod_{j=1}^r \left[\frac{\gamma_j}{F(x_j)-F(x_j-)}\int_{F(x_j-)}^{F(x_j)} (1-u_j)^{R_j}\,du_j)\right]$$

$$= \left[\prod_{j=1}^{r-1} \frac{\gamma_j[\overline{F}^{R_j+1}(x_j-)-\overline{F}^{R_j+1}(x_j)]}{(R_j+1)[\overline{F}(x_j-)-\overline{F}(x_j)]}\right]\frac{\overline{F}^{\gamma_r}(x_r-)-\overline{F}^{\gamma_r}(x_r)}{\overline{F}(x_r-)-\overline{F}(x_r)}. \tag{2.51}$$

Example 2.8.6. Suppose F is the cumulative distribution function of a geometric distribution with parameter $p \in (0,1)$ and support $\mathbb{N}$. Then, $\overline{F}(t) = (1-p)^t$, $t \in \mathbb{N}_0$, and (2.51) simplifies to

$$f^{X_{1:m:n},\dots,X_{r:m:n}}(\mathbf{x}_r) = \Big(\prod_{j=1}^{r-1}\frac{\gamma_j}{R_j+1}(1-p)^{R_j(x_j-1)}\Big)(1-p)^{(\gamma_r-1)(x_r-1)}h(p)$$

$$= \Big(\prod_{j=1}^{r-1}\frac{\gamma_j}{R_j+1}\Big)(1-p)^{\sum_{j=1}^{r-1}R_j(x_j-1)+(\gamma_r-1)(x_r-1)}h(p),$$

where $h(p) = p^{-r}\big(\prod_{j=1}^{r-1}[1-(1-p)^{R_j+1}]\big)[1-(1-p)^{\gamma_r}]$.

Finally, we present a result on the Markovian structure due to Tran [854] (see also Cramer and Tran [307]). It extends a result of Rüschendorf [761] for order statistics from arbitrary cumulative distribution function F. It shows that progressively Type-II censored order statistics from discontinuous cumulative distribution function form a Markov chain if the ties are taken into account.

Theorem 2.8.7. Let $\mathbf{X}^{\mathscr{R}}$ be the random vector of progressively Type-II censored order statistics. For $1 \le r \le m$ and $\mathbf{x}_r = (x_1,\dots,x_r) \in \mathbb{R}^r_\le$, let $\tau_1,\dots,\tau_k \in \mathbb{N}$ denote the lengths of the occurring ties in $\mathbf{x}_r$. Moreover, let $\tilde{\tau}_r : \mathbb{R}^r_\le \longrightarrow \{1,\dots,r\}$ be a map defined by

$$\tilde{\tau}_r = \tilde{\tau}_r(\mathbf{x}_r) = \sum_{j=1}^r \mathbb{1}_{\{x_r\}}(x_j), \quad \mathbf{x}_r \in \mathbb{R}^r_\le,$$

and let $T_r = \tilde{\tau}_r\big(X_{1:m:n},\dots,X_{r:m:n}\big)$. Then,

(i) the joint $\bigotimes_{j=1}^r (P^F \otimes \sum_{l=1}^r \varepsilon_l)$-density of progressively Type-II censored order statistics $X_{1:m:n},\dots,X_{r:m:n}$ and $T_1,\dots,T_r$ is given by

$$h(x_1, t_1, \ldots, x_r, t_r) = \mathbb{1}_{\{(\tilde{\tau}_1(\mathbf{x}_r), \ldots, \tilde{\tau}_r(\mathbf{x}_r))\}}(\mathbf{t}_r) f^{X_{1:m:n}, \ldots, X_{r:m:n}}(\mathbf{x}_r),$$
$$t_1, \ldots, t_r, \in \mathbb{N},\ x_1, \ldots, x_r \in \mathbb{R},$$

where ε_l denotes the probability measure associated with the degenerate distribution in l, $l = 1, \ldots, r$;

(ii) $(X_{j:m:n}, T_j)_{1 \le j \le m}$ forms a Markov chain.

Remark 2.8.8.

(i) If F is continuous, then $(T_j)_{1 \le j \le n}$ equals $(1, \ldots, 1)$ P^F a.e. This yields the Markovian property of $(X_{j:m:n})_{1 \le j \le m}$ as given in Theorem 2.5.1.

(ii) Tran [854] and Balakrishnan and Dembińska [96] showed that progressively Type-II censored order statistics (or, more generally, generalized order statistics) do not form a Markov chain when the support of F has at least three points (see also Balakrishnan and Dembińska [95]). For order statistics, this result is due to Nagaraja [661] (see also Arnold et al. [58]). For further details on the dependence structure of order statistics, we refer to Arnold et al. [57] and Nagaraja [662, 665].

2.9 Exceedances

Bairamov and Eryılmaz [78] addressed the problem of exceedance statistics. Consider progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{m:m:n}$ from a continuous cumulative distribution function F and an IID sample $Y_1, \ldots, Y_k$ from a continuous cumulative distribution function G. Then, for $1 \le r < s \le m$, the statistics $V_{r:m:n}^{(k)}$ and $W_{r,s:m:n}^{(k)}$ are defined as

$$V_{r:m:n}^{(k)} = \sum_{i=1}^{k} \mathbb{1}_{(-\infty, X_{r:m:n})}(Y_i), \quad W_{r,s:m:n}^{(k)} = \sum_{i=1}^{k} \mathbb{1}_{(X_{r:m:n}, X_{s:m:n})}(Y_i),$$

i.e., the number of Y's not exceeding $X_{r:m:n}$ and included in the interval $(X_{r:m:n}, X_{s:m:n})$, respectively. Notice that $W_{r,s:m:n}^{(k)} = V_{s:m:n}^{(k)} - V_{r:m:n}^{(k)}$. Clearly,

$$E V_{r:m:n}^{(k)} = P(Y_1 \le X_{r:m:n}) = \int [1 - F_{r:m:n}(t)] dG(t) = p_r^{\mathscr{R}}, \quad \text{say.}$$

Using arguments as in Bairamov [75] and Bairamov and Eryılmaz [77], Bairamov and Eryılmaz [78] showed that

$$\lim_{k \to \infty} \frac{1}{k} V_{r:m:n}^{(k)} \xrightarrow{\mathsf{d}} P^{G(X_{r:m:n})}, \quad \lim_{k \to \infty} \frac{1}{k} W_{r:m:n}^{(k)} \xrightarrow{\mathsf{d}} P^{G(X_{s:m:n}) - G(X_{r:m:n})}.$$

For $F = G$, the probability $p_r^{\mathscr{R}}$ can be written as

$$p_r^{\mathscr{R}} = \int_0^1 [1 - F^{U_{r:m:n}}(t)]dt = EU_{r:m:n} = 1 - \prod_{j=1}^{r} \frac{\gamma_j}{\gamma_j + 1}$$

(see Theorem 7.2.3). Since $\sum_{i=1}^{k} \mathbb{1}_{(-\infty,u)}(F(Y_i))$ is $\mathsf{bin}(k, u)$-distributed, the exact distribution of $V_{r:m:n}^{(k)}$ under the hypothesis $F = G$ is given by

$$P(V_{r:m:n}^{(k)} = j) = \binom{k}{j}\Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{i=1}^{r} a_{i,r} \mathsf{B}(j + 1, \gamma_i + k - j), \quad j \in \{0, \dots, k\}.$$

Chapter 3
Further Distributional Results on Progressive Type-II Censoring

3.1 Characterizations by Progressively Type-II Censored Order Statistics

Characterizations of distributions by properties of order statistics have received great attention in the literature. Many different results have been obtained, and the situation is rather confusing due to the various assumptions, conditions, and distributions. For reviews on developments, one may refer to Galambos [390], Reiss [750], Arnold et al. [58], Balakrishnan and Basu [88], Gather et al. [393], Kamps [500], David and Nagaraja [327], and Ahsanullah and Hamedani [24]. In the following, we present only some major results which were established for progressively Type-II censored order statistics. Some of the results may also be valid in the more general setting of generalized order statistics (where possibly some additional assumptions may have to be imposed on the parameters). Thus, the results may be slightly more general than presented here.

3.1.1 Characterizations by Independence Properties

Characterizations of generalized Pareto distributions by independence of certain random variables are discussed in many settings. For instance, independence of spacings of order statistics characterizes an exponential distribution. In the following, we summarize some characterizations based on independence properties of progressively Type-II censored order statistics.

To begin with, we present a characterization established by Marohn [638]. It shows that, for instance, the independence of spacings in the exponential case is a characterizing property of the exponential distribution. The result reverses Theorem 2.3.2 (exponential distribution), Corollary 2.3.11 (reflected power distribution), and Corollary 2.3.14 (Pareto distribution).

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_3,

Theorem 3.1.1. (i) $X_{1:m:n}, \dots, X_{m:m:n}$ are progressively Type-II censored order statistics from a two-parameter exponential distribution Exp(μ, ϑ) iff, with $X_{0:m:n} = \mu$,

$$S_j^{\mathscr{R}} = \gamma_j (X_{j:m:n} - X_{j-1:m:n}), \quad 1 \le j \le m, \tag{3.1}$$

are independent random variables with $S_j^{\mathscr{R}} \sim$ Exp(ϑ), $1 \le j \le m$;

(ii) $X_{1:m:n}, \dots, X_{m:m:n}$ are progressively Type-II censored order statistics from a reflected power function distribution RPower(β) iff, with $X_{0:m:n} = 0$,

$$V_j = \left(\frac{1 - X_{j:m:n}}{1 - X_{j-1:m:n}} \right)^{\beta}, \quad 1 \le j \le m,$$

are independent random variables with $V_j \sim$ Beta$(\gamma_j, 1)$, $1 \le j \le m$;

(iii) $X_{1:m:n}, \dots, X_{m:m:n}$ are progressively Type-II censored order statistics from a Pareto distribution Pareto(α) iff, with $X_{0:m:n} = 1$,

$$W_j = \frac{X_{j:m:n}^{\alpha}}{X_{j-1:m:n}^{\alpha}}, \quad 1 \le j \le m,$$

are independent random variables with $W_j \sim$ Pareto(γ_j), $1 \le j \le m$.

Kamps and Keseling [504] established the following extension of a theorem due to Rogers [756] for the usual order statistics.

Proposition 3.1.2. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F, $1 \le r < s \le m$, and let $h : \mathbb{R}^{m-s+1} \longrightarrow \mathbb{R}$ be a measurable function. Then, if $X_{r:m:n}$ and $h(X_{s:m:n}, \dots, X_{m:m:n})$ are independent, $X_{j:m:n}$ and $h(X_{s:m:n}, \dots, X_{m:m:n})$ are also independent for every j with $r + 1 \le j \le s$.

This result is used to prove an extension of Rossberg's theorem for the usual order statistics (see Rossberg [760]). It was extended to progressively Type-II censored order statistics by Balakrishnan and Malov [112] and Kamps and Keseling [504] with the latter discussing generalized order statistics.

Theorem 3.1.3. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F, $1 \le p \le r < s \le m$, and $c_j \in \mathbb{R}$, $r \le j \le s$ with $\sum_{j=r}^{s} c_j = 0$, $c_r, c_s \ne 0$. Then, $X_{p:m:n}$ and the contrast $L_{rs} = \sum_{j=r}^{s} c_j X_{j:m:n}$ are independent iff the baseline distribution is an Exp(μ, ϑ)-distribution with $\mu \in \mathbb{R}$ and $\vartheta > 0$.

The preceding result is very general and includes many interesting characterizations as special cases. For illustration, we give the following examples.

Corollary 3.1.4. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F, $1 \le p \le r < s \le m$.

Then, each of the following properties provides a characterization of an Exp(μ, ϑ)-distribution with $\mu \in \mathbb{R}$ and $\vartheta > 0$:

(i) $X_{p:m:n}$ and $X_{s:m:n} - X_{r:m:n}$ are independent,
(ii) $X_{m-1:m:n}$ and $X_{m:m:n} - X_{m-1:m:n}$ are independent,
(iii) $X_{1:m:n}$ and $\sum_{j=2}^{m}(X_{j:m:n} - X_{1:m:n})$ are independent.

Kamps and Keseling [504] pointed out that characterizations of other distributions result via monotone transformations. For details, one may refer to their article.

Hashemi and Asadi [433] used a result of Oakes and Dasu [693] to establish characterizations of generalized Pareto distribution in terms of the mean residual life function defined as

$$m(t) = E(X - t|X > t),$$

where $X \sim F$. They proved, among other results, the following theorem.

Theorem 3.1.5. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from an absolutely continuous cumulative distribution function F with $F(0) = 0$ and $1 \le r < s \le m$. Let $\theta(x) = m(x)/m(0)$, $x \ge 0$, where $m(\cdot)$ denotes the mean residual life function of F. Then, $X_{r:m:n}$ and $(X_{s:m:n} - X_{r:m:n})/\theta(X_{r:m:n})$ are independent iff the baseline distribution is a generalized Pareto distribution.

3.1.2 Characterizations by Distributional Properties

Characterizations of the exponential distribution by distributional properties have been widely discussed. Cramer et al. [311] obtained the following characterization extending a result of Ahsanullah [21] for the usual order statistics.

Theorem 3.1.6. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from a cumulative distribution function F with $F^{\leftarrow}(0+) = 0$ and $F(x) < 1$, $x > 0$. Moreover, suppose that F is increasing failure rate (IFR) or DFR. Denote by $S_j^{\mathscr{R}}$, $1 \le j \le m$, the normalized spacings as in (3.1). Then, if $ES_r^{\mathscr{R}} = ES_{r+1}^{\mathscr{R}}$ for some $r \in \{1, \ldots, m-1\}$, the population distribution is an Exp(ϑ)-distribution for some $\vartheta > 0$.

A similar characterization result refers to the new better than used (NBU) or NWU property. For order statistics, either the identity of the distribution of the subrange $X_{s:n} - X_{r:n}$, $r < s$, and of the order statistic $X_{s-r:n-r}$ or the identity $EX_{s:n} - EX_{r:n} = EX_{s-r:n-r}$ is a characteristic property of exponential distributions (see Ahsanullah [22], Gajek and Gather [385], Iwińska [472], Gather et al. [393, p. 266/267], and Kamps [501]). An extension to progressively Type-II censored order statistics has been provided by Kamps and Cramer [503].

Theorem 3.1.7. Let $X_{1:m:n}^{\mathscr{R}}, \ldots, X_{m:m:n}^{\mathscr{R}}$ be progressively Type-II censored order statistics from an absolutely continuous, strictly increasing cumulative distribution function F on $(0, \infty)$ with $F(0) = 0$ and with censoring scheme $\mathscr{R}$.

Suppose that F is NBU or NWU. Then, the population distribution is an Exp(ϑ)-distribution for some $\vartheta > 0$ iff integers r, s, and m, $1 \leq r < s \leq m$, exist such that

(i) $X^{\mathscr{R}}_{s:m:n} - X^{\mathscr{R}}_{r:m:n} \stackrel{\mathrm{d}}{=} X^{\mathscr{R}_{\lhd r}}_{s-r:m-r:\gamma_{r+1}}$, or
(ii) $EX^{\mathscr{R}}_{s:m:n} - EX^{\mathscr{R}}_{r:m:n} = EX^{\mathscr{R}_{\lhd r}}_{s-r:m-r:\gamma_{r+1}}$, assuming that all the expected values exist.

A similar characterization in terms of mean residual lifetime has been established by Tavangar and Hashemi [840].

3.1.3 Characterizations via Regression

In this section, we present characterizations of continuous cumulative distribution functions via regression -type relations

$$E(h(X_{r+l:m:n})|X_{r:m:n} = x) = g(x) \tag{3.2}$$

for given continuous functions h and g and fixed r and l, $1 \leq r \leq m-1$, $1 \leq l \leq m-r$. h is supposed to be strictly monotone. The problem is to determine all (continuous) distributions P^F satisfying the regression relation in (3.2). We formulate the following results in terms of progressively Type-II censored order statistics being aware that most of them also hold more generally for generalized order statistics. We follow the presentation in Cramer et al. [312].

The problem has been addressed for various forms of ordered data. In particular, the case $l = 1$ has received great attention. For order statistics, Ferguson [362] apparently studied it first when $h(t) = t$. Early references for record values are Nagaraja [660, 663]. Keseling [516, 517] considered regressions of generalized order statistics subject to some restrictions on the parameters which includes m-generalized order statistics and progressively Type-II censored order statistics. The reversed regression $E(X_{r:m:n}|X_{r+l:m:n} = x)$ is investigated in the literature for order statistics and record values as well (see, e.g., Ferguson [362], Nagaraja [663], and Franco and Ruiz [377, 378]). Keseling [516] established respective results for m-generalized order statistics. Bieniek [202] considered the reversed regression problem for adjacent generalized order statistics. The case $l = 2$ is considered in Pudeg [732] (order statistics) and Keseling [517] (m-generalized order statistics).

The problem in (3.2) has been considered for adjacent random variables, i.e., $l = 1$, in a more general framework. Namely, let $H : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a given bivariate function and $g : \mathbb{R} \longrightarrow \mathbb{R}$ be a known univariate function. Then, (3.2) can be written in the form

$$E(H(X_{r:m:n}, X_{r+1:m:n})|X_{r:m:n} = x) = g(x) \quad \text{for all } x$$

for some fixed $r \in \{1, \ldots, n-1\}$. Since a general solution is not possible, special choices of H and g will lead to the desired results (see Keseling [516, p. 32]).

Franco and Ruiz [377] considered order statistics when $H(s,t) = h(t)$ and h is continuous and strictly increasing. They characterized the continuous baseline cumulative distribution function F completely given h and g. Similar results were derived in Franco and Ruiz [378] for record values and continuous F. For a discrete cumulative distribution function, the characterization problem was solved by Nagaraja [664] for $h(x) = x$. This work was extended by Franco and Ruiz [379] to a continuous and strictly increasing h. Further results with different kinds of assumptions are provided by Rogers [757], Ferguson [362], Nagaraja [663], Khan and Abu-Salih [526], and Ouyang [696, 697]. Finally, Franco and Ruiz [380] presented a unifying approach to the described situation for arbitrary cumulative distribution function F.

For order statistics, Rao and Shanbhag [740] considered $H(s,t) = \phi(t-s)$ and a constant conditional expectation $c > 0$ and characterized the exponential distribution among the continuous distributions. A similar result for record values was obtained by Rao and Shanbhag [739].

Conditional Expectations and Characterization Problems

Cramer et al. [312] and Keseling [516] presented conditions such that the conditional expectation $E(h(X^{\mathscr{R}}_{r+l:m:n})|X^{\mathscr{R}}_{r:m:n} = \cdot)$ with a continuous strictly increasing function h specifies the distribution uniquely. Clearly, one has to consider only $h(x) = x$ because for continuous strictly increasing h

$$E(h(X^{\mathscr{R}}_{r+l:m:n})|X^{\mathscr{R}}_{r:m:n} = \cdot) = g \quad P^{X^{\mathscr{R}}_{r:m:n}} \text{ a.e.}$$

is equivalent to

$$E(Y^{\mathscr{R}}_{r+1:m:n}|Y^{\mathscr{R}}_{r:m:n} = \cdot) = g \circ h^{-1} \quad P^{Y^{\mathscr{R}}_{r:m:n}} \text{ a.e.,}$$

where $Y^{\mathscr{R}}_{1:m:n}, \ldots, Y^{\mathscr{R}}_{m:m:n}$ denote progressively Type-II censored order statistics from the cumulative distribution function

$$F^Y(y) = \begin{cases} 0, & y < h(\alpha(F)) \\ F(h^{-1}(y)), & h(\alpha(F)) \le y < h(\omega(F)) \\ 1, & h(\omega(F)) \le y \end{cases}$$

and with the same censoring scheme $\mathscr{R}$ (for order statistics, see Pudeg [732, p. 92]).

Cramer et al. [312] showed that the cumulative distribution function cannot be specified uniquely by an arbitrary version of the conditional expectation. A similar example for reversed regression of order statistics is given in Keseling [516].

However, a one-to-one correspondence between the conditional expectation ξ and the cumulative distribution function F holds if the conditional expectation ξ is supposed to be an increasing and continuous version of the conditional expectation. Lemma 3.1.8 states that such a version always exists (cf. Keseling [516, Lemma 2.2]) and is unique under weak conditions.

Lemma 3.1.8. Let $1 \le r \le m-1$ and $l \in \{1,\dots,m-r\}$. Let $X_{1:m:n},\dots,X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous distribution function F with censoring scheme $\mathscr{R}$. For $x \in (\alpha(F),\omega(F))$, let Z_x be a random variable such that $Z_x \sim P^{X_{r+l:m:n}|X_{r:m:n}=x}$. Moreover, let $g:\mathbb{R}\to\mathbb{R}$ be a (finite) version of the conditional expectation $E(X_{r+l:m:n}|X_{r:m:n}=\cdot)$. Then,

(i) EZ_x is an increasing and continuous version of the conditional expectation. Moreover, $EX_{r:m:n} \le EZ_x < \infty$ for all $x \in (\alpha(F),\omega(F))$;
(ii) If g is increasing and continuous, then the equation $g(x) = Z_x$ holds for all $x \in (\alpha(F),\omega(F))$. Hence, $EZ_{(\cdot)}$ is the unique increasing and continuous version of the conditional expectation $E(X_{r+l:m:n}|X_{r:m:n}=\cdot)$ on the support of F.

Remark 3.1.9. From the preceding results, it follows that there is a unique continuous and increasing version of the conditional expectation. This version is calculated as the expectation of the regular version of the conditional probability $P^{X_{r+l:m:n}|X_{r:m:n}=x}$. Using these properties, the underlying distribution is characterized. However, in general a unique characterization is not possible since the support of the distribution is generally not determined by the conditional expectation. A specification of the conditional expectation of progressively Type-II censored order statistics by an increasing and continuous function g does not lead to a unique specification of the support (see Cramer et al. [312]). To avoid this ambiguity, the support can be prescribed explicitly such that a unique solution results. Otherwise, the form of the support has to be discussed separately using properties of the regression function.

Adjacent Progressively Type-II Censored Order Statistics

The characterization result for the conditional expectation of adjacent progressively Type-II censored order statistics, i.e., $E(X_{r+1:m:n}|X_{r:m:n}=\cdot)$, is based on the approach of Franco and Ruiz [377] for order statistics. It is due to Keseling [516] who established it for generalized order statistics. First, an explicit formula for the cumulative distribution function F is given subject to a prescribed proper conditional expectation ξ. The second result answers the question which functions can be seen as conditional expectations of a progressively Type-II censored order statistic.

Let $1 \le r \le m-1$, $h:\mathbb{R}\longrightarrow\mathbb{R}$ be a continuous monotone function, and $f_{r+1|r:m:n}(\cdot|x)$ be the P^F-density of $P^{X_{r+1:m:n}|X_{r:m:n}=x}$ given in (2.37). Then, the order mean function $\xi^F_{\gamma_{r+1}}:(-\infty,\omega(F))\longrightarrow\mathbb{R}$ is defined by

$$\xi^F_{\gamma_{r+1}}(x) = \int_{[x,\infty)} h(t) f_{r+1|r:m:n}(t|x)\, dP^F(t), \quad x < \omega(F),$$

for $F \in \mathscr{F}_h^{\gamma_{r+1}}$, where

$$\mathscr{F}_h^{\gamma_{r+1}} = \Big\{ F \mid F \text{ is a continuous cumulative distribution function such that}$$
$$\int_{[x,\infty)} |h(t)|(1-F(t))^{\gamma_{r+1}-1}\, dP^F(t) < \infty \text{ for all } x \in \mathbb{R} \Big\}$$

(see Franco and Ruiz [377]). Notice that, according to Lemma 3.1.8, $\xi^F_{\gamma_{r+1}}$ is the continuous increasing continuation of the conditional expectation $E(X_{r+1:m:n}|X_{r:m:n} = \cdot)$ on the interval $(-\infty, \omega(F))$. If $(-\infty, \alpha(F)] \neq \emptyset$, then the function $\xi^F_{\gamma_{r+1}}(x) = \gamma_{r+1} \int h(t)(1-F(t))^{\gamma_{r+1}-1}\, dP^F(t)$ is defined to be constant for all $x \in (-\infty, \alpha(F)]$. The set of admissible functions $\xi^F_{\gamma_{r+1}}$ is denoted by

$$\mathscr{M}_h^{\gamma_{r+1}} = \Big\{ \xi \mid \text{There exists } F \in \mathscr{F}_h^{\gamma_{r+1}} \text{ such that } \xi = \xi^F_{\gamma_{r+1}} \Big\}.$$

Properties of $\xi \in \mathscr{M}_h^{\gamma_{r+1}}$ are given in Lemma 3.1.10 which is due to Keseling [516]. In particular, it shows the connection to the setting of order statistics discussed in Franco and Ruiz [377].

Lemma 3.1.10. Let $1 \le r \le m-1$, $\mathbb{D} \subseteq \mathbb{R}$, $\xi : \mathbb{D} \longrightarrow \mathbb{R}$, h be a continuous and monotone function and $h(\beta) = \lim_{x\to\beta-} h(x)$.

Then,

$$\xi \in \mathscr{M}_h^{\gamma_{r+1}} \quad \text{iff} \quad \xi \in \mathscr{M}_h^{n-r}.$$

In this case, ξ has the following properties:

(i) $\mathbb{D} = (-\infty, \beta)$ for some $\beta \in (-\infty, \infty]$;
(ii) ξ is continuous on $\mathbb{D}$;
(iii) if h is increasing, then $h(\beta) > \xi(x) > h(x)$ for all $x \in \mathbb{D}$;
if h is decreasing, then $h(\beta) < \xi(x) < h(x)$ for all $x \in \mathbb{D}$;
(iv) ξ is increasing (decreasing) if h is increasing (decreasing);
(v) the Riemann–Stieltjes integral

$$\int_{-\infty}^{x} \frac{1}{\xi(t) - h(t)}\, d\xi(t)$$

converges for any $x \in \mathbb{D}$. It converges to infinity for $x \to \beta$;
(vi) if $\mathbb{D} = \mathbb{R}$, then

$$\lim_{x\to\infty} \xi(x) \exp\Big\{ \int_{-\infty}^{x} \frac{1}{\xi(t) - h(t)}\, d\xi(t \Big\} = 0.$$

The following theorem establishes the desired inversion result which relates a function $\xi \in \mathscr{M}_h^{\gamma_{r+1}}$ uniquely to a cumulative distribution function F. For kth record values, the corresponding result is given in Grudzień and Szynal [415]. For order statistics, Khan and Abu-Salih [526] established an inversion formula subject to some restrictive smoothness assumptions. For a proof, one may refer to Cramer et al. [312].

Theorem 3.1.11. Let $1 \le r \le m-1$, $\mathbb{D} \subseteq \mathbb{R}$, $\xi : \mathbb{D} \longrightarrow \mathbb{R}$ with $\xi \in \mathscr{M}_h^{\gamma_{r+1}}$ for $\gamma_{r+1} > 0$, h be a continuous and monotone function and $F \in \mathscr{F}_h^{\gamma_{r+1}}$ with $\xi = \xi_{\gamma_{r+1}}^F$. Then,

(i) $\mathbb{D} = (-\infty, \omega(F))$;
(ii) $F(x) = 1 - \exp\left\{-\dfrac{1}{\gamma_{r+1}} \displaystyle\int_{-\infty}^x \frac{1}{\xi(t) - h(t)}\, d\xi(t)\right\}$, $x \in \mathbb{D}$.

Now, we apply Theorem 3.1.11 to the linear regression problem. Suppose now that $h(x) = x$, $x \in \mathbb{R}$, and

$$\xi(x) = \begin{cases} a\alpha + b, & x < \alpha \\ ax + b, & \alpha \le x < \beta \end{cases}, \qquad \beta \in (-\infty, \infty].$$

where $\alpha < \beta$ and a, b are given real numbers:

$$E(X_{r+l:m:n} | X_{r:m:n} = \cdot) = \xi \quad P^{X_{r:m:n}} \text{ a.e.}$$

It should be mentioned that ξ is defined for $x < \alpha$ as a constant in order to meet the conditions given in Lemma 3.1.10. Since h is an increasing function, ξ must be increasing according to Part (iv) of Lemma 3.1.10. Thus, $a > 0$. This proves that the generalized Pareto distributions are characterized by the linear regression property of progressively Type-II censored order statistics.

Corollary 3.1.12. (i) Let $a = 1$. Then, $\xi \in \mathscr{M}_h^{\gamma_{r+1}}$ iff $b > 0$ and $\beta = \infty$.
The cumulative distribution function is that of a two-parameter exponential distribution, i.e.,

$$F(x) = 1 - \exp\left\{-\frac{x-\alpha}{\gamma_{r+1} b}\right\}, \quad x \ge \alpha;$$

(ii) Let $0 < a < 1$. Then, $\xi \in \mathscr{M}_h^{\gamma_{r+1}}$ iff $\alpha(F) = \alpha < \frac{b}{1-a} = \beta = \omega(F) < \infty$.
The cumulative distribution function is that of a reflected two-parameter power function distribution, i.e.,

$$F(x) = 1 - \left(\frac{\beta - x}{\beta - \alpha}\right)^\theta, \quad \alpha \le x \le \beta,$$

where $\theta = \frac{a}{\gamma_{r+1}(1-a)}$;

$h = -\frac{1}{\lambda}\log(1-G)$	(α, β)	$G(x), x \in (\alpha, \beta)$	References
$-\frac{1}{\lambda}\log\left(1-\left(\frac{x}{\theta}\right)^{\lambda}\right)$	$(0, \theta)$	$\left(\frac{x}{\theta}\right)^{\lambda}$ (power-function)	
$\log\left(\frac{x}{\theta}\right)$	(θ, ∞)	$1-\left(\frac{\theta}{x}\right)^{\lambda}$ (Pareto)	Ouyang [697]
x^{θ}	$(0, \infty)$	$1-\exp(-\lambda x^{\theta})$ (Weibull)	Mohie El-Din et al. [653]
$\frac{1}{\lambda}\log(1+\exp(\lambda x))$	$\mathbb{R}$	$(1+\exp(-\lambda x))^{-1}$ (logistic)	Khan and Abu-Salih [526], Ouyang [697]
$\log(1+x^{\theta})$	$(0, \infty)$	$1-(1+x^{\theta})^{-\lambda}$ (Burr XII)	Ouyang [697]

Table 3.1 Characterizations for order statistics related to the characterization of the exponential distribution given in Corollary 3.1.12 ($\theta > 0$)

(iii) Let $1 < a$. Then, $\xi \in \mathscr{M}_h^{\gamma_{r+1}}$ iff $\alpha(a-1) > -b$ and $\beta = \infty$.
The cumulative distribution function is that of a Pareto distribution, i.e.,

$$F(x) = 1 - \left(\frac{\alpha + \frac{b}{a-1}}{x + \frac{b}{a-1}}\right)^{\theta}, \quad \alpha \le x < \infty,$$

where $\theta = \frac{a}{\gamma_{r+1}(a-1)}$.

Remark 3.1.13. It should be noted that the type of characterized distributions does not depend on the parameters of the underling progressively Type-II censored order statistics except for γ_{r+1}. This is due to the Markov property (see Theorem 2.5.1).

For order statistics and record values, the results of Corollary 3.1.12 were established by Nagaraja [660, 663], wherein γ_{r+1} is given by $n-r$ and 1, respectively. The parameters of the power function and Pareto distribution read $\theta = \frac{a}{(n-r)|a-1|}$ and $\theta = \frac{a}{|a-1|}$ for order statistics and record values, respectively.

Many characterizations using the regression approach can be found in the literature. For illustration, Table 3.1 taken from Cramer et al. [312] subsumes some results (cf. Keseling [516, p. 45]). It shows characterization results based on the preceding characterization of the exponential distribution via the regression

$$E(h(X_{r+1:m:n})|X_{r:m:n} = x) = h(x) + \frac{1}{\gamma_{r+1}\lambda} \quad P^G \text{ a.e.},$$

where $h : (\alpha, \beta) \longrightarrow (0, \infty)$, $-\infty \le \alpha < \beta \le \infty$, is a strictly increasing and continuous function with $\lim_{x\to\alpha} h(x) = 0$ and $\lim_{x\to\beta} h(x) = \infty$ ($\gamma_{r+1}, \lambda > 0$). The references are related to the corresponding results for order statistics.

Finally, it should be mentioned that many results known in the literature are monotone transformations of the characterizations given in Corollary 3.1.12 which for order statistics are due to Ferguson [362]. Results of that type are given in Dallas

[321], Khan and Khan [527, 528], Khan and Abu-Salih [525, 526], Mohie El-Din et al. [653], and Ouyang [697]. For further results on characterizations, one may refer to Keseling [516].

Progressively Type-II Censored Order Statistics Based on Higher-Order Gap

Up to now, conditional expectations $E(X_{r+l:m:n}|X_{r:m:n} = \cdot)$ with adjacent progressively Type-II censored order statistics have been considered. Now , we address a regression based on higher-order gap, i.e., we assume that the distance l between the considered progressively Type-II censored order statistics may be larger than one, i.e., $1 < l \leq m - r$.

The characterization result given in Theorem 3.1.14 is derived along the lines of the respective result in Dembińska and Wesołowski [333] using integrated Cauchy functional equations (cf. Rao and Shanbhag [740, Theorem 2.2.2]). For restricted generalized order statistics, this result can be found in Bieniek and Szynal [204] and, in the most general setting, in Cramer et al. [312]. Detailed proofs can be found in these references.

Theorem 3.1.14. Let $1 \leq r \leq m-1$, $1 \leq l \leq m-r$, and $X_{1:m:n}, \ldots X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F with censoring scheme $\mathscr{R}$.

If constants $a > 0$ and $b \in \mathbb{R}$ exist such that

$$E(X_{r+l:m:n}|X_{r:m:n} = x) = ax + b \quad P^F \text{ a.e.},$$

then F is the cumulative distribution function of a generalized Pareto distribution. Thus, up to an affine transformation of the argument, it is given by one of the following cumulative distribution functions:

(i) For $a = 1$, $F(x) = 1 - \exp(-x)$, $x \geq 0$;
(ii) For $0 < a < 1$, $F(x) = 1 - (-x)^\theta$, $x \in [-1, 0]$;
(iii) For $1 < a$, $F(x) = 1 - x^\theta$, $x \in [1, \infty)$.

The parameter θ is given by $\theta = -\frac{1}{\eta}$, where η is the unique solution in η of the polynomial equation

$$\prod_{j=r+1}^{r+l} (\gamma_j - \eta) = \frac{1}{a} \prod_{j=r+1}^{r+l} \gamma_j, \quad \eta \in (-\infty, \gamma_{r+l}).$$

Remark 3.1.15. For order statistics, the result in Theorem 3.1.14 was derived by Dembińska and Wesołowski [333]. Pudeg [732] considered the case $l = 2$. López Blázquez and Moreno Rebollo [617] provided a different proof of the characterization assuming that the baseline cumulative distribution functions are

l-times differentiable. The Cauchy functional equation approach is in favor of this one because the regression problem is solved for any continuous cumulative distribution function (see Cramer et al. [312]). Dembińska and Wesołowski [334] successfully applied the preceding approach to record values as well.

The characterization of the exponential distribution given in Theorem 3.1.14 can be rewritten in the form ($a = 1$)

$$E(X_{r+l:m:n} - X_{r:m:n} | X_{r:m:n} = x) = b \quad P^F \text{ a.e.}$$

with some constant $b > 0$. This observation leads to a different characterization of the exponential distribution via the regression

$$E(X_{r+l+\nu:m:n} - X_{r+\nu:m:n} | X_{\nu:m:n} = x) = b \quad P^F \text{ a.e.}$$

for some r, l, ν. A proof can be found in Cramer and Kamps [302].

Theorem 3.1.16. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from an absolutely continuous cumulative distribution function F with density function f and with censoring scheme $\mathscr{R}$. Suppose for $r, l, \nu \in \mathbb{N}$ with $r + l + \nu \le n$ the expectation $E|X_{r+l+\nu:m:n}|$ is finite.

If a constant $b > 0$ exists such that

$$E(X_{r+l+\nu:m:n} - X_{r+\nu:m:n} | X_{\nu:m:n} = \cdot) = b \quad P^F \text{ a.e.},$$

then

(i) for any $j \in \{1, \ldots, r\}$

$$E(X_{r+l+\nu:m:n} - X_{r+\nu:m:n} | X_{\nu+j:m:n} = \cdot) = b \quad P^F \text{ a.e.;}$$

(ii) F is the cumulative distribution function of a two-parameter exponential distribution.

Reversed Regression of Adjacent Progressively Type-II Censored Order Statistics

For $1 \le r \le m - 1$, Bieniek [202] has considered the reversed regression problem

$$E(h(X_{r:m:n}) | X_{r+1:m:n} = x) = g(x)$$

for generalized order statistics. He provided a characterization in terms of the strictly increasing function

$$h_{r+1}(x) = (1 - x)^{1-\gamma_{r+1}} f^{U_{r+1:m:n}}(x), \quad x \in (0, 1).$$

From the representation of $f^{U_{r+1:m:n}}$ in (2.24) it follows that

$$h_{r+1}(1) = \lim_{x \to 1-} h_{r+1}(x) = \gamma_{r+1} \prod_{j=1}^{r} \frac{\gamma_j}{\gamma_j - \gamma_{r+1}}$$

(see also Cramer et al. [313, Eq. (2.6)]). The corresponding cumulative distribution function is given by

$$F(x) = h_{r+1}^{-1}\left(h_{r+1}(1) \exp\left\{-\int_x^\beta \frac{dg(y)}{h(y) - g(y)}\right\}\right), \quad x \le \beta,$$

where $\beta = \omega(F)$ denotes the right endpoint of support of F. For linear regression, i.e., $g(x) = ax + b$, Bieniek [202] proved the following theorem.

Theorem 3.1.17. Let $1 \le r \le m-1$, and $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F with censoring scheme $\mathscr{R}$. (α, β) denotes the support of F.

If constants $a > 0$ and $b \in \mathbb{R}$ exist such that

$$E(X_{r:m:n} | X_{r+1:m:n} = x) = ax + b \quad P^F \text{ a.e.},$$

then F is one of the following cumulative distribution functions:

(i) For $a = 1$, $\alpha = -\infty$ and $F(x) = h_{r+1}^{-1}\big(h_{r+1}(1)\exp(-(x-\beta)/b)\big)$, $x < \beta$, $\beta \in \mathbb{R}$;
(ii) For $0 < a < 1$, $\alpha = b/(1-c)$ and $F(x) = h_{r+1}^{-1}\Big(h_{r+1}(1)\big(\frac{x-\alpha}{\beta-\alpha}\big)^\theta\Big)$, $x \in [\alpha, \beta]$, $\beta > \alpha$, $\theta = a/(1-a)$;
(iii) For $1 < a$, $\alpha = -\infty$ and $F(x) = h_{r+1}^{-1}\Big(\frac{A}{(\beta-x)^\theta}\Big)$, $x < \beta$, $\beta = b/(1-a)$, $\theta = a/(a-1)$, $A > 0$.

Notice that, for order statistics, $h_{r+1}(u) = (r+1)\binom{n}{r+1}u^r$. Then, the result of Theorem 3.1.17 corresponds to those given in Ferguson [362] and Franco and Ruiz [377, Remark 5.8].

3.1.4 Characterizations for Discrete Parents

Characterizations of distributions by means of progressively Type-II censored order statistics from discrete parents have been discussed in Tran [854] and Balakrishnan et al. [149] with the former dealing with generalized order statistics. Tran's results can be applied to discrete progressively Type-II censored order statistics using the quantile representation of progressively Type-II censored order statistics due to Balakrishnan and Dembińska [96, 97] (see also Sect. 2.8).

As for a continuous cumulative distribution function F, we organize the characterizations by their type, i.e., characterizations via *regression, independence*, and *distribution properties*. Without going into details, we mention that a basic tool in deriving them is Shanbhag's Lemma. Each section states first a result due to Tran [854] followed by results given in Balakrishnan et al. [149].

Theorem 3.1.18 (Rao and Shanbhag [740]). Let $\{(v_n, w_n) : n \in \mathbb{N}_0)\}$ be a sequence of vectors with nonnegative real components such that $v_n \neq 0$ at least for one $n \in \mathbb{N}$ and $w_1 \neq 0$. Then,

$$v_n = \sum_{m=0}^{\infty} w_m v_{m+n}, \quad n \in \mathbb{N}_0,$$

iff $\sum_{m=0}^{\infty} w_m b^m = 1$ and $v_n = v_0 b^n$, $n \in \mathbb{N}$, for some $b > 0$.

Characterization by Regression

For continuous cumulative distribution function, a constant regression characterizes the exponential distribution (see Theorem 3.1.14). For discrete distributions, we get a corresponding result leading to geometric-type distributions.

Theorem 3.1.19 (Tran [854]). Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from a discrete cumulative distribution function F with support $\mathbb{N}_0$. Then:

$$E\left(X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n} | X^{\mathscr{R}}_{i:m:n}, X^{\mathscr{R}}_{i+1:m:n} > X^{\mathscr{R}}_{i:m:n}\right) = c \quad \text{a.e.} \tag{3.3}$$

for some $c \in \mathbb{R}$ and $1 \leq i < m$ iff F is a modified geometric distribution with cumulative distribution function $F(0) = \theta$ and $F(j) = 1 - (1-\theta)(1-p)^j$, $j \in \mathbb{N}$, for some $\theta \in (0,1)$ and $p \in (0,1)$.

A generalization of the characterizing identity (3.3) is due to Balakrishnan et al. [149]. The conditioning event $\left\{X^{\mathscr{R}}_{i+1:m:n} > X^{\mathscr{R}}_{i:m:n}\right\}$ in (3.3) is replaced by $\left\{X^{\mathscr{R}}_{i:m:n} > X^{\mathscr{R}}_{i:m:n} + l\right\}$ with some fixed $l \in \mathbb{N}_0$. Furthermore, $\Phi\left(X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n}\right)$ replaces $X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n}$, where $(\Phi(k))_k$ is a suitably chosen sequence. For order statistics, the result has been proved by Rao and Shanbhag [741] for $l = 0$ and by Nagaraja [664] for $\Phi(j) - j \equiv 0$.

Theorem 3.1.20. Let $X^{\mathscr{R}}_{1:m:n}, \ldots X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from cumulative distribution function F with support $\mathbb{N}_0$. Fix $l \in \mathbb{N}_0$ and suppose $(\Phi(k))_{k \geq l}$ is a nondecreasing sequence such that $\Phi(l+2) > \Phi(l+1)$.

Then,

$$E\left[\Phi\left(X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n}\right) \mid X^{\mathscr{R}}_{i:m:n}, X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n} > l\right] = c \quad \text{a.e.} \tag{3.4}$$

for some $c \in \mathbb{R}$ and $1 \le i < m$ iff F is a modified geometric distribution with $F(l+j) = 1-(1-\theta)q^j$, $j \in \mathbb{N}_0$, for some $\theta \in (0,1)$ and $q \in (0,1)$.

For $l < 0$, the condition in (3.4) reads $E\left[\Phi\left(X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n}\right) \mid X^{\mathscr{R}}_{i:m:n}\right] = c$ a.e. for some $c \in \mathbb{R}$. As pointed out by Nagaraja [664], the geometric distribution has the property

$$E(X_{i+1:n} - X_{i:n}|X_{i:n}) = c \quad \text{a.e. for some } c \in \mathbb{R} \tag{3.5}$$

iff $i = 1$. This observation has been utilized by López-Blázquez and Miño [616]. They showed that if (3.5) holds with $i = 1$ and the underlying cumulative distribution function has support on $\mathbb{N}_0$, then F is geometric.

Theorem 3.1.21. Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from cumulative distribution function F with support $\{0, 1, \ldots, N\}$ for some $1 \le N \le \infty$. Then,

$$E\left(X^{\mathscr{R}}_{2:m:n} - X^{\mathscr{R}}_{1:m:n}|X^{\mathscr{R}}_{1:m:n}\right) = c \quad \text{a.e. for some } c \in \mathbb{R}$$

iff F is geometric, that is, $F(j) = 1 - q^{j+1}$, $j \in \mathbb{N}_0$, for some $q \in (0,1)$.

Characterization by Distribution Properties

The first characterization utilizes a condition on the hazard rate of discrete distribution defined by

$$\lambda_F(t) = \frac{P(X=t)}{P(X \ge t)} = \frac{F(t) - F(t-)}{1 - F(t-)}, \quad t \in \mathbb{N}_0.$$

Notice that $\lambda_F(t) = P(X = 0)$, $t \in \mathbb{N}_0$, is constant for a geometric distribution (see Gupta et al. [425]).

Theorem 3.1.22 (Tran [854]). Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from a discrete cumulative distribution function F with support $\mathbb{N}_0$, and let $\mathscr{R}_{\lhd i} = (R_{i+1}, \ldots, R_m)$ denote a left truncated censoring scheme.

If $\lambda_F(0) \ge (\le) \lambda_F(t)$ for all $x \in \mathbb{N}_0$, then for some $1 \le i < j \le m$, the conditional distribution of the spacing $X^{\mathscr{R}}_{j:m:n} - X^{\mathscr{R}}_{i:m:n}$, given the event $\{X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n} > 0\}$, is the same as the unconditional distribution of $X^{\mathscr{R}_{\lhd i}}_{j-i:m-i:\gamma_{i+1}} + 1$ iff F is a geometric cumulative distribution function.

The next characterization provides an extension of results for order statistics to progressively Type-II censored order statistics (see Puri and Rubin [734] and Zijlstra [948]). It has been established by Balakrishnan et al. [149].

Theorem 3.1.23. Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from cumulative distribution function F with support being a subset of $\mathbb{N}_0$. Assume $P(X = 1) > 0$ and $P(X > 1) > 0$. Then, for some $1 \le i < m$,

$$X^{\mathscr{R}}_{i+1:m:n} - X^{\mathscr{R}}_{i:m:n} \stackrel{\mathsf{d}}{=} X^{\mathscr{R}_{\lhd i}}_{1:m-i:\gamma_{i+1}}$$

iff F is a modified geometric distribution with cumulative distribution function $F(j) = 1 - (1 - p_0)q^j$, $j \ge 0$, where $p_0 \in (0, 1)$ and $q \in (0, 1)$ are such that

$$1 = \left(\prod_{j=1}^{i} \gamma_j\right) \sum_{s=1}^{i} \left[\frac{a_{s,i}}{\gamma_s - \gamma_{i+1}} \left(1 - (1 - p_0)^{\gamma_s - \gamma_{i+1}} \frac{1 - q^{\gamma_{i+1}}}{1 - q^{\gamma_s}}\right)\right].$$

Characterization by (Conditional) Independence

The following characterizations are based on independence properties. For order statistics, the first result was established by Nagaraja and Srivastava [669].

Theorem 3.1.24 (Tran [854]). Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from a discrete cumulative distribution function F with support $\mathbb{D}$. Then, if $|\mathbb{D}| \ge 3$ and for some $2 \le i < j \le m$, the random variable $X^{\mathscr{R}}_{i:m:n}$ and the event $\{X^{\mathscr{R}}_{j:m:n} = X^{\mathscr{R}}_{i:m:n}\}$ are conditionally independent, given the event $\{X^{\mathscr{R}}_{i:m:n} > X^{\mathscr{R}}_{i-1:m:n}\}$, F is a modified geometric distribution with $\mathbb{D} = \{a_i \in \mathbb{R} | i \in \mathbb{N}_0$ and $a_i < a_j$ for $i < j\}$, $F(a_0) = \theta$ and $F(a_j) = 1 - (1 - \theta)(1 - p)^j$, $j \in \mathbb{N}$, for some $\theta \in (0, 1)$ and $p \in (0, 1)$.

The next characterization is a generalization of a result due to Govindarajulu [412].

Theorem 3.1.25. Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from cumulative distribution function F with support $\mathbb{N}_0$. Then, for a fixed $k \ge 2$ and $1 < i \le m$, the random variable $X^{\mathscr{R}}_{1:m:n}$ and the event $\{X^{\mathscr{R}}_{i:m:n} - X^{\mathscr{R}}_{1:m:n} \ge k\}$ are independent and

$$\overline{F}(j) = q^{j+1}, \ \ 0 \le j < k,$$

iff F is a geometric distribution.

Theorem 3.1.26 extends Theorem 3.1 of El-Neweihi and Govindarajulu [350] (for order statistics with $s = n$, see Galambos [388]). An extension to generalized order statistics has been established by Tran [854] under a stronger assumption.

Theorem 3.1.26. Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from cumulative distribution function F with support $\{a_0, a_1, \ldots\}$, where $(a_k)_{k\geq 0}$ is an increasing sequence of real numbers. Then, the random variable $X^{\mathscr{R}}_{1:m:n}$ and the event $\{X^{\mathscr{R}}_{i:m:n} = X^{\mathscr{R}}_{1:m:n}\}$ are independent for some $1 < i \leq m$ iff F is a geometric-type distribution, i.e., $\overline{F}(a_j) = q^{j+1}$, $j \in \mathbb{N}_0$, with $q \in (0,1)$.

3.2 Stochastic Ordering of Progressively Type-II Censored Order Statistics

An important notion in reliability theory is stochastic ordering, which has received great attention in comparing ordered random variables. In particular, many results have been established for generalized order statistics and progressively Type-II censored order statistics extending properties for order statistics. In this section, we are going to review results obtained for progressively Type-II censored order statistics. For surveys on stochastic ordering of order statistics and record values, we refer to the monographs of Shaked and Shanthikumar [799] and Müller and Stoyan [659] and, in particular, to the articles of Arnold and Villaseñor [56] and Boland et al. [212] in the *Handbook of Statistics* Vol. 16 and Vol. 17 by Balakrishnan and Rao [116, 117] devoted to order statistics, as well as to the references cited in these works.

Details on stochastic ordering, failure rate ordering, likelihood ratio ordering, and dispersive ordering can be found in many references including Shaked and Shanthikumar [799], Ross [759, Chap. 9], and Müller and Stoyan [659]. The definitions of these orders are given in the Appendix A.2.2.

In the following, let X and Y be random variables with continuous cumulative distribution functions F and G, respectively, satisfying

$$F^{-1}(0+), G^{-1}(0+) \geq 0,$$

i.e., their supports are contained in the positive real line. Sometimes, the cumulative distribution functions are supposed to be absolutely continuous with density functions f and g.

3.2.1 Univariate Stochastic Orders and Its Applications to Progressively Type-II Censored Order Statistics

Stochastic Order

The following result has been established by Khaledi [518] in terms of generalized order statistics. In terms of the parametrization of progressively Type-II censored order statistics, he assumed that, for censoring schemes $\mathscr{R}, \mathscr{S}$ and $j \leq i$, the condition

$$(\gamma_1(\mathscr{S}),\dots,\gamma_j(\mathscr{S})) \preccurlyeq_{\mathsf{p}} (\gamma_{\ell_1}(\mathscr{R}),\dots,\gamma_{\ell_j}(\mathscr{R})) \quad \text{for some set } \{\ell_1,\dots,\ell_j\} \subseteq \{1,\dots,i\}, \tag{3.6}$$

holds. However, since in progressive censoring the parameters are ordered, e.g.,$\gamma_1(\mathscr{R}) \geq \dots \geq \gamma_{m_1}(\mathscr{R})$, (3.6) is, for $j \leq i$, equivalent to

$$(\gamma_1(\mathscr{S}),\dots,\gamma_j(\mathscr{S})) \preccurlyeq_{\mathsf{p}} (\gamma_{i-j+1}(\mathscr{R}),\dots,\gamma_i(\mathscr{R})). \tag{3.7}$$

In the following, we use the notation $n_1 = \gamma_1(\mathscr{R}), n_2 = \gamma_2(\mathscr{S})$.

Theorem 3.2.1. Let F, G be continuous cumulative distribution functions with $F \leq_{\mathsf{st}} G$ and $X \sim F, Y \sim G$. Moreover, let $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}, \mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$ with $m_1, m_2 \in \mathbb{N}$ be censoring schemes. Then,

(i) $X^{\mathscr{R}}_{i:m_1:n_1} \leq_{\mathsf{st}} Y^{\mathscr{R}}_{i:m_1:n_1}$, $1 \leq i \leq m_1$;
(ii) if $j \leq i$ and (3.7) holds, then $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{st}} Y^{\mathscr{R}}_{i:m_1:n_1}$.

Proof. We first prove (i). For brevity, we suppress the indices and write n and m instead of n_1 and m_1, respectively. Denote by $Z^{\mathscr{R}}_{i:m:n}$ an exponential progressively Type-II censored order statistic with censoring scheme $\mathscr{R}$. Then, the cumulative distribution function of $X^{\mathscr{R}}_{i:m:n}$ and $Y^{\mathscr{R}}_{i:m:n}$ are given by $F^{X^{\mathscr{R}}_{i:m:n}} = H^{Z^{\mathscr{R}}_{i:m:n}} \circ (-\log \overline{F})$ and $F^{Y^{\mathscr{R}}_{i:m:n}} = H^{Z^{\mathscr{R}}_{i:m:n}} \circ (-\log \overline{G})$, respectively.

Since $F \leq_{\mathsf{st}} G$, we have $\overline{F}(t) \leq \overline{G}(t)$ so that $-\log \overline{F}(t) \geq -\log \overline{G}(t)$, $t \in \mathbb{R}$. Thus, for $t \in \mathbb{R}$,

$$\overline{F}^{X^{\mathscr{R}}_{i:m:n}}(t) = \overline{F}^{Z^{\mathscr{R}}_{i:m:n}}(-\log \overline{F}(t)) \leq \overline{F}^{Z^{\mathscr{R}}_{i:m:n}}(-\log \overline{G}(t)) = \overline{F}^{Y^{\mathscr{R}}_{i:m:n}}(t),$$

because $\overline{F}^{Z^{\mathscr{R}}_{i:m:n}}$ is a nonincreasing function.

The proof of (ii) is as follows. From Corollary 3.2.4, we obtain $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{hr}} X^{\mathscr{R}}_{i:m_1:n_1}$ which implies $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{st}} X^{\mathscr{R}}_{i:m_1:n_1}$ (see Fig. A.1). Using the transitivity of the stochastic order and (i), we obtain directly the desired result. □

Property (i) can also be deduced from Theorem 3.2.21 by marginalization.

Corollary 3.2.2. Under the conditions of Theorem 3.2.1, the following result holds:

If $(\gamma_1(\mathscr{S}),\dots,\gamma_j(\mathscr{S})) \preccurlyeq_{\mathsf{p}} (\gamma_1(\mathscr{R}),\dots,\gamma_j(\mathscr{R}))$, then $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{st}} Y^{\mathscr{R}}_{j:m_1:n_1}$.

Since $\gamma_k(\mathscr{S}) \geq \gamma_k(\mathscr{R})$, $k = 1,\dots,j$, implies p-majorization of the respective vectors, the preceding corollary implies Theorem 5.4.2 of Burkschat [226].

Hazard Rate Order

Theorem 3.2.3. Let F, G be continuous cumulative distribution functions with $F \le_{\mathsf{hr}} G$ and $X \sim F, Y \sim G$. Moreover, let $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}, \mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$ with $m_1, m_2 \in \mathbb{N}$ be censoring schemes. Then,

(i) $X^{\mathscr{R}}_{i:m_1:n_1} \le_{\mathsf{hr}} Y^{\mathscr{R}}_{i:m_1:n_1}$, $1 \le i \le m_1$;
(ii) if $j \le i$ and condition (3.7) holds, then $X^{\mathscr{S}}_{j:m_2:n_2} \le_{\mathsf{hr}} Y^{\mathscr{R}}_{i:m_1:n_1}$.

Proof. By analogy with the proof of Theorem 3.2.1, we make use of the representations $F_{i,*} = F^{X^{\mathscr{R}}_{i:m:n}} = F^{Z^{\mathscr{R}}_{i:m:n}}(-\log \overline{F})$ and $G_{i,*} = F^{Y^{\mathscr{R}}_{i:m:n}} = F^{Z^{\mathscr{R}}_{i:m:n}}(-\log \overline{G})$ (indices are suppressed for brevity). From Theorem 2.7.7, we know that the density function of a progressively Type-II censored order statistic from an exponential distribution is log-concave. Thus, $\overline{H}_{i,*} = \overline{F}^{Z^{\mathscr{R}}_{i:m:n}}$ is a differentiable and log-concave function. Then, we find for $x < y$,

$$\begin{aligned}
\frac{\overline{G}_{*,i}(x)}{\overline{F}_{*,i}(x)} &= \frac{\overline{H}_{*,i}(-\log \overline{G}(x))}{\overline{H}_{*,i}(-\log \overline{F}(x))} = \frac{\overline{H}_{*,i}(-\log \overline{G}(x))}{\overline{H}_{*,i}\left(-\log \overline{G}(x) + \log \frac{\overline{G}(x)}{\overline{F}(x)}\right)} \\
&\le \frac{\overline{H}_{*,i}(-\log \overline{G}(y))}{\overline{H}_{*,i}\left(-\log \overline{G}(y) + \log \frac{\overline{G}(x)}{\overline{F}(x)}\right)} \\
&\le \frac{\overline{H}_{*,i}(-\log \overline{G}(y))}{\overline{H}_{*,i}\left(-\log \overline{G}(y) + \log \frac{\overline{G}(y)}{\overline{F}(y)}\right)} = \frac{\overline{H}_{*,i}(-\log \overline{G}(y))}{\overline{H}_{*,i}(-\log \overline{F}(y))} = \frac{\overline{G}_{*,i}(y)}{\overline{F}_{*,i}(y)}.
\end{aligned}$$

The first inequality can be seen as follows: $F \le_{\mathsf{hr}} G$ implies $F \le_{\mathsf{st}} G$ such that $\frac{\overline{F}(t)}{\overline{G}(t)} \le 1$. Thus, $a = \log \frac{\overline{G}(t)}{\overline{F}(t)} \ge 0$. Since $\overline{H}_{*,i}$ is a differentiable and log-concave function, $v = \log \overline{H}_{*,i}$ is differentiable and concave such that v' is decreasing. In particular, $v'(t) \ge v'(t+a)$ for $a \ge 0$. Hence, $w(t) = v(t) - v(t+a)$ has (for fixed $a \ge 0$) a nonnegative derivative

$$w'(t) = v'(t) - v'(t+a) \ge 0.$$

Thus, w is nondecreasing in t which yields the first inequality. In the last inequality, we use the fact that $F \le_{\mathsf{hr}} G$, i.e., $\frac{\overline{G}(t)}{\overline{F}(t)}$ is nondecreasing in t, and that $\overline{H}_{*,i}$ is decreasing.

In order to prove (ii), we consider first exponential progressively Type-II censored order statistics. According to the proof of Theorem 3.2.11, condition (3.7) yields $Z^{\mathscr{S}}_{j:m_2:n_2} \le_{\mathsf{disp}} Z^{\mathscr{R}}_{i:m_1:n_1}$. Then, using a theorem due to Bagai and Kochar [66] (see also Shaked and Shanthikumar [799, Theorem 3.B.20 (b)]), i.e., $X \le_{\mathsf{hr}} Y$ if $X \le_{\mathsf{disp}} Y$ and either the cumulative distribution function of X or Y is IFR, we obtain $Z^{\mathscr{S}}_{j:m_2:n_2} \le_{\mathsf{hr}} Z^{\mathscr{R}}_{i:m_1:n_1}$. Notice that the convolution of exponentials has an IFR distribution (cf. Theorem 3.3.2).

Finally, we apply the fact that the hazard rate order is preserved under increasing transformations (cf. Shaked and Shanthikumar [799, Theorem 1.B.2]). Therefore, $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{hr}} X^{\mathscr{R}}_{i:m_1:n_1}$ so that (i) and the transitivity of the hazard rate order yields the desired result. □

The first result (i) is due to Hu and Zhuang [457] in terms of generalized order statistics (for absolutely continuous cumulative distribution functions, see Belzunce et al. [188] and Khaledi [518]). A version of the second result can be found in Khaledi [518].

Theorem 3.2.3 yields directly the following result upon choosing $F = G$.

Corollary 3.2.4. Let $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}, \mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$ with $m_1, m_2 \in \mathbb{N}$ be censoring schemes and $X^{\mathscr{S}}_{j:m_2:n_2}, X^{\mathscr{R}}_{i:m_1:n_1}$ be progressively Type-II censored order statistics from the same continuous cumulative distribution function F. If $j \leq i$ and condition (3.7) holds, then $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{hr}} X^{\mathscr{R}}_{i:m_1:n_1}$.

The preceding result implies the following corollary since the condition $\gamma_k(\mathscr{R}) \leq \gamma_k(\mathscr{S})$, $k = 1, \ldots, j$, implies (3.7). For some particular cases in the model of generalized order statistics, we refer to Hu and Zhuang [457, Theorem 3.2] (see also Cramer and Kamps [301]).

Corollary 3.2.5. Let the assumptions of Corollary 3.2.4 be satisfied. If $j \leq i$ and $\gamma_k(\mathscr{R}) \leq \gamma_k(\mathscr{S})$, $k = 1, \ldots, j$, then $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{hr}} X^{\mathscr{R}}_{i:m_1:n_1}$.

Likelihood Ratio Order

The following result can be found in Korwar [544] and Hu and Zhuang [457] (for generalized order statistics). The particular case $i = j + 1$ and $\mathscr{R} = \mathscr{S}$ was considered in Cramer et al. [311].

Theorem 3.2.6. Let $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}, \mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$ with $m_1, m_2 \in \mathbb{N}$ be censoring schemes and $X^{\mathscr{S}}_{j:m_2:n_2}, X^{\mathscr{R}}_{i:m_1:n_1}$ be progressively Type-II censored order statistics from the same absolutely continuous cumulative distribution function F. If $j \leq i$ and $\gamma_k(\mathscr{R}) \leq \gamma_k(\mathscr{S})$, $k = 1, \ldots, j$, then $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{lr}} X^{\mathscr{R}}_{i:m_1:n_1}$.

Proof. The proof is based on the following result which can be found in Shaked and Shanthikumar [799, p. 46] (see also Keilson and Sumita [513]):

> Let $X_1, \ldots, X_j$ and $Y_1, \ldots, Y_i$ be two sets of independent random variables with $j \leq i$ and $X_r \leq_{\mathsf{lr}} Y_r$, $r = 1, \ldots, j$. If all random variables have a log-concave density function, then
>
> $$\sum_{r=1}^{j} X_r \leq_{\mathsf{lr}} \sum_{r=1}^{i} Y_r.$$

Let $Z^{\mathscr{S}}_r \sim \mathsf{Exp}(\gamma_r(\mathscr{S}))$, $r = 1, \ldots, j$, and $Z^{\mathscr{R}}_r \sim \mathsf{Exp}(\gamma_r(\mathscr{R}))$, $r = 1, \ldots, i$, be independent exponential random variables. Then, we obtain from $\gamma_r(\mathscr{R}) \leq \gamma_r(\mathscr{S})$ the result $Z^{\mathscr{S}}_r \leq_{\mathsf{lr}} Z^{\mathscr{R}}_r$, $r = 1, \ldots, j$. Since progressively Type-II censored order statistics based on an exponential distribution can be written as

$$Z^{\mathscr{S}}_{j:m_2:n_2} = \sum_{r=1}^{j} Z^{\mathscr{S}}_r \text{ and } Z^{\mathscr{R}}_{i:m_1:n_1} = \sum_{r=1}^{i} Z^{\mathscr{R}}_r,$$

the assumption $\gamma_r(\mathscr{R}) \le \gamma_r(\mathscr{S})$, $r = 1, \ldots, j$, leads to

$$Z^{\mathscr{S}}_{j:m_2:n_2} = \sum_{r=1}^{j} Z^{\mathscr{S}}_r \le_{\mathsf{lr}} \sum_{r=1}^{j} Z^{\mathscr{R}}_r + \sum_{r=j+1}^{i} Z^{\mathscr{R}}_r = Z^{\mathscr{R}}_{i:m_1:n_1}$$

(cf. Shaked and Shanthikumar [799, p. 46/47]). Since the function $H^{\leftarrow} = F^{\leftarrow}(1 - \exp(\cdot))$ is nondecreasing and $H^{\leftarrow}(Z^{\mathscr{S}}_{j:m_2:n_2}) \stackrel{\mathsf{d}}{=} X^{\mathscr{S}}_{j:m_2:n_2}$, the result follows from Theorem 1.C.8 in Shaked and Shanthikumar [799]. □

In Hu and Zhuang [457, Theorem 3.1], the preceding result can be found for generalized order statistics with parameters

(i) $i = j + 1$, $\gamma_r(\mathscr{R}) = \gamma_r(\mathscr{S})$, $r = 1, \ldots, j$;
(ii) $i = j$, $\gamma_r(\mathscr{R}) = \gamma_r(\mathscr{S}) + c$, $r = 1, \ldots, j$, with $c = m_n \ge 0$;
(iii) $i = j + 1$, $\gamma_{r+1}(\mathscr{R}) = \gamma_r(\mathscr{S}) - c_r$, with $c_r = R_{r+1} \ge 0$, $r = 1, \ldots, j$.

For order statistics, Theorem 3.2.6 reads as follows:

(i) Let $\gamma_r(\mathscr{S}) = m - r + 1$, $r = 1, \ldots, j$, and $\gamma_r(\mathscr{R}) = n - r + 1$, $r = 1, \ldots, i$, then $X_{j:m} \le_{\mathsf{lr}} X_{i:n}$ for $j \le i$ and $m \ge n$ (cf. Chan et al. [242] for $m = n$);
(ii) Let $\gamma_r(\mathscr{S}) = m - j + r$, $r = 1, \ldots, j$, and $\gamma_r(\mathscr{R}) = n - i + r$, $r = 1, \ldots, i$, then $X_{j:m} \le_{\mathsf{lr}} X_{i:n}$ for $j \le i$ and $i - j \ge n - m$ (cf. Raqab and Amin [743]; see also Khaledi and Kochar [519], Lillo et al. [598], and Shaked and Shanthikumar [799], Theorem 1.C.37);
(iii) Let $\gamma_r(\mathscr{S}) = n - r + 2$, $r = 1, \ldots, n+1$, and $\gamma_r(\mathscr{R}) = n - r + 1$, $r = 1, \ldots, n$. Then, $\gamma_1(\mathscr{S}) = n + 1 \ge n = \gamma_1(\mathscr{R})$ so that $X_{1:n+1} \le_{\mathsf{lr}} X_{1:n}$. Moreover, $\gamma_{r+1}(\mathscr{S}) = \gamma_r(\mathscr{R}) = n - r + 1$ such that $X_{n:n} \le_{\mathsf{lr}} X_{n+1:n+1}$ (cf. Boland et al. [212, Theorem 3.4]).

Remark 3.2.7. As is obvious from the proof of Theorem 3.2.6, the condition in this theorem can be weakened to

$$(\gamma_{i-j+1}(\mathscr{R}), \ldots, \gamma_i(\mathscr{R})) \le (\gamma_1(\mathscr{S}), \ldots, \gamma_j(\mathscr{S})).$$

From the monotonicity of the γ's, this is equivalent to the condition

$$(\gamma_{\ell_1}(\mathscr{R}), \ldots, \gamma_{\ell_j}(\mathscr{R})) \le (\gamma_1(\mathscr{S}), \ldots, \gamma_j(\mathscr{S})) \qquad \text{for some set } \{\ell_1, \ldots, \ell_j\} \subseteq \{1, \ldots, i\},$$

imposed by Izadi and Khaledi [473] to prove a generalization of Theorem 3.2.6 [see also condition (3.6)].

The likelihood ratio order of progressively Type-II censored order statistics can be discussed in the two-sample case as well. The following comparison result has been obtained by Hu and Zhuang [457] for generalized order statistics.

Theorem 3.2.8. Let F, G be absolutely continuous cumulative distribution functions with density functions f and g, respectively, and $X \sim F, Y \sim G$. Moreover, let $\mathscr{R}$ be a censoring scheme. If the hazard rates $\lambda_F = \frac{f}{\overline{F}}$ and $\lambda_G = \frac{g}{\overline{G}}$ have an increasing ratio $\frac{\lambda_G}{\lambda_F}$, then

$$F \leq_{\mathsf{hr}} G \quad\Longrightarrow\quad X^{\mathscr{R}}_{i:m:n} \leq_{\mathsf{lr}} Y^{\mathscr{R}}_{i:m:n}.$$

Proof. Denote by $h_{*,i}$, $f_{*,i}$, and $g_{*,i}$ the density functions of the ith progressively Type-II censored order statistic from a standard exponential distribution, F and G, respectively. Then, $h_{*,i}$ is log-concave and

$$\frac{g_{*,i}(x)}{f_{*,i}(x)} = \frac{h_{*,i}(-\log \overline{G}(x))}{h_{*,i}(-\log \overline{F}(x))} \cdot \frac{\lambda_G(x)}{\lambda_F(x)}.$$

Since $F \leq_{\mathsf{hr}} G$, the same argument as in the proof of Theorem 3.2.3 applied to the ratio $\frac{h_{*,i}(-\log \overline{G}(x))}{h_{*,i}(-\log \overline{F}(x))}$ proves that this ratio increases. Here, it has to be noticed that $h_{*,i}$ is a decreasing function. This is due to the fact that $H_{*,i}$ is a DFR cumulative distribution function and thus the hazard rate $\lambda_{H_{*,i}}$ is decreasing. Hence, $h_{*,i} = \lambda_{H_{*,i}} \overline{H}_{*,i}$ is decreasing. Since $\frac{\lambda_G}{\lambda_F}$ is increasing by assumption and all functions are nonnegative, this yields the desired result. □

Given $F \leq_{\mathsf{hr}} G$ and the assumption "*λ_G/λ_F is increasing,*" Lemma 3.5 in Belzunce et al. [186] yields that $F \leq_{\mathsf{lr}} G$. The following theorem is due to Belzunce et al. [188] who established it for generalized order statistics.

Theorem 3.2.9. Let $X^{\mathscr{R}}_{i:m:n}, Y^{\mathscr{R}}_{i:m:n}$, $1 \leq i \leq m$, be progressively Type-II censored order statistics from absolutely continuous cumulative distribution functions F and G, respectively, with $F \leq_{\mathsf{lr}} G$ and censoring scheme $\mathscr{R}$. Then, $X^{\mathscr{R}}_{i:m:n} \leq_{\mathsf{lr}} Y^{\mathscr{R}}_{i:m:n}$, $1 \leq i \leq m$.

Remark 3.2.10. It is not sufficient to assume $F \leq_{\mathsf{lr}} G$ in order to obtain likelihood ratio ordering of generalized order statistics. Belzunce et al. [188] established the result for generalized order statistics provided that one of the following assumptions hold:

(i) $m_r = \gamma_r - \gamma_{r+1} - 1 \geq 0$ for $r = 1, \ldots, i-1$ and $F \leq_{\mathsf{lr}} G$, or
(ii) $m_r = \gamma_r - \gamma_{r+1} - 1 \geq -1$ for $r = 1, \ldots, i-1$ and $F \leq_{\mathsf{hr}} G$ and $\frac{\lambda_G}{\lambda_F}$ is increasing.

Part (i) is an extension of Franco et al. [381, Theorem 3.4], while Part (ii) is strengthened by a corresponding version of Theorem 3.2.8 since the condition $m_k \geq -1$ can be dropped (cf. Hu and Zhuang [457]).

Dispersive Order

The following dispersive ordering result has been proved for generalized order statistics by Khaledi [518].

Theorem 3.2.11. Let $Z^{\mathscr{R}}_{i:m_1:n_1}$ and $Z^{\mathscr{S}}_{j:m_2:n_2}$ be progressively Type-II censored order statistics from a standard exponential distribution with censoring schemes $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}, \mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$, $m_1, m_2 \in \mathbb{N}$, respectively. Moreover, let F be a continuous DFR cumulative distribution function and $X^{\mathscr{R}}_{i:m_1:n_1}$ and $X^{\mathscr{S}}_{j:m_2:n_2}$ be progressively Type-II censored order statistics from F. If $j \le i$ and condition (3.7) holds, then

$$\begin{aligned} Z^{\mathscr{S}}_{j:m_2:n_2} &\le_{\mathsf{disp}} Z^{\mathscr{R}}_{i:m_1:n_1}, \\ X^{\mathscr{S}}_{j:m_2:n_2} &\le_{\mathsf{disp}} X^{\mathscr{R}}_{i:m_1:n_1}. \end{aligned} \tag{3.8}$$

Proof. Let $Z_1, \ldots, Z_j, Z_1^*, \ldots, Z_i^* \overset{\text{iid}}{\sim} \mathsf{Exp}(1)$. Then, according to (2.13),

$$Z^{\mathscr{S}}_{j:m_2:n_2} \overset{\mathsf{d}}{=} \sum_{\ell=1}^{j} \frac{1}{\gamma_\ell(\mathscr{S})} Z_\ell,$$

$$Z^{\mathscr{R}}_{i:m_1:n_1} \overset{\mathsf{d}}{=} \sum_{\ell=1}^{j} \frac{1}{\gamma_{i-j+\ell}(\mathscr{R})} Z_\ell^* + \sum_{\ell=j+1}^{i} \frac{1}{\gamma_{i-\ell+1}(\mathscr{R})} Z_\ell^*.$$

In the next step, we use the fact that a random variable X satisfies $X \le_{\mathsf{disp}} X+Y$ for any random variable Y independent of X iff X has a log-concave density function (cf. Saunders [780]): $\sum_{\ell=1}^{j} \frac{1}{\gamma_{i-j+\ell}(\mathscr{R})} Z_\ell^*$ is independent of $\sum_{\ell=j+1}^{i} \frac{1}{\gamma_{i-\ell+1}(\mathscr{R})} Z_\ell^*$ and has a log-concave density function since a convolution of independent exponentially distributed random variables has a log-concave density function. Hence,

$$\sum_{\ell=1}^{j} \frac{1}{\gamma_{i-j+\ell}(\mathscr{R})} Z_\ell^* \le_{\mathsf{disp}} Z^{\mathscr{R}}_{i:m_1:n_1}.$$

According to Khaledi and Kochar [521, Theorem 2.1], condition (3.7) implies

$$\sum_{\ell=1}^{j} \frac{1}{\gamma_\ell(\mathscr{S})} Z_\ell \le_{\mathsf{disp}} \sum_{\ell=1}^{j} \frac{1}{\gamma_{i-j+\ell}(\mathscr{R})} Z_\ell^*$$

so that the transitivity of the dispersive order proves $Z^{\mathscr{S}}_{j:m_2:n_2} \le_{\mathsf{disp}} Z^{\mathscr{R}}_{i:m_1:n_1}$.

Moreover, since $Z^{\mathscr{S}}_{j:m_2:n_2} \le_{\mathsf{st}} Z^{\mathscr{R}}_{i:m_1:n_1}$ and

$$S_F = F^{\leftarrow}\big(1 - \exp(\cdot)\big) \tag{3.9}$$

is an increasing convex transformation, Theorem 2.2 (ii) of Rojo and He [758] yields the desired result in (3.8) (see also Hu and Zhuang [457, Lemmas 2.1 and 2.3]). □

In terms of order statistics, the preceding result shows that $X_{j:n} \le_{\mathsf{disp}} X_{i:n}$ if $j \le i$ and F is DFR. This result was obtained by Kochar [538].

Since dispersive order $X \le_{\mathsf{disp}} Y$ of random variables X and Y implies an ordering of the variances, i.e., $\mathsf{Var}\, X \le \mathsf{Var}\, Y$, the preceding result proves that

$$\mathsf{Var}\, X^{\mathscr{S}}_{j:m_2:n_2} \le \mathsf{Var}\, X^{\mathscr{R}}_{i:m_1:n_1}$$

given $j \le i$ and condition (3.7). This extends a result of David and Groeneveld [323] for order statistics.

The following stochastic comparison can be found in Belzunce et al. [188, Theorem 3.12] and Khaledi [518, Theorem 3.9].

Proposition 3.2.12. Let F, G be continuous cumulative distribution functions with $F \le_{\mathsf{disp}} G$ and $X \sim F, Y \sim G$. Moreover, let $\mathscr{R}$ be a censoring scheme. Then, $X^{\mathscr{R}}_{i:m:n} \le_{\mathsf{disp}} Y^{\mathscr{R}}_{i:m:n}$, $1 \le i \le m$.

Proof. As in the proof of Theorem 3.2.3, we have the representations $F_{i,*} = F^{X^{\mathscr{R}}_{i:m:n}} = F^{Z^{\mathscr{R}}_{i:m:n}}(-\log \overline{F})$ and $G_{i,*} = F^{Y^{\mathscr{R}}_{i:m:n}} = F^{Z^{\mathscr{R}}_{i:m:n}}(-\log \overline{G})$. Then,

$$F^{\leftarrow}_{*,i} \circ G_{*,i} = F \leftarrow \circ \Big(F^{Z^{\mathscr{R}}_{i:m:n}}\Big)^{\leftarrow} \circ F^{Z^{\mathscr{R}}_{i:m:n}} \circ G = F^{\leftarrow} \circ G.$$

An application of (A.1) yields the claimed result. □

As pointed out by Belzunce et al. [188], the method used in the proof of Proposition 3.2.12 can be applied to several other stochastic orders, e.g., convex $\le_{\mathsf{c}}$, star-shaped $\le_{*}$, and superadditive order $\le_{\mathsf{su}}$, respectively (see Cramer [286]; for order statistics, see Barlow and Proschan [168, p. 107/108] and Dharmadhikari and Joag-dev [339]). This yields the following results.

Theorem 3.2.13. Let F, G be continuous cumulative distribution functions and $X \sim F, Y \sim G$. Moreover, let $\mathscr{R}$ be a censoring scheme. If $F \le_{\mathsf{c}} [\le_{*}, \le_{\mathsf{su}}]\, G$, then $X^{\mathscr{R}}_{r:m:n} \le_{\mathsf{c}} [\le_{*}, \le_{\mathsf{su}}]\, Y^{\mathscr{R}}_{r:m:n}$, $1 \le r \le m$.

A combination of Theorem 3.2.11 and Proposition 3.2.12 leads directly to the following result.

Theorem 3.2.14. Let F, G be continuous cumulative distribution functions with $F \le_{\mathsf{disp}} G$ where either F or G is DFR. Moreover, let $X^{\mathscr{S}}_{j:m_2:n_2}$ and $Y^{\mathscr{R}}_{i:m_1:n_1}$ be progressively Type-II censored order statistics from F and G with censoring schemes $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}$, $\mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$, $m_1, m_2 \in \mathbb{N}$, respectively. Then, if $j \le i$ and condition (3.7) holds, then

$$X^{\mathscr{S}}_{j:m_2:n_2} \le_{\mathsf{disp}} Y^{\mathscr{R}}_{i:m_1:n_1}.$$

Minimal Bounds w.r.t. Stochastic, Hazard, and Likelihood Ratio Order

Applying the results of Bon and Păltănea [213], we can identify minimal generalized order statistics w.r.t. stochastic, hazard, and likelihood ratio order, respectively. It turns out that these generalized order statistics are kth record values with appropriately chosen value for k. In particular, we have the following stochastic bounds.

Theorem 3.2.15. Let $X_{m:m:n}$ be a progressively Type-II censored order statistic from a (absolutely) continuous cumulative distribution function F with censoring scheme $\mathscr{R}$. Moreover, let $X_m^{(k)}$ be an mth k-record value drawn from F.

Then:

(i) $X_m^{(k)} \leq_{\mathsf{st}} X_{m:m:n}$ with $k = \sqrt[m]{\prod_{j=1}^m \gamma_j}$;

(ii) $X_m^{(k)} \leq_{\mathsf{hr}} X_{m:m:n}$ with $k = \sqrt[m]{\prod_{j=1}^m \gamma_j}$;

(iii) $X_m^{(k)} \leq_{\mathsf{lr}} X_{m:m:n}$ with $k = \frac{1}{m} \sum_{j=1}^m \gamma_j$.

Lorenz Order and Convex Orders

In the area of order statistics and record values, some results on Lorenz ordering are known (cf. Arnold and Nagaraja [52], Arnold and Villaseñor [55,56], Kochar [539]). We present some extensions to progressively Type-II censored order statistics.

It is known that star ordering implies Lorenz ordering (cf. Arnold and Villaseñor [56, p. 80]). A direct application of Theorem 3.2.13 shows that

$$F \leq_* G \Longrightarrow X_{r:m:n}^{\mathscr{R}} \leq_{\mathsf{L}} Y_{r:m:n}^{\mathscr{R}}, \quad 1 \leq r \leq m.$$

Hence, progressively Type-II censored order statistics are Lorenz ordered if their parent distributions are star ordered. Another general result is the following.

Theorem 3.2.16. Let $X_{1:m:n}^{\mathscr{R}}, \dots, X_{m:m:n}^{\mathscr{R}}$ be progressively Type-II censored order statistics from F with censoring scheme $\mathscr{R}$ such that $\frac{g(z,x)}{x}$ is increasing for every $z \in (0,1)$, where $g(z,x) = F^{\leftarrow}(1 - z\overline{F}(x))$. Moreover, suppose that the moments of the progressively Type-II censored order statistics exist. Then, $X_{r:m:n}^{\mathscr{R}} \leq_{\mathsf{L}} X_{r+1:m:n}^{\mathscr{R}}$, $1 \leq r \leq m-1$.

Proof. First, we have from Theorem 2.3.6 the connection

$$X_{r+1:m:n}^{\mathscr{R}} \stackrel{\mathsf{d}}{=} F^{\leftarrow}\big(1 - U_{r+1}^{1/\gamma_{r+1}} \overline{F}(X_{r:m:n}^{\mathscr{R}})\big) = g(U_{r+1}^{1/\gamma_{r+1}}, X_{r:m:n}^{\mathscr{R}}),$$

where $U_{r+1}^{1/\gamma_{r+1}}$ and $X_{r:m:n}^{\mathscr{R}}$ are independent random variables. By assumption, $\frac{g(z,x)}{x}$ is increasing in x so that Theorem 2.7 in Arnold and Villaseñor [56] yields $X_{r:m:n}^{\mathscr{R}} \leq_{\mathsf{L}} X_{r+1:m:n}^{\mathscr{R}}$ for $1 \leq r \leq m-1$. □

Example 3.2.17. The preceding theorem holds, e.g., for a Pareto(α)-distribution with parameter $\alpha > 1$. We have $g(z,x) = z^{-1/\alpha}x$ so that the ratio $\frac{g(z,x)}{x}$ is constant w.r.t. x. Hence, progressively Type-II censored order statistics based on a Pareto distribution are Lorenz ordered if $\alpha > 1$. Notice that the condition $\alpha > 1$ guarantees the existence of the moments. For order statistics, this yields $X_{r:n} \leq_{\mathsf{L}} X_{r+1:n}$, $1 \leq r \leq n-1$, which is given in Arnold and Villaseñor [56, Theorem 4.1]. Alternatively, the Lorenz ordering of Pareto progressively Type-II censored order statistics may be obtained from Strassen's theorem (see Arnold and Villaseñor [55, 56]) using the product representation in Corollary 2.3.13.

Since stochastic order $\leq_{\mathsf{st}}$ implies increasing convex order $\leq_{\mathsf{icx}}$, we obtain from Theorem 3.2.1 the following result.

Corollary 3.2.18. Let F, G be continuous cumulative distribution functions with $F \leq_{\mathsf{st}} G$. Moreover, let $X^{\mathscr{S}}_{j:m_2:n_2}$ and $Y^{\mathscr{R}}_{i:m_1:n_1}$ be progressively Type-II censored order statistics from cumulative distribution functions G and F with censoring schemes $\mathscr{R} \in \mathscr{C}^{m_1}_{m_1,n_1}$, $\mathscr{S} \in \mathscr{C}^{m_2}_{m_2,n_2}$, $m_1, m_2 \in \mathbb{N}$, respectively. If $j \leq i$ and if (3.7) holds, then $X^{\mathscr{S}}_{j:m_2:n_2} \leq_{\mathsf{icx}} Y^{\mathscr{R}}_{i:m_1:n_1}$.

For the next result, we need the following auxiliary results which can be found in Shaked and Shanthikumar [799, Theorem 3.A.35].

Theorem 3.2.19. Let $k \leq m$ and $(\gamma_1^{-1}, \ldots, \gamma_m^{-1}) \preccurlyeq_{\mathsf{m}} (\eta_1^{-1}, \ldots, \eta_k^{-1}, 0^{*m-k})$ and $X_1, \ldots, X_m$ be independent and identically distributed random variables with finite expectation. Then,

$$\sum_{j=1}^{k} \frac{1}{\eta_j} X_j \leq_{\mathsf{cx}} \sum_{j=1}^{m} \frac{1}{\gamma_j} X_j \quad \text{and} \quad \sum_{j=1}^{k} \frac{1}{\eta_j} X_j \leq_{\mathsf{L}} \sum_{j=1}^{m} \frac{1}{\gamma_j} X_j.$$

From this result and the sum representation in (2.13) of exponential progressively Type-II censored order statistics, we obtain the following ordering result.

Corollary 3.2.20. Let $k \leq m$ and $X^{\mathscr{S}}_{m:m:n}$ and $Y^{\mathscr{R}}_{k:k:s}$ be progressively Type-II censored order statistics based on a standard exponential distribution with censoring schemes $\mathscr{S}$ and $\mathscr{R}$ such that

$$(\gamma_1(\mathscr{R})^{-1}, \ldots, \gamma_m(\mathscr{R})^{-1}) \preccurlyeq_{\mathsf{m}} (\gamma_1(\mathscr{S})^{-1}, \ldots, \gamma_k(\mathscr{S})^{-1}, 0^{*m-k}).$$

Then,

$$X^{\mathscr{S}}_{m:m:n} \leq_{\mathsf{cx}} Y^{\mathscr{R}}_{k:k:s} \quad \text{and} \quad X^{\mathscr{S}}_{m:m:n} \leq_{\mathsf{L}} Y^{\mathscr{R}}_{k:k:s}.$$

Results for comparisons in the increasing convex directional order have been established by Balakrishnan et al. [153]. Dependence orderings are addressed in Khaledi and Kochar [522]. In particular, they showed that, under certain conditions on the parameters, some measures of concordance are monotone. These measures include Spearman's ρ, Kendall's τ, and Gini's coefficient of association (see Khaledi and Kochar [522, Corollary 3.1]).

3.2.2 Multivariate Stochastic Orderings and Its Applications to Progressively Type-II Censored Order Statistics

In this section, we consider applications of multivariate stochastic orders to progressively Type-II censored order statistics, i.e., stochastic and likelihood ratio order. Applications of multivariate versions of the hazard order and dispersive order can be found in Belzunce et al. [188].

Stochastic Order

Theorem 3.2.21 (Belzunce et al. [188]). Let $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{Y}^{\mathscr{R}}$ be vectors of progressively Type-II censored order statistics from continuous cumulative distribution functions F and G with censoring scheme $\mathscr{R}$, respectively. Then, if $F \le_{\mathsf{st}} G$, then $\mathbf{X}^{\mathscr{R}} \le_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}$.

Proof. The proof is based on Theorem 6.B.1 in Shaked and Shanthikumar [799]:

> Two random vectors $\mathbf{X}$, $\mathbf{Y}$ satisfy $\mathbf{X} \le_{\mathsf{st}} \mathbf{Y}$ if random vectors $\tilde{\mathbf{X}}$, $\tilde{\mathbf{Y}}$, defined on the same probability space, exist such that $\mathbf{X} \stackrel{\mathsf{d}}{=} \tilde{\mathbf{X}}$, $\mathbf{Y} \stackrel{\mathsf{d}}{=} \tilde{\mathbf{Y}}$ and $P(\tilde{\mathbf{X}} \le \tilde{\mathbf{Y}}) = 1$.

Let $\tilde{\mathbf{X}} = \big(F^{\leftarrow}(U^{\mathscr{R}}_{r:m:n})\big)_{1\le r\le m}$, $\tilde{\mathbf{Y}} = \big(G^{\leftarrow}(U^{\mathscr{R}}_{r:m:n})\big)_{1\le r\le m}$, where $U^{\mathscr{R}}_{1:m:n}, \ldots, U^{\mathscr{R}}_{m:m:n}$ denote uniform progressively Type-II censored order statistics with censoring scheme $\mathscr{R}$. Then, $\tilde{\mathbf{X}} \stackrel{\mathsf{d}}{=} \mathbf{X}^{\mathscr{R}}$ and $\tilde{\mathbf{Y}} \stackrel{\mathsf{d}}{=} \mathbf{Y}^{\mathscr{R}}$.

The assumption $F \le_{\mathsf{st}} G$ leads directly to the identity

$$P(\tilde{\mathbf{X}} \le \tilde{\mathbf{Y}}) = P\bigg(\bigcap_{r=1}^{m} \big\{F^{\leftarrow}(U^{\mathscr{R}}_{r:m:n}) \le G^{\leftarrow}(U^{\mathscr{R}}_{r:m:n})\big\}\bigg) = 1$$

which proves the result. □

Since the multivariate stochastic order is preserved under marginalization, this yields directly Part (i) of Theorem 3.2.1.

Likelihood Ratio Order

The following result is taken from a more general result established by Belzunce et al. [188] for generalized order statistics.

Theorem 3.2.22. Let $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{Y}^{\mathscr{R}}$ be vectors of progressively Type-II censored order statistics from absolutely continuous cumulative distribution functions F and G with censoring scheme $\mathscr{R}$, respectively. Then, each of the conditions

(i) $F \le_{\mathsf{lr}} G$,
(ii) $F \le_{\mathsf{hr}} G$ and $\frac{\lambda_G}{\lambda_F}$ is increasing

implies $\mathbf{X}^{\mathscr{R}} \le_{\mathsf{lr}} \mathbf{Y}^{\mathscr{R}}$.

Proof. The joint density of uniform progressively Type-II censored order statistics is given by [cf. (2.3)]

$$f^{\mathbf{U}^{\mathscr{R}}}(\mathbf{u}_m) = \Big(\prod_{j=1}^m \gamma_j\Big)\prod_{j=1}^m (1-u_j)^{R_j},$$

where $0 < u_1 < \cdots < u_m < 1$. Moreover,

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}_m) = f^{\mathbf{U}^{\mathscr{R}}}(F(x_1),\ldots,F(x_m)) \cdot \prod_{j=1}^m f(x_j),$$

$$f^{\mathbf{Y}^{\mathscr{R}}}(\mathbf{y}_m) = f^{\mathbf{U}^{\mathscr{R}}}(G(y_1),\ldots,G(y_m)) \cdot \prod_{j=1}^m g(y_j).$$

According to the definition of multivariate likelihood ratio order, (A.2) has to be checked. As can be seen from the product structure of the joint density function, it is sufficient to consider the inequalities

$$\overline{F}^{R_j}(x_j)f(x_j)\overline{G}^{R_j}(y_j)g(y_j) \le \overline{F}^{R_j}(x_j\wedge y_j)f(x_j\wedge y_j)\overline{G}^{R_j}(x_j\vee y_j)g(x_j\vee y_j), \tag{3.10}$$

$j = 0,\ldots,m$, separately. For $x_j \le y_j$, both sides of (3.10) are identical and so nothing remains to be shown. Suppose $x_j \ge y_j$. Then, we have to verify the inequality

$$\overline{F}^{R_j}(x_j)f(x_j)\overline{G}^{R_j}(y_j)g(y_j) \le \overline{F}^{R_j}(y_j)f(y_j)\overline{G}^{R_j}(x_j)g(x_j) \tag{3.11}$$

$$\iff \left(\frac{\overline{F}(x_j)}{\overline{G}(x_j)}\right)^{R_j} \cdot \frac{f(x_j)}{g(x_j)} \le \left(\frac{\overline{F}(y_j)}{\overline{G}(y_j)}\right)^{R_j} \cdot \frac{f(y_j)}{g(y_j)}.$$

Hence, it remains to be shown that $\left(\frac{\overline{F}}{\overline{G}}\right)^{R_j} \cdot \frac{f}{g}$ is a nonincreasing function. By assumption, R_j is nonnegative. Moreover, the assumption $F \le_{\mathsf{lr}} G$ implies that $\frac{f}{g}$ is nonincreasing and that $F \le_{\mathsf{hr}} G$. Therefore, $\frac{\overline{F}}{\overline{G}}$ is nonincreasing. Since all functions are nonnegative, this yields the result provided that condition (i) is satisfied.

Notice that $\left(\frac{\overline{F}}{\overline{G}}\right)^{R_j} \cdot \frac{f}{g} = \left(\frac{\overline{F}}{\overline{G}}\right)^{R_j+1} \cdot \frac{\lambda_F}{\lambda_G}$ so that (3.11) and, thus, the result follows from condition (ii), too. □

Remark 3.2.23. For generalized order statistics, the following assumptions have been imposed on the model parameters by Belzunce et al. [188] to prove the above theorem:

(i) $m_r = \gamma_r - \gamma_{r+1} - 1 \geq 0$ for $r = 1, \ldots, m-1$, $\gamma_m \geq 1$, and $F \leq_{\text{lr}} G$, or
(ii) $m_r = \gamma_r - \gamma_{r+1} - 1 \geq -1$ for $r = 1, \ldots, m-1$ and $F \leq_{\text{hr}} G$ and $\frac{\lambda_G}{\lambda_F}$ is increasing.

Balakrishnan et al. [148] applied these results in combination with Theorem 3.11.4 in Müller and Stoyan [659] to establish the following conditional ordering result.

Theorem 3.2.24. Let $L \subseteq \mathbb{R}^n$ be a sublattice, i.e., $\mathbf{x}, \mathbf{y} \in L$ implies $\mathbf{x} \vee \mathbf{y} \in L$ and $\mathbf{x} \wedge \mathbf{y} \in L$. Then, under the conditions of Theorem 3.2.22, for $1 \leq s \leq m$,

$$[X_{s:m:n} | \mathbf{X}^{\mathscr{R}} \in L] \leq_{\text{lr}} [Y_{s:m:n} | \mathbf{Y}^{\mathscr{R}} \in L].$$

The result implies likelihood ratio ordering of residual lifetimes. For $r \leq s$, and given the conditions of Theorem 3.2.22,

$$[X_{s:m:n} - t | X_{r:m:n} > t] \leq_{\text{lr}} [Y_{s:m:n} - t | Y_{r:m:n} > t].$$

Here, the sublattice L is defined by $\{\mathbf{x}_m | x_{r:m} > t\} \subseteq \mathbb{R}^m_{\leq}$. In this context, Hashemi et al. [434] studied also the residual life. They obtained stochastic orderings for the quantity $[X_{s:m:n} - t | X_{r:m:n} \leq t < X_{r+1:m:n}]$, $0 \leq r < s \leq m$.

Since the multivariate likelihood ratio order is preserved under marginalization, Theorem 3.2.22 leads directly to results for marginal distributions. In fact, this provides an alternative proof of Theorem 3.2.9.

The same approach can be utilized to prove a result for different sets of parameters but the same underlying distribution. First, uniform generalized order statistics are considered.

Proposition 3.2.1 (Belzunce et al. [188]). *Let* $\mathbf{U}^{\mathscr{R}}$ *and* $\mathbf{U}^{\mathscr{S}}$ *be vectors of uniform progressively Type-II censored order statistics with censoring schemes* $\mathscr{R}$ *and* $\mathscr{S}$*, respectively. Then,*

$$\mathbf{U}^{\mathscr{R}} \leq_{lr} \mathbf{U}^{\mathscr{S}} \iff R_j \geq S_j, \ j = 1, \ldots, m.$$

Proof. As in the proof of Theorem 3.2.22, we obtain for all $1 > x_j \geq y_j > 0$, $j = 1, \ldots, m$,

$$(1 - x_j)^{R_j} (1 - y_j)^{S_j} \leq (1 - y_j)^{R_j} (1 - x_j)^{S_j}, \quad j = 1, \ldots, m.$$

This leads to the inequalities

$$\left(\frac{1-x_j}{1-y_j}\right)^{R_j-S_j} \le 1 \quad j=1,\ldots,m.$$

Since $\frac{1-x_j}{1-y_j} \le 1$, this holds iff $R_j \ge S_j$, $j=1,\ldots,m$. □

Since the multivariate likelihood ratio order is preserved under strictly increasing transformations of each component (cf. Shaked and Shanthikumar [799, Theorem 6.E.3]), this result leads directly to a theorem for underlying continuous cumulative distribution functions.

Theorem 3.2.25 (Belzunce et al. [188]). Let $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{X}^{\mathscr{S}}$ be vectors of progressively Type-II censored order statistics from an absolutely continuous and strictly increasing cumulative distribution function F with censoring schemes $\mathscr{R}$ and $\mathscr{S}$, respectively. Then, $R_j \ge S_j$, $j=1,\ldots,m$, implies $\mathbf{X}^{\mathscr{R}} \le_{\mathsf{lr}} \mathbf{X}^{\mathscr{S}}$.

Theorems 3.2.25 and 3.2.22 can be combined to compare vectors of progressively Type-II censored order statistics (generalized order statistics) with different cumulative distribution functions and different censoring schemes (cf. Belzunce et al. [188, Theorem 3.10]). Further multivariate orders are also discussed in the literature. Multivariate hazard rate ordering is considered in Belzunce et al. [188]. Results in terms of the multivariate dispersive order are discussed in Belzunce et al. [189], Chen and Hu [250], and Xie and Hu [925]. Multivariate dispersive ordering for spacings of progressively Type-II censored order statistics is discussed in Zhuang and Hu [946].

3.2.3 Applications to Spacings of Progressively Type-II Censored Order Statistics

In this section, we consider stochastic orders of (p-)spacings. For progressively Type-II censored order statistics $X_{1:m:n},\ldots,X_{m:m:n}$, the associated spacings $\mathbf{S}^{\star\mathscr{R}} = (S_1^{\star\mathscr{R}},\ldots,S_m^{\star\mathscr{R}})$ are as defined in (2.44). A p-spacing is given by a contrast $S_{r+p-1,r-1}^{\star\mathscr{R}} = X_{r+p-1:m:n} - X_{r-1:m:n}$, where $p \in \{1,\ldots,m-r+1\}$ [see (2.38)]. Further, we write $\mathbf{S}_X^{\star\mathscr{R}}$ ($\mathbf{S}_Y^{\star\mathscr{R}}$) if the baseline distribution is specified by the cumulative distribution function of the random variable X (Y).

The first result on stochastic ordering of spacings of progressively Type-II censored order statistics is motivated by Kamps [498, Theorem V.2.7, p. 183]. Although this result was formulated in terms of m-generalized order statistics, the proof holds for arbitrary generalized order statistics. Moreover, Kamps [498] assumed that F, G are tail ordered. Since tail order and dispersive order are equivalent under the conditions considered here, we can note the theorem in terms of dispersive ordering.

Theorem 3.2.26. Let $X_{1:m:n}, \dots, X_{m:m:n}$ and $Y_{1:m:n}, \dots, Y_{m:m:n}$ be progressively Type-II censored order statistics based on F and G, respectively, and with censoring scheme $\mathscr{R}$. Then: If $F \le_{\mathsf{disp}} G$, then

$$S^{\star\mathscr{R}}_{r+p-1,r-1;X} \le_{\mathsf{st}} S^{\star\mathscr{R}}_{r+p-1,r-1;Y}, \quad 1 \le p \le m-r+1.$$

Proof. Since $F \le_{\mathsf{disp}} G$, it follows from $Y_{r:m:n} \le Y_{r+p-1:m:n}$ that

$$\begin{aligned} & F^{\leftarrow}(G(Y_{r+p-1:m:n})) - Y_{r+p-1:m:n} \le F^{\leftarrow}(G(Y_{r:m:n})) - Y_{r:m:n} \\ \Longleftrightarrow\ & F^{\leftarrow}(G(Y_{r+p-1:m:n})) - F^{\leftarrow}(G(Y_{r:m:n})) \le Y_{r+p-1:m:n} - Y_{r:m:n}. \end{aligned}$$

Since G is continuous, we have $X_{i:m:n} \stackrel{\mathsf{d}}{=} F^{\leftarrow}(G(Y_{i:m:n}))$, $i = 1, \dots, m$. Combining these, we obtain for arbitrary $t \in \mathbb{R}$

$$\begin{aligned} P(S^{\star\mathscr{R}}_{r+p-1,r-1;X} > t) &= P(F^{\leftarrow}(G(Y_{r+p-1:m:n})) - F^{\leftarrow}(G(Y_{r:m:n})) > t) \\ &\le P(Y_{r+p-1:m:n} - Y_{r:m:n} > t) = P(S^{\star\mathscr{R}}_{r+p-1,r-1;Y} > t). \end{aligned}$$

Thus, $S^{\star\mathscr{R}}_{r+p-1,r-1;X} \le_{\mathsf{st}} S^{\star\mathscr{R}}_{r+p-1,r-1;Y}$. □

As a particular case, the result of Bartoszewicz [173] for the usual order statistics is included.

Results on multivariate stochastic ordering of spacings are provided by Belzunce et al. [186] in terms of sequential order statistics. They considered interepoch times of nonhomogeneous pure birth processes which can be seen as spacings of sequential order statistics (see also Pellerey et al. [715]). The particular results in terms of generalized order statistics are presented in Belzunce et al. [188] which, due to the simpler structure, lead to more general results in the model of generalized order statistics. Since the proofs of these results are quite technical, we only state the results and refer for the details to the original literature. For $p = 1$, the result of Kamps [498] is a particular case of Theorem 3.2.27. The result is presented in terms of progressively Type-II censored order statistics.

Theorem 3.2.27. Let $X_{1:m:n}, \dots, X_{m:m:n}$ and $Y_{1:m:n}, \dots, Y_{m:m:n}$ be progressively Type-II censored order statistics based on F and G, respectively, and with censoring scheme $\mathscr{R}$. Then: If $F \le_{\mathsf{disp}} G$, then $\mathbf{S}^{\star\mathscr{R}}_X \le_{\mathsf{st}} \mathbf{S}^{\star\mathscr{R}}_Y$.

Finally, we note a result for the multivariate likelihood ratio order of spacings of progressively Type-II censored order statistics taken from Belzunce et al. [188] who established the result for generalized order statistics. Notice that, for generalized order statistics, additional assumptions may be necessary.

Theorem 3.2.28. Let $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{Y}^{\mathscr{R}}$ be vectors of progressively Type-II censored order statistics from absolutely continuous cumulative distribution function F and G with censoring schemes $\mathscr{R}$, respectively. The density function are denoted by f and g and the hazard rates by λ_F and λ_G. If f or g or both are log-convex and either

(i) $F \leq_{\text{lr}} G$, or
(ii) $F \leq_{\text{hr}} G$ and $\frac{\lambda_G}{\lambda_F}$ is increasing,

then $\mathbf{S}_X^{\star\mathscr{R}} \leq_{\text{lr}} \mathbf{S}_Y^{\star\mathscr{R}}$.

Contributions to ordering of spacings of order statistics have been made by Hu and Wei [455], Hu and Zhuang [456], and Misra and van der Meulen [651]. Xie and Zhuang [926] discussed the mean residual life order and the excess wealth order for spacings of generalized order statistics.

The following results for normalized spacings [cf. (2.9)]

$$S_1^{\mathscr{R}} = \gamma_1(\mathscr{R})X_{1:m:n}^{\mathscr{R}}, \qquad S_r^{\mathscr{R}} = \gamma_r(\mathscr{R})(X_{r:m:n}^{\mathscr{R}} - X_{r-1:m:n}^{\mathscr{R}}),\ r = 2,\dots,m,$$

from the same population cumulative distribution function but with different censoring schemes are due to Burkschat [228, see Theorem 4.5].

Theorem 3.2.29. Let F be a cumulative distribution function and $\mathscr{R}, \mathscr{S} \in \mathscr{C}_{m,n}^m$ such that

$$(\gamma_1(\mathscr{R}),\dots,\gamma_m(\mathscr{R})) \geq (\gamma_1(\mathscr{S}),\dots,\gamma_m(\mathscr{S})).$$

Then, $\mathbf{S}^{\mathscr{R}} \geq_{\text{st}} \mathbf{S}^{\mathscr{S}}$ if F is IFR, and $\mathbf{S}^{\mathscr{R}} \leq_{\text{st}} \mathbf{S}^{\mathscr{S}}$ if F is DFR.

For the non-normalized spacings the following result holds.

Corollary 3.2.30. Let F be a DFR-cumulative distribution function and $\mathscr{R}, \mathscr{S} \in \mathscr{C}_{m,n}^m$ such that

$$(\gamma_1(\mathscr{R}),\dots,\gamma_m(\mathscr{R})) \geq (\gamma_1(\mathscr{S}),\dots,\gamma_m(\mathscr{S})).$$

Then, $\mathbf{S}^{\star\mathscr{R}} \leq_{\text{st}} \mathbf{S}^{\star\mathscr{S}}$.

3.3 Aging Properties

Aging properties are closely connected to stochastic ordering. Theorem 3.2.13 has been used by Cramer [286] to establish aging properties of progressively Type-II censored order statistics. Subsequently, we discuss the following aging notions.

Definition 3.3.1. Let F be a continuous cumulative distribution function with support $[0,\infty)$. Then, F (or a random variable $X \sim F$) is said to have

(i) an IFR iff the hazard function $-\log \overline{F}$ is convex on $[0,\infty)$;

(ii) an increasing failure rate on the average (IFRA) iff the hazard function $-\log \overline{F}$ is star shaped on $[0, \infty)$;
(iii) the NBU property iff the hazard function $-\log \overline{F}$ is superadditive on $[0, \infty)$, i.e., $\overline{F}(x+y) \leq \overline{F}(x)\overline{F}(y)$ for $x, y \geq 0$.

Denoting by G the cumulative distribution function of a standard exponential distribution, then $F \leq_{\mathsf{c}} [\leq_{*}, \leq_{\mathsf{su}}]\, G$ is equivalent to the IFR [IFRA, NBU] property of F (see Shaked and Shanthikumar [799, Theorem 4.B.11]). Then, Theorem 3.2.13 yields the following result. For order statistics, parts of Theorem 3.3.2 can be found in Barlow and Proschan [168] and Takahasi [835] (see also Shaked and Shanthikumar [799, Theorem 4.B.15]).

Theorem 3.3.2. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F with support $[0, \infty)$. Then, F is IFR [IFRA, NBU] implies that $F^{X_{r:m:n}}$ is IFR [IFRA, NBU], $1 \leq r \leq m$.

Tavangar and Asadi [839] have established characterizations of aging properties. In particular, they proved the following result.

Theorem 3.3.3. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from a continuous cumulative distribution function F and censoring scheme $\mathscr{R}$ and $1 \leq r \leq s \leq m$. Then:

(i) If F is IFR, then $[X_{s:m:n} - t | X_{r:m:n} \geq t]$ is stochastically decreasing in $t \geq 0$;
(ii) If F is NBU, then $[X_{s:m:n} - t | X_{r:m:n} \geq t] \leq_{\mathsf{st}} X_{s:m:n}$, $t \geq 0$;
(iii) If F is NWU, then $X^{\mathscr{R}_{\lhd r-1}}_{s-r+1:m-r+1:\gamma_r} \leq_{\mathsf{st}} [X_{s:m:n} - t | X_{r:m:n} \geq t]$, $t \geq 0$.

Belzunce et al. [187] studied multivariate aging properties in terms of non-homogeneous birth processes. They applied their results to generalized order statistics. A restriction to progressive censoring shows that progressively Type-II censored order statistics $\mathbf{X}^{\mathscr{R}}$ is MIFR if F is an IFR-cumulative distribution function. Moreover, $\mathbf{X}^{\mathscr{R}}$ is multivariate Polya frequency of order 2 (MPF_2) if the density function of F is log-concave. Further notions of multivariate IFR and its applications to generalized order statistics are discussed in Arias-Nicolás et al. [47]. Some additional results in terms of sequential order statistics are established in Burkschat and Navarro [233] and Torrado et al. [853].

3.4 Asymptotic and Extreme Value Results

3.4.1 Extreme Value Analysis for Order Statistics

Extreme value theory of order statistics deals with limiting distributions of the appropriately normalized maximum $X_{n:n}$ of an IID sample $X_1, \ldots, X_n$, $X_1 \sim F$. Supposing that a limiting distribution exists, the problem is to find those nondegenerate distributions L which are limiting laws of the linearly normalized maximum $(X_{n:n} - b_n)/a_n$, i.e.,

$$F^n(a_n x + b_n) \xrightarrow{n\to\infty} L(x), \quad x \in \mathbb{R},$$

where $a_n > 0$ and $b_n \in \mathbb{R}$ are normalizing constants. The results for order statistics are due to Fisher and Tippett [375] and Gnedenko [404] who showed that the so-called extreme value distributions

$$H_{1,\rho}(x) = \begin{cases} 0, & x \le 0 \\ \exp(-x^{-\rho}), & x > 0 \end{cases}, \quad \rho > 0, \qquad \text{(Frechét)}$$

$$H_{2,\rho}(x) = \begin{cases} \exp(-|x|^{\rho}), & x < 0 \\ 1, & x \ge 0 \end{cases}, \quad \rho > 0, \qquad \text{(Weibull)}$$

$$H_{3,0}(x) = \exp(-\exp(-x)), \quad x \in \mathbb{R}. \qquad \text{(Gumbel)}$$

are the only possible limiting distributions. Subsequently, many results on extreme value theory for order statistics have been established. Improvements have been made in several aspects. In particular, it has been shown that the notion of regularly varying functions is important in characterizing distributions converging to a particular type of distribution (see, e.g., Resnick [751]). Moreover, convergence results have been established w.r.t. other norms, e.g., the variational distance (see Reiss [750]). A comprehensive presentation of this topic and many further results can be found in the monographs by Galambos [389, 390], Leadbetter et al. [576], Pfeifer [718], Resnick [751], Reiss [750], and De Haan and Ferreira [331].

3.4.2 Extreme Value Analysis for Progressively Type-II Censored Order Statistics

Extreme value analysis for progressively Type-II censored order statistics has been developed by Cramer [290] by utilizing results for generalized order statistics obtained in Cramer [285]. According to Theorem 2.3.6, the distribution of the mth progressively Type-II censored order statistic $X_{m:m:n}$ can be written as a quantile transformation of a product of IID transformed uniform random variables $U_1, \ldots, U_m$

$$X_{m:m:n} \stackrel{\mathsf{d}}{=} F^{\leftarrow}\Big(1 - \prod_{j=1}^{m} U_j^{1/\gamma_j}\Big).$$

Therefore, the cumulative distribution function $F^{X_{m:m:n}}$ has the representation

$$F^{X_{m:m:n}}(x) = P\Big(\sum_{j=1}^{m} Z_j \le -\log \overline{F}(x)\Big), \quad x \in \mathbb{R},$$

where $Z_1, \dots, Z_m$ are independent random variables with $Z_j = -\frac{1}{\gamma_j}\log U_j \sim \mathsf{Exp}(\gamma_j)$, $1 \le j \le m$.

In the following, the parameters γ_j may depend on m, which we emphasize by the notation $\gamma_{j,m}$. Therefore, Cramer [290] considered a triangular scheme (3.12) of independent random variables $Z_{j,m} \sim \mathsf{Exp}(\gamma_{j,m})$, where $\gamma_{j,m} = \sum_{i=j}^{m}(R_{i,m}+1)$, $1 \le j \le m$, and $\mathscr{R}_m = (R_{1,m}, \dots, R_{m,m})$ denote the corresponding censoring plan, $m \in \mathbb{N}$.

Exponential random variables	Censoring plans	
$Z_{1,1}$	$R_{1,1}$	
$Z_{1,2}\ Z_{2,2}$	$R_{1,2}\ R_{2,2}$	
$Z_{1,3}\ Z_{2,3}\ Z_{3,3}$	$R_{1,3}\ R_{2,3}\ R_{3,3}$	(3.12)
$Z_{1,4}\ Z_{2,4}\ Z_{3,4}\ Z_{4,4}$	$R_{1,4}\ R_{2,4}\ R_{3,4}\ R_{4,4}$	
$Z_{1,5}\ Z_{2,5}\ Z_{3,5}\ Z_{4,5}\ Z_{5,5}$	$R_{1,5}\ R_{2,5}\ R_{3,5}\ R_{4,5}\ R_{5,5}$	
$\vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \ddots$	$\vdots\ \vdots\ \vdots\ \vdots\ \vdots\ \ddots$	

The problem is to identify the nondegenerate limits L of

$$F^{X_{m:m:n}}(a_m x + b_m) = P\Big(\sum_{j=1}^{m} Z_{j,m} \le -\log \overline{F}(a_m x + b_m)\Big) \xrightarrow{m\to\infty} L(x)$$

for all continuity points x of some nondegenerate cumulative distribution function L. In order to find these distributions, Cramer [290] considered the limiting behavior of the sum $\sum_{j=1}^{m} Z_{j,m}$ and of $-\log \overline{F}(a_m x + b_m)$ separately. Notice that the latter illustrates the relation to the extreme value theory of order statistics.

Introducing the notation $t_m^{(k)} = \sum_{j=1}^{m} \frac{1}{\gamma_{j,m}^k}$, $k \in \mathbb{N}$, we get

$$t_m^{(1)} = E\sum_{j=1}^{m} Z_{j,m}, \quad t_m^{(2)} = \mathsf{Var}\Big(\sum_{j=1}^{m} Z_{j,m}\Big).$$

Cramer [290] showed that two settings have to be handled separately:

$$\lim_{m\to\infty} t_m^{(1)} < \infty, \tag{C1}$$

$$\lim_{m\to\infty} t_m^{(1)} = \infty, \quad \lim_{m\to\infty} t_m^{(2)} < \infty. \tag{C2}$$

Remark 3.4.1. In the model of progressive censoring, only the cases (C1) and (C2) are possible. This can be seen from the relation $\gamma_{j,m} = m - j + 1 + \sum_{i=j}^{m} R_i \geq m-j+1$ or, equivalently, $\frac{1}{\gamma_{j,m}} \leq \frac{1}{m-j+1}$ with equality if the censoring scheme is given by $\mathscr{R}_m = (0^{*m})$. In this sense, progressively Type-II censored order statistics can be seen as dominated by order statistics since $t_m^{(k)} = \sum_{j=1}^{m} \gamma_{j,m}^{-k} \leq \sum_{j=1}^{m} \frac{1}{j^k}$, $k \in \mathbb{N}$. Moreover, it is clear that the limits $t_\infty^{(k)}$, $k \geq 2$, are finite provided they exist.

In Cramer [290], it is shown that both scenarios are possible. As an example, the following one-step plans are considered.

(i) For (C1), the following cases are distinguished:

$$(1)\ t_\infty^{(1)} > 0, t_\infty^{(2)} > 0, \quad (2)\ t_\infty^{(1)} > 0, t_\infty^{(2)} = 0, \quad (3)\ t_\infty^{(1)} = 0, t_\infty^{(2)} = 0.$$

These cases are generated by the following one-step censoring schemes:

(1) $\mathscr{R}_m = (0^{*m-2}, m, 0)$, $m \geq 3$;
(2) $\mathscr{R}_m = (0^{*m-1}, m)$, $m \geq 2$;
(3) $\mathscr{R}_m = (0^{*m-1}, m^2)$, $m \geq 3$.

(ii) For progressively Type-II censored order statistics with equi-balanced censoring plan $\mathscr{R}_m = (R^{*m})$, $\lim_{m\to\infty} t_m^{(1)} = \infty$. For $s \geq 2$, we have

$$t_\infty^{(s)} = \frac{1}{(R+1)^s}\zeta(s) < \infty,$$

where ζ denotes the Riemann-zeta function. For $s = 2$, one gets $t_\infty^{(2)} = \frac{\pi^2}{6(R+1)^2}$.

(iii) As pointed out in (i), $t_\infty^{(2)} = 0$ is possible. However, the censoring schemes $\mathscr{R}_m = (0^{*m-1}, \lfloor\sqrt{m}\rfloor)$, $m \geq 2$, show that this holds also when $t_\infty^{(1)} = \infty$.

Remark 3.4.2. Cramer [285] pointed out that a third scenario is possible for generalized order statistics

$$\lim_{m\to\infty} t_m^{(1)} = \infty, \quad \lim_{m\to\infty} t_m^{(2)} = \infty. \tag{C3}$$

Comparing these settings, it turns out that (C2) is similar to the setup of order statistics provided $\lim_{m\to\infty} t_m^{(2)} > 0$, whereas (C3) resembles the case of record values (see Arnold et al. [59], Nevzorov [680], Resnick [751]).

The scenarios (C1) and (C2) lead to a different limiting behavior: For (C1), a limiting distribution depending on the underlying cumulative distribution function F results if $\lim_{m\to\infty} t_m^{(2)} > 0$. If $\lim_{m\to\infty} t_m^{(2)} = 0$, i.e., the variance of $\sum_{j=1}^{m} Z_{j,m}$ converges to zero, the behavior is connected to central and intermediate order statistics.

In what follows, we establish the limiting results and the associated normalizing sequences $(a_m)_{m\in\mathbb{N}}$ and $(b_m)_{m\in\mathbb{N}}$. For (C1) and (C2), we calculate the characteristic function of the limiting distribution. Given the assumption $\lim_{m\to\infty} t_m^{(2)} = 0$, we present a Lyapunov-type condition to guarantee asymptotic normality. Applying the δ-method, this approach yields generalizations of limiting results for central and intermediate order statistics (see, e.g., Reiss [750]).

3.4.3 Extreme Value Analysis for Exponential Progressively Type-II Censored Order Statistics

The extreme value analysis for progressively Type-II censored order statistics is developed in two steps. First, an exponential population distribution is considered. Let the normalized exponential progressively Type-II censored order statistic $Z_{m:m:n}$ be given by the linear transformation

$$M_m = \frac{Z_{m:m:n} - t_m^{(1)}}{\sqrt{t_m^{(2)}}} = \frac{\sum_{j=1}^m Z_{j,m} - t_m^{(1)}}{\sqrt{t_m^{(2)}}}, \quad m \in \mathbb{N},$$

with $(Z_{j,m})_{j,m}$ as in (3.12). Then, due to the independence assumption, the characteristic function φ_m of M_m is given by

$$\varphi_m(s) = E\mathrm{e}^{isM_m} = \exp\left(-is\frac{t_m^{(1)}}{\sqrt{t_m^{(2)}}}\right)\prod_{j=1}^m \frac{\gamma_{j,m}\sqrt{t_m^{(2)}}}{\gamma_{j,m}\sqrt{t_m^{(2)}} - is}, \quad s \in \mathbb{R}.$$

Let log denote the principal branch of the logarithm in the complex plane. Then, Cramer [290] established the following representation of the characteristic function of the limiting cumulative distribution function L.

Theorem 3.4.3. Let $t_\infty^{(2)} > 0$ be finite. Then, for $|s| < 1$,

(i) the limit $\lim_{m\to\infty}[-\log\varphi_m(s)]$ exists. It is given by

$$\lim_{m\to\infty}[-\log\varphi_m(s)] = \frac{s^2}{2} - \sum_{\nu=3}^\infty \frac{1}{\nu}(is)^\nu \frac{t_\infty^{(\nu)}}{(t_\infty^{(2)})^{\nu/2}} = \kappa(s), \quad \text{say},$$

given that $t_\infty^{(2)}$ is positive and finite. Otherwise, the limit of the ratio $t_m^{(\nu)}/(t_m^{(2)})^{\nu/2}$ has to be considered as a whole in order to find the limiting function;

(ii) the limit $\lim_{m\to\infty}\varphi_m(s)$ exists. It is given by

$$\lim_{m\to\infty} \varphi_m(s) = \exp\{-\kappa(s)\} = \varphi(s), \quad |s| < 1. \tag{3.13}$$

The function φ is uniquely determined on the real line by (3.13) and defines a characteristic function. Moreover, $EM_m^\nu \to EM^\nu$ for all $\nu \in \mathbb{N}$, where $M \sim L$. In particular, $F^{M_m} \xrightarrow{\text{d}} L$, where the characteristic function of L is given by φ.

3.4.4 Extreme Value Analysis for Progressively Type-II Censored Order Statistics from a Cumulative Distribution Function F

Positive Asymptotic Variance

Let $0 < t_\infty^{(2)} < \infty$, $(u_m)_{m\in\mathbb{N}} \subseteq \mathbb{R}$, and $\tau_m = \overline{F}(u_m)$, $m \in \mathbb{N}$. Then,

$$F^{X_{m:m:n}}(u_m) = P\Big(\sum_{j=1}^m C_{j,m} \le -\log\tau_m\Big) = F^{M_m}\Big(\frac{-\log\tau_m - t_m^{(1)}}{\sqrt{t_m^{(2)}}}\Big)$$

From Theorem 3.4.3, it follows that the cumulative distribution functions F^{M_m} converge weakly to a nondegenerate cumulative distribution function L. Let $x \in \mathbb{R}$, $(a_m)_{m\in\mathbb{N}} \subseteq (0,\infty)$, $(b_m)_{m\in\mathbb{N}} \subseteq \mathbb{R}$, and $u_m = u_m(x) = a_m x + b_m$, $m \in \mathbb{N}$, so that

$$z_m(x) = \frac{-\log\tau_m - t_m^{(1)}}{\sqrt{t_m^{(2)}}} = \frac{-\log\overline{F}(a_m x + b_m) - t_m^{(1)}}{\sqrt{t_m^{(2)}}}.$$

If normalizing sequences $(a_m)_{m\in\mathbb{N}} \subseteq (0,\infty)$, $(b_m)_{m\in\mathbb{N}} \subseteq \mathbb{R}$ can be chosen with $\lim_{m\to\infty} z_m(x) = z(x)$, then Slutsky's lemma (cf. Serfling [793, p. 19]) yields that $F^{M_m}(z_m(x)) \longrightarrow L(z(x))$ for all continuity points $z(x)$ of L.

Theorem 3.4.4. Let $t_\infty^{(1)} = \infty$ and $0 < t_\infty^{(2)} < \infty$. Then, $P^{(X_{m:m:n}-a_m)/b_m}$ converges to a nondegenerate limiting distribution iff F is in the domain of attraction of an extreme value distribution H.

In that case, up to an affine transformation, the limiting distribution has the representation $L\big(-\log(-\log H)/\sqrt{t_\infty^{(2)}}\big)$, where L is the limiting distribution of M_m. Appropriate normalizing constants are defined by (3.14).

Proof. First, let $t_\infty^{(1)} < \infty$. Then, $a_m = 1$, $b_m = 0$, $m \in \mathbb{N}$, and we get

$$z_m(x) = \frac{-\log\overline{F}(x) - t_m^{(1)}}{\sqrt{t_m^{(2)}}} \longrightarrow \frac{-\log\overline{F}(x) - t_\infty^{(1)}}{\sqrt{t_\infty^{(2)}}} = z(x), \quad x \in \mathbb{R},$$

so that the limiting distribution depends on the population cumulative distribution function F.

For $t_\infty^{(1)} = \infty$, we get $z_m(x) = -\log\left(\overline{F}(a_m x + b_m)\exp(t_m^{(1)})\right)/\sqrt{t_m^{(2)}}$. Supposing that F is in the domain of attraction of an extreme value distribution H, normalizing sequences $(\alpha_m)_{m\in\mathbb{N}} \subseteq (0,\infty)$, $(\beta_m)_{m\in\mathbb{N}} \subseteq \mathbb{R}$ exists with

$$m\overline{F}(\alpha_m x + \beta_m) \longrightarrow -\log H(x), \quad x \in \mathbb{R}$$

(cf. De Haan and Ferreira [331, p. 4]). Then, $\lim_{m\to\infty} t_m^{(1)} = \infty$ implies $\lim_{m\to\infty}\exp(t_m^{(1)}) = \infty$. Choosing normalizing constants as

$$a_m = \alpha_{\lfloor \exp(t_m^{(1)})\rfloor}, \quad b_m = \beta_{\lfloor \exp(t_m^{(1)})\rfloor}, \qquad m \in \mathbb{N}, \tag{3.14}$$

the limit $\lim_{m\to\infty} z_m(x) = -\log(-\log H(x))/\sqrt{t_\infty^{(2)}}$ holds. Hence,

$$F^{X_{m:m:n}}(z_m(x)) \longrightarrow L\left(-\log(-\log H(x))/\sqrt{t_\infty^{(2)}}\right)$$

for all continuity points $y = -\log(-\log H(x))/\sqrt{t_\infty^{(2)}}$ of L. □

Theorem 3.4.4 illustrates that the possible limiting distributions of progressively Type-II censored order statistics are directly connected to those of order statistics. The problem of deriving the limiting distribution can be divided into two tasks. First, it has to be ensured that the random variable M_m converges to the cumulative distribution function L. In the second step, the underlying cumulative distribution function F has to be analyzed w.r.t. the domain of attraction of an extreme value distribution H. If so, the normalizing sequences given in Theorem 3.4.4 can be used to ensure convergence to a nondegenerate distribution.

Zero Asymptotic Variance

Let $t_\infty^{(2)} = 0$. Then, $t_\infty^{(\nu)} = 0$ for all $\nu \geq 2$. Now, the Lyapunov condition for the central limiting theorem leads to a sufficient condition for convergence. With $Z_{j,m} - \gamma_{j,m}^{-1}$, $1 \leq j \leq m$, the Lyapunov condition reads

$$\frac{1}{(t_m^{(2)})^{3/2}} \sum_{j=1}^{n} E\left|Z_{j,m} - \gamma_{j,m}^{-1}\right|^3 \xrightarrow{n\to\infty} 0.$$

Since $Z_j = \gamma_{j,m} Z_{j,m}$, $1 \leq j \leq m$, are IID Exp(1)-random variables, this simplifies to

$$\sum_{j=1}^{m} E\left|Z_{j,m} - \gamma_{j,m}^{-1}\right|^3 = E|Z_1 - 1|^3 \sum_{j=1}^{m} \frac{1}{\gamma_{j,m}^3} = E|Z_1 - 1|^3 t_m^{(3)}.$$

Hence, the Lyapunov condition is equivalent to assuming that

$$\frac{t_m^{(3)}}{(t_m^{(2)})^{3/2}} \xrightarrow{n\to\infty} 0. \tag{3.15}$$

This condition implies convergence to a normal distribution as well as

$$\frac{t_m^{(\nu)}}{(t_m^{(2)})^{\nu/2}} \xrightarrow{n\to\infty} 0 \quad \text{for all } \nu \geq 3. \tag{3.16}$$

Remark 3.4.5. The ratios in (3.16) are bounded from below by zero and from above by one. In particular, one can show (if all limits exist) that

$$t_m^{(3)}/(t_m^{(2)})^{3/2} \xrightarrow{n\to\infty} c_3 > 0$$

implies $t_m^{(\nu)}/(t_m^{(2)})^{\nu/2} \xrightarrow{n\to\infty} c_\nu > 0$ for all $\nu \geq 3$. Notice that $c_{\nu_0} = 0$ for some $\nu_0 \geq 3$ implies the validity of the Lyapunov condition. Hence, the limiting distribution is either normal or all cumulants for $\nu \geq 2$ of the limiting distribution are positive.

Theorem 3.4.3 yields the following asymptotic result.

Theorem 3.4.6. Suppose condition (3.15) holds. Then,

$$M_m \xrightarrow{\mathsf{d}} \mathsf{N}(0,1). \tag{3.17}$$

If (3.16) holds with limits $c_\nu > 0$, $\nu \geq 2$, then the characteristic function of the limiting law is given by (3.13).

Writing $G = -\log \overline{F}$, Corollary 2.3.7 yields the identity $X_{m:m:n} \stackrel{\mathsf{d}}{=} G^{\leftarrow}\big(\sum_{j=1}^{m} C_{j,m}\big)$. The δ-method (see, e.g., Sen and Singer [791, p. 131]), Theorem 3.4.6 and Slutsky's lemma imply the following result.

Theorem 3.4.7. Let F be an absolutely continuous and continuously differentiable cumulative distribution function with density function f. Suppose that (3.15) holds and that $\xi > 0$ exists with

$$\frac{\xi - t_m^{(1)}}{\sqrt{t_m^{(2)}}} \xrightarrow{m\to\infty} 0, \tag{3.18}$$

and $f(G^{\leftarrow}(\xi)) > 0$. Then, the mth progressively Type-II censored order statistic from F is asymptotically normal, i.e.,

$$\frac{X_{m:m:n} - G^{\leftarrow}(\xi)}{\sqrt{t_m^{(2)}}} \xrightarrow{\mathsf{d}} \mathsf{N}\left(0, \mathrm{e}^{-2\xi} f^{-2}(G^{\leftarrow}(\xi))\right).$$

Appropriate normalizing constants are given by $a_m = \sqrt{t_m^{(2)}}$ and $b_m = G^{\leftarrow}(\xi)$, $m \in \mathbb{N}$.

Condition (3.18) implies that $t_\infty^{(1)} = \xi > 0$ since $t_\infty^{(2)} = 0$. For $t_\infty^{(1)} \in \{0, \infty\}$, the situation is different. As an example, consider $t_m^{(1)} \to \infty$, $G^{\leftarrow}(t_m^{(1)})$ tends to $G^{\leftarrow}(\infty) = F^{\leftarrow}(1) = \omega(F)$. Then, for $\omega(F) < \infty$, a convergence result can be obtained (for order statistics, cf. Reiss [750, p. 109], Smirnov [808], and Falk [358]). A similar result can be established for $t_\infty^{(1)} = 0$ and $\alpha(F) > -\infty$.

Theorem 3.4.8. Suppose F is absolutely continuous with density function f and that F has a finite right endpoint $\omega(F)$ of its support. Let F be differentiable as well as f be uniformly continuous on $(\omega(F) - \varepsilon, \omega(F))$ for some $\varepsilon > 0$. Moreover, let $f(t) > \delta$ for $t \in (\omega(F) - \varepsilon, \omega(F))$ and some $\delta > 0$. Then, if (3.17) holds with $t_\infty^{(1)} = \infty$, then $X_{m:m:n} \xrightarrow{\mathsf{d}} \mathsf{N}(0, 1)$. Appropriate normalizing constants are given by

$$a_m = \frac{\sqrt{t_m^{(2)}}}{(G^{\leftarrow})'(t_m^{(1)})}, \quad b_m = G^{\leftarrow}(t_m^{(1)}), \quad m \in \mathbb{N}.$$

The conditions of Theorem 3.4.8 are satisfied by the uniform distribution. The normalizing constants $a_{m,\mathrm{uni}} = \sqrt{t_m^{(2)}} \exp(-t_m^{(1)})$ and $b_{m,\mathrm{uni}} = 1 - \exp(-t_m^{(1)})$, $m \in \mathbb{N}$, ensure asymptotic normality.

3.4.5 Applications to Upper, Lower, Central, and Intermediate Progressively Type-II Censored Order Statistics

Cramer [290] has shown that the preceding results can be applied not only to upper extreme progressively Type-II censored order statistics but also to upper, lower, central, and intermediate progressively Type-II censored order statistics.

Upper Progressively Type-II Censored Order Statistics

The results for the maximal progressively Type-II censored order statistic $X_{m:m:n}$ can be directly applied to an upper progressively Type-II censored order statistic $X_{m-r+1:m:n}$ with $r \in \mathbb{N}, r \leq m$ fixed. As mentioned in Nasri-Roudsari and Cramer [671] for m-generalized order statistics, this corresponds to a sample of progressively Type-II censored order statistics with $\gamma_{1,m}, \ldots, \gamma_{m-r+1,m}$ or censoring scheme $\mathscr{R}_m^* = (R_{1,m}, \ldots, R_{m-r,m}, \gamma_{m-r+1,m})$ (see Theorem 2.4.1). Therefore, the preceding theory can directly be applied by defining the sums

$$t_{m-r+1}^{(\nu)} = \sum_{j=1}^{m-r+1} \frac{1}{\gamma_{j,m}^{\nu}}, \quad \nu \in \mathbb{N}.$$

In the case of equi-balanced censoring plans, i.e., $\gamma_{j,m} = (m-j+1)(R+1)$, $1 \leq j \leq m$, the result for the convergence of the exponential sum is given by

$$L(z) = \frac{1}{\Gamma(r)} \Gamma\left(r, z^{R+1}\right), \quad z \in \mathbb{R},$$

which is the cumulative distribution function of a generalized three-parameter gamma distribution introduced by Stacy [823] (cf. Nasri-Roudsari [670], Nasri-Roudsari and Cramer [671]). For order statistics, this yields the result of Smirnov [807].

Lower Progressively Type-II Censored Order Statistics

As above, we have to restrict the censoring scheme to the particular scenario. For $1 \leq r \leq m$ fixed, it is clear that the progressively Type-II censored order statistic $X_{r:m:n}^{\mathscr{R}_m}$ is only affected by the parameters $\gamma_{1,m}, \ldots, \gamma_{r,m}$. Now, the sums $t_r^{(\nu)}$ have only a finite number of terms so that convergence of the parameters $\gamma_{1,m}, \ldots, \gamma_{r,m}$ has to be considered. Cramer [290] has shown that for $(\delta_m)_{m \in \mathbb{N}}$ with $\lim_{m\to\infty} \delta_m = \infty$ and $\lim_{m\to\infty} \gamma_{j,m}/\delta_m = \eta_j \in \mathbb{N}$, $1 \leq j \leq r$,

$$Z_m^* = \delta_m Z_{r:m:n}^{\mathscr{R}_m} = \delta_m \sum_{j=1}^{r} \frac{1}{\gamma_{j,m}} Z_j \xrightarrow{\mathsf{d}} \sum_{j=1}^{r} \frac{1}{\eta_j} Z_j,$$

where $Z_1, \ldots, Z_r$ are IID standard exponential random variables. Since the limit on the right can be interpreted as exponential generalized order statistics with parameters $\eta_1, \ldots, \eta_r$, it follows that exponential generalized order statistics can always be seen as a limit of exponential progressively Type-II censored order statistics. Since

$$F^{X_{r:m:n}}(a_m t + b_m) = P\left(Z_m^* \leq -\log \overline{F}^{\delta_m}(a_m t + b_m)\right)$$

and $\delta_m \to \infty$, it follows that a nondegenerate limiting law results iff F is in the domain of attraction of a minimum-stable distribution. If $\lim_{m\to\infty} \delta_m < \infty$, then the limiting distribution is either degenerate or it depends on the baseline distribution F.

The preceding conditions are satisfied for equi-balanced censoring schemes with $\delta_m = m(R+1)$, $m \in \mathbb{N}$ (for m-generalized order statistics, see Houben [452]). Provided that F is in the domain of attraction of a minimum-stable cumulative distribution function H, the limiting distribution is a transformed gamma distribution

$$L(x) = \frac{1}{\Gamma(r)} \Gamma\left(r-1, -\log(1-H(x))\right).$$

Central and Intermediate Progressively Type-II Censored Order Statistics

Here, $r = r_m$ is increasing w.r.t. m so that $\lim_{m\to\infty} r_m = \infty$ and $\lim_{m\to\infty}(m - r_m) = \infty$. As above, the sums

$$t_{m-r_m+1}^{(\nu)} = \sum_{j=1}^{m-r_m+1} \frac{1}{\gamma_{j,m}^{\nu}} = \sum_{j=r_m}^{m} \frac{1}{\gamma_{m-j+1,m}^{\nu}}, \quad \nu \in \mathbb{N},$$

have to be studied. As for order statistics, the resulting progressively Type-II censored order statistics are called central progressively Type-II censored order statistics if $\lim_{m\to\infty} \frac{r_m}{m} = \xi \in (0,1)$ and intermediate progressively Type-II censored order statistics if $\lim_{m\to\infty} \frac{r_m}{m} = \xi \in \{0,1\}$.

For equi-balanced censoring schemes, Cramer [290] showed that

$$\frac{t_{m-r_m+1}^{(\nu)}}{(t_{m-r_m+1}^{(2)})^{\nu/2}} \approx \frac{1}{r_m^{\nu/2-1}} \frac{1-\left(\frac{r_m}{m}\right)^{\nu-1}}{\left(1-\frac{r_m}{m}\right)^{\nu/2}} \to 0 \text{ for } \nu \geq 3, \ \frac{r_m}{m} \to \xi \in [0,1),$$

which illustrates that normal laws are involved in the limiting distribution.

Let $r_m/m \to \xi \in (0,1)$. If condition (3.18) holds, Theorem 3.4.7 can be applied showing that

$$\sqrt{n}(X_{m-r_m+1:m:n} - F^{\leftarrow}(1-\xi^{R+1})) \xrightarrow{d} \mathsf{N}\left(0, \frac{\xi^{-2R} f^{-2}(F^{\leftarrow}(1-\xi^{R+1}))}{(R+1)^2 \xi(1-\xi)}\right).$$

Since the expectation of the non-normalized progressively Type-II censored order statistic is approximately $F^{\leftarrow}(1-\xi^{R+1})$, this result illustrates that, under certain regularity conditions, a central progressively Type-II censored order statistic can be used to approximate a quantile of the baseline cumulative distribution function F. For order statistics, the preceding result was obtained by Smirnov [807].

Suppose now $\frac{r_m}{m} \to 0$. For intermediate progressively Type-II censored order statistics, the following corollary is deduced from Theorem 3.4.8.

Corollary 3.4.9. Suppose F is absolutely continuous with density function f and that F has a finite right endpoint $\omega(F)$ of its support. Let F be differentiable as well as f be uniformly continuous on $(\omega(F)-\varepsilon, \omega(F))$ for some $\varepsilon > 0$. Moreover, let $f(t) > \delta$ for $t \in (\omega(F)-\varepsilon, \omega(F))$ and for some $\delta > 0$. Then, $\frac{X_{r_m:m:n}-a_m}{b_m}$ is asymptotically normal. Appropriate normalizing constants are given by

$$a_m = \left(\frac{r_m}{m}\right)^{\frac{1}{R+1}} \sqrt{\frac{m-r_m}{m r_m}} \frac{1}{(R+1) f(F^{\leftarrow}(1-(r_m/m)^{1/(R+1)}))},$$

$$b_m = F^{\leftarrow}\left(1-\left(\frac{r_m}{m}\right)^{\frac{1}{R+1}}\right), \quad m \in \mathbb{N}.$$

The assumptions of Corollary 3.4.9 are fulfilled for the uniform distribution so that intermediate uniform progressively Type-II censored order statistics are asymptotically normal for equi-balanced censoring schemes. In this case, the normalizing constants are given by

$$a_{m,\text{uni}} = \frac{1}{(R+1)}\left(\frac{r_m}{m}\right)^{\frac{1}{R+1}} \sqrt{\frac{m-r_m}{m r_m}}, \quad b_{m,\text{uni}} = 1-\left(\frac{r_m}{m}\right)^{\frac{1}{R+1}}.$$

Remark 3.4.10. Intermediate order statistics have been investigated by many authors including Čibisov [265], Wu [903], and Smirnov [808]. Falk [358] has given a very general result involving von Mises-type conditions, which leads to simple representations of the normalizing constants as well. Moreover, this paper provides a good survey of the literature on this topic.

3.4.6 Limits for Central Progressively Type-II Censored Order Statistics with Blocked Observations

Hofmann et al. [444] proposed a progressive censoring model with blocked observations represented by the numbers $\overline{R}_1, \ldots, \overline{R}_m$. This model corresponds to progressive Type-II censoring with censoring scheme

$$\mathscr{R}_m = (0^{*\overline{R}_1-1}, R_1, 0^{*\overline{R}_2-1}, R_2, \ldots, 0^{*\overline{R}_m-1}, R_m).$$

The initial sample size may be reduced by withdrawing R_0 items before the experiment starts. Hence, $\gamma_1 = n_\bullet = n - R_0$. The observed sample size is given by $m_\bullet = \overline{R}_{\bullet m} = \sum_{j=1}^m \overline{R}_j$. The number of censored items is given by $R_{\bullet m} = \sum_{j=1}^m R_j$. Hence, $n = R_0 + \sum_{j=1}^m (\overline{R}_j + R_j)$ is the original sample size.

The standard model is included by choosing $\overline{R}_j = 1$, $j = 1, \dots, m$. The resulting sample will be denoted by the random variables $X_{1:m_\bullet:n_\bullet}, \dots, X_{m_\bullet:m_\bullet:n_\bullet}$.

The asymptotic scenario is now designed as follows. Let $n \to \infty$ and the number of blocks m be fixed. Then, the limiting proportional block sizes and the limiting proportional censoring sizes are defined by

$$\overline{\lambda}_i = \lim_{n\to\infty} \frac{\overline{R}_i}{n}, \ i = 1, \dots, m, \quad \lambda_i = \lim_{n\to\infty} \frac{R_i}{n}, \ i = 0, \dots, m,$$

respectively.

In order to get nondegenerate limits, suppose $\lambda_i \in [0, 1)$ and $\overline{\lambda}_i \in (0, 1)$ for $1 \le i \le m$ and the last included observation is a central progressively Type-II censored order statistic leading to the assumption $\lambda_m > 0$. Then, $\lambda_0 + \sum_{i=1}^m (\overline{\lambda}_i + \lambda_i) = 1$. Further, let

$$S_i = \sum_{j=i}^m (\overline{R}_j + R_j), \quad \overline{R}_{\bullet i} = \sum_{j=1}^i \overline{R}_j, \quad \delta_i = \sum_{j=i}^m (\overline{\lambda}_j + \lambda_j), \ i = 1, \dots, m;$$

$$p_i = 1 - \frac{\overline{\lambda}_i}{\delta_i}, \ i = 1, \dots, m; \tag{3.19}$$

$$t_0 = 1 - \lambda_0, \ t_i = 1 - \frac{\lambda_i}{\delta_i - \overline{\lambda}_i}, \ i = 1, \dots, m-1.$$

Therefore, $p_i \in (0, 1)$, $t_i \in (0, 1]$ for all i, without additional restrictions. The proportion of remaining units still in the experiment before the ith block of failures is represented by δ_i. The proportion of the ith block of failures when compared to all remaining observations at this point is denoted by p_i. Similarly, t_i stands for the proportion of remaining items being censored after the ith block. It should be mentioned that if no progressive censoring and no reduction of the sample size takes place, $R_i = \lambda_i = 0$ for $i = 0, \dots, m-1$. This corresponds to the setting $t_0 = \dots = t_{m-1} = 1$.

Let $Y_{i:m:n}$ be the last failure time observed in the ith block, i.e., $Y_{i:m:n} = X_{\overline{R}_{\bullet i}:m_\bullet:n_\bullet}$, $i = 1, \dots, m$. Hence, the data are given by the progressive block Type-II censored sample

$$\mathbf{Y}_\bullet^{\mathscr{R}} = (Y_{1:m:n}, \dots, Y_{m:m:n}) = (X_{\overline{R}_{\bullet i}:m_\bullet:n_\bullet}, \dots, X_{\overline{R}_{\bullet m}:m_\bullet:n_\bullet}).$$

It is important for the statistical analysis that only the largest observed failure time within each block and not all the failures in each block are available.

It is clear that the above sample is a marginal sample. Assuming that the population cumulative distribution function F is absolutely continuous with density function f, the density function of $\mathbf{Y}_\bullet^{\mathscr{R}}$ is given by

$$f^{\mathbf{Y}_\bullet^{\mathscr{R}}}(\mathbf{y}_m) = c\prod_{i=1}^m (F(y_i)-F(y_{i-1}))^{\overline{R}_i-1}(1-F(y_i))^{R_i} f(y_i),\ y_1 \le \dots \le y_m,$$

where $F(y_0)=0$, $c=\prod_{i=1}^m \frac{S_i!}{(\overline{R}_i-1)!(S_i-\overline{R}_i)!}$.

If the population distribution is exponential, Hofmann et al. [444] showed that the density function of the spacings $V_1 = Z_{1:m_\bullet:n_\bullet}$, $V_j = Z_{j:m_\bullet:n_\bullet} - Z_{j-1:m_\bullet:n_\bullet}$, $j=2,\dots,m$, is given by

$$f^{V_1,\dots,V_m}(\mathbf{v}_m) = \prod_{i=1}^m \left(\frac{S_i!}{(\overline{R}_i-1)!(S_i-\overline{R}_i)!}(1-\mathrm{e}^{-v_i})^{\overline{R}_i-1}\mathrm{e}^{-v_i(S_i-\overline{R}_i)}\right),$$

where $v_1,\dots,v_m \ge 0$. Therefore, the resulting distribution corresponds to the joint distribution of m exponential order statistics $Z_{\overline{R}_1:S_1},\dots,Z_{\overline{R}_m:S_m}$ from m independent samples [see (2.30)].

Since $\overline{R}_i \approx n\overline{\lambda}_i$ and $S_i \approx n\delta_i$, we obtain

$$\lim_{n\to\infty} EV_i = -\log\left(1-\frac{\overline{\lambda}_i}{\delta_i}\right) = -\log p_i,$$

$$\lim_{n\to\infty} n\,\mathsf{Var}(V_i) = \frac{1}{\delta_i-\overline{\lambda}_i} - \frac{1}{\delta_i} = \nu_i.$$

The asymptotic joint distribution of the appropriately normalized progressive block Type-II censored order statistics $\mathbf{Y}_\bullet^{\mathscr{R}}$ is obtained by applying the central limit theorem and the δ-method. Let $\Sigma = \mathrm{diag}(\nu_1,\dots,\nu_m)$, $I_m = \mathrm{diag}(1,\dots,1)$, and $D=(d_{ij})_{i,j}$ be a lower triangular matrix such that $d_{ij}=1$, $1\le i\le j\le m$, and $d_{ij}=0$ otherwise.

Theorem 3.4.11. Let $\mathbf{Y}_\bullet^{\mathscr{R}}$ be a progressive block Type-II censored sample drawn from an absolutely continuous differentiable distribution function F with density function f. Moreover, let $u_i = F^{\leftarrow}(1-\prod_{j=1}^i p_j)$ and $\Delta = \mathrm{diag}(\Delta_1,\dots,\Delta_m)$ with $\Delta_i^{-1} = f(u_i)/\prod_{j=1}^i p_j$, $1\le i\le m$. Suppose $f(u_i)\neq 0$, $1\le i\le m$, and the following condition holds:

$$\sqrt{n}\Big(-\log p_i - \sum_{j=1}^{\overline{R}_i}\frac{1}{S_i-j+1}\Big) \overset{n\to\infty}{\longrightarrow} 0, \quad 1\le i\le m. \tag{3.20}$$

Then,

$$\sqrt{n}(\mathbf{Y}_\bullet^{\mathscr{R}}-\mathbf{u}) \overset{\mathrm{d}}{\longrightarrow} \mathsf{N}_m(\mathbf{0}, \Delta D\Sigma D'\Delta),$$

where $\mathsf{N}_m(\mu, A)$ denotes a multivariate normal distribution with expectation μ and variance–covariance matrix A.

Remark 3.4.12. Hofmann et al. [444] pointed out the following properties:

(i) Condition (3.20) can be replaced by

$$\sqrt{n}\left(\log(1-\overline{R}_i/S_i)-\log p_i\right)\overset{n\to\infty}{\longrightarrow}0,\ \ 1\le i\le m.$$

Assuming that $\overline{R}_i=\lfloor n\overline{\lambda}_i\rfloor+1$, $R_i=\lfloor n\lambda_i\rfloor+1$, $1\le i\le m$, for given values of λ_i and $\overline{\lambda}_i$, respectively, this assertion holds;

(ii) If $\lambda_1=\cdots=\lambda_{m-1}=0$, i.e., no progressive censoring takes place, it is easy to show that $\prod_{j=1}^{i}p_j=\sum_{j=i+1}^{m}\overline{\lambda}_j+\lambda_m=\delta_i=1-q_i$, say, $1\le i\le m-1$, and $\delta_m=\lambda_m>0$, $q_m=1-\lambda_m$. Thus, the above expressions simplify to

$$u_i=F^{\leftarrow}\Big(1-\prod_{j=1}^{i}p_j\Big)=F^{\leftarrow}(q_i)\quad\text{and}\quad \Delta_i=\frac{1-q_i}{f(F^{\leftarrow}(q_i))},$$

$$\nu_i=\frac{1}{1-q_i}-\frac{1}{1-q_{i-1}},\quad 1\le i\le m,\quad q_0=0.$$

This yields the results of Mosteller [658].

3.5 Near Minimum Progressively Type-II Censored Order Statistics

Generalizing the notion of near order statistics (see Pakes and Steutel [699] and Pakes and Li [698], and [124], Balakrishnan and Stepanov [125] introduced the quantity

$$\mu_{m,n}(w)=\#\{X_{j:m:n}\mid X_{1:m:n}<X_{j:m:n}<X_{1:m:n}+w,2\le j\le m\}$$

for some $w>0$. It describes the number of progressively Type-II censored order statistics which have a distance of at most w from the minimum of the data. Balakrishnan and Stepanov [125] motivated their analysis of $\mu_{m,n}(w)$ by interpreting it as a measure for the number of "poor-quality items". It seems reasonable that early failures correspond to units with poor quality. Hence, those items which fail close to the minimum can be classified as such items, and $\mu_{m,n}(w)$ seems to be a reasonable quality measure.

Denoting by $S_{k,i}^{\star\mathscr{R}}=X_{k:m:n}-X_{i:m:n}$ the (i,k)-spacing of the progressively Type-II censored order statistics, it follows that

$$P(\mu_{m,n}(x)\ge k)=P(S_{k+1,1}^{\star\mathscr{R}}\le x),\quad x>0.$$

An explicit expression for this probability can be derived directly from (2.39) with $r=1$ and $s=k+1$, i.e., from the representation

$$F^{S^{\star\mathscr{R}}_{k+1,1}}(w) = 1 - \int_{\mathbb{R}} H_{k+1}\Big(\frac{1-F(v+w)}{1-F(v)}\Big)\gamma_1 \overline{F}^{\gamma_1-1}(v)dF(v)\,,$$

with $H_{k+1}(z) = \big(\prod_{j=2}^{k+1}\gamma_j\big)\sum_{i=2}^{k+1} a^{(1)}_{i,k+1}\frac{1}{\gamma_i}z^{\gamma_i}$, $z \in (0,1)$. Hence,

$$P(\mu_{m,n}(w) < k) = \Big(\prod_{j=1}^{k+1}\gamma_j\Big)\sum_{i=2}^{k+1} a^{(1)}_{i,k+1}\frac{1}{\gamma_i}\int_{\mathbb{R}} \overline{F}^{\gamma_i}(v+w)\overline{F}^{\gamma_1-\gamma_i-1}(v)dF(v)\,.$$

Balakrishnan and Stepanov [125] studied asymptotic properties of this distribution provided that the sample size n tends to infinity. In particular, they assumed a scheme similar to the triangle scheme of censoring plans given in (3.12), i.e., $\mathscr{R}_{m_j} = (R_{1,m_j},\ldots,R_{m_j,m_j})$ and $\gamma_{k,m_j} = \sum_{i=k}^{m_j}(R_{i,m_j}+1)$, $1 \le k \le m_j$, $j \in \mathbb{N}$. Assuming that the left endpoint $\alpha(F)$ of the baseline distribution is finite, they proved the following result.

Theorem 3.5.1. Let $w > 0$ be fixed, $n_j = \gamma_{1,m_j}$, and $\lim_{j\to\infty} n_j = \infty$, $\lim_{j\to\infty} m_j = \infty$. Moreover, suppose $\lim_{j\to\infty}\max_{1\le k\le m_j} R_{k,m_j}/n_j = 0$. Then, $\mu_{m_j,n_j}(w)$ converges to zero in probability as $j \to \infty$.

For $\alpha(F) = -\infty$, they showed that $\mu_{m_j,n_j}(w) \xrightarrow{j\to\infty} 0$ almost surely provided that

$$\int_{\mathbb{R}} \frac{F(x+w)-F(x)}{F^2(x)}dF(x) < \infty.$$

In this area, this type of condition has first been established in Li [585] who investigated almost sure convergence of number of records near the maxima (convergence in probability has been established in Pakes and Steutel [699]). Further, given the condition

$$\lim_{x\to-\infty}\frac{F(x)}{F(x+w)} = \beta \tag{3.21}$$

with $\beta = 1$, they found convergence in probability. Finally, they considered the situation that the sequences of censoring numbers $(R_{k,m_j})_k$ are bounded for any k, i.e., given a sequence $(B_k)_{k\in\mathbb{N}}$ of bounds, they assumed that $\sup_{j\in\mathbb{N}} R_{k,m_j} \le B_k < \infty$, $k \in \mathbb{N}$.

Theorem 3.5.2. Let $w > 0$ be fixed and $\lim_{j\to\infty} m_j = \infty$. Suppose that a sequence $(B_k)_{k\in\mathbb{N}}$ of bounds exists such that $\sup_{j\in\mathbb{N}} R_{k,m_j} \le B_k < \infty$, $k \in \mathbb{N}$. Moreover, assume that the limit in (3.21) exists with $\beta \in (0,1)$. Then, $\mu_{m_j,n_j}(w)$ converges in distribution to a random variable Y as $j \to \infty$, where Y is geometrically distributed with probability β.

Chapter 4
Progressive Type-I Censoring: Basic Properties

Progressive Type-I censoring, as introduced in Sect. 1.1.2, generalizes Type-I right censoring by introducing additional inspection times $T_1 < \cdots < T_{k-1}$ where some units in the life test are removed from the experiment according to a prespecified censoring scheme $\mathscr{R}^0 = (R_1^0, \ldots, R_{k-1}^0)$. The threshold T_k defines the termination time of the life test so that all observed failures do not exceed T_k. The generation process is depicted in Fig. 1.6.

As mentioned above, the model reduces to Type-I right censoring for $k = 1$. In this case, given n lifetimes $X_1, \ldots, X_n$, the sample size is random and defined by the random variable $M = \sum_{i=1}^n \mathbb{1}_{(-\infty,T_1]}(X_{i:n})$. Assuming that $M \geq 1$, the data is given by the sample $X_{1:n}, \ldots, X_{M:n}$ of order statistics with random sample size M. This kind of data is extensively studied in survival analysis (see, e.g., Klein and Moeschberger [536]).

The generalization to progressive Type-I censored data has been first proposed in Cohen [267] and further discussed in a series of papers in the late 1960s and early 1970s. Most results are related to likelihood inference as detailed in Chap. 13. In the following, we present a detailed account of the distribution theory of progressively Type-I censored order statistics. We assume throughout that the original lifetimes $X_1, \ldots, X_n$ are IID random variables with cumulative distribution function F. For some purposes, we make additional assumptions like absolute continuity on F with density function f.

4.1 Distribution and Block Independence

As for Type-I right censoring, it may happen that all random variables are (progressively) censored during the experiment and no failure is observed. This happens if all items which fail until T_k are progressively censored at one of the censoring times $T_1, \ldots, T_{k-1}$ before they fail. This event occurs with a positive probability P_O. At this point, it is worth mentioning that the originally planned

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_4,

censoring scheme $\mathscr{R}^0 = (R_1^0, \ldots, R_{k-1}^0)$ and the effectively applied censoring plan $\mathscr{R} = (R_1, \ldots, R_k)$ do not coincide. In particular, the dimension is different. This is the result of the construction process since the number of surviving units at T_k is random and, thus, cannot be specified in advance. However, in a particular life test, R_k is a random variable with (random) support $\{0, \ldots, n - R_{\bullet k-1}\}$, where $R_{\bullet j} = \sum_{i=1}^{j} R_i$, $j = 0, \ldots, k-1$. Further, let $R_k = n - R_{\bullet k-1}$.

In order to calculate P_O, suppose for the moment $X_{R^0_{\bullet j-1}+1}, \ldots, X_{R^0_{\bullet j}}$ are progressively censored at time T_j, $1 \le j \le k$. Then, the event *no observation* is given by the conditions

$$\min(X_{R^0_{\bullet j-1}+1}, \ldots, X_{R^0_{\bullet j}}) \ge T_j, \quad j = 1, \ldots, k,$$

so that the corresponding probability is given by $\prod_{i=1}^{k}[1-F(T_i)]^{R_i^0}$. Notice that this value does not depend on the indices $R^0_{\bullet j-1}+1, \ldots, R^0_{\bullet j}$ of the particular observed lifetimes $X_{R^0_{\bullet j-1}+1}, \ldots, X_{R^0_{\bullet j}}$. Now, a random permutation π of $\{1, \ldots, n\}$ yields the respective result in general so that

$$P_O = P(\text{no observation}) = \frac{1}{n!}\sum_{\pi \in \mathfrak{S}_n}\prod_{i=1}^{k}[1 - F(T_i)]^{R_i^0} = \prod_{i=1}^{k}\overline{F}^{R_i^0}(T_i). \qquad (4.1)$$

Denoting by $D_1, \ldots, D_k$ the (random) number of observations in the intervals $(-\infty, T_1), [T_1, T_2), \ldots, [T_{k-1}, T_k]$ (see Sect. 1.1.2), P_O can be written as $P_O = P(D_j = 0, 1 \le j \le k)$. Now, we study the joint distribution of the progressively Type-I censored order statistics $X^{\mathscr{R},T}_{1:M:n}, \ldots, X^{\mathscr{R},T}_{M:M:n}$ and $D_1, \ldots, D_k$ provided that $d_{\bullet k} \ge 1$ ($x_i \in \mathbb{R}$, $i = 1, \ldots, d_{\bullet k}$)

$$P(X^{\mathscr{R},T}_{i:d_{\bullet k}:n} \le x_i, 1 \le i \le d_{\bullet k}, D_j = d_j, 1 \le j \le k).$$

Notice that with $D_j = d_j$, $j = 1, \ldots, k$, the sample is given by $X^{\mathscr{R},T}_{i:d_{\bullet k}:n}$, $1 \le i \le d_{\bullet k}$ and $M = D_{\bullet k}$. Thus, specifying the outcomes of $D_1, \ldots, D_k$, the random variable M is fixed. Denoting by $U^{\mathscr{R},F(T)}_{j:M:n}$ the corresponding progressively Type-I censored order statistic from a uniform distribution with censoring times $F(T_1), \ldots, F(T_k)$, we arrive at

$$\begin{aligned} &P(X^{\mathscr{R},T}_{i:d_{\bullet k}:n} \le x_i, 1 \le i \le d_{\bullet k}, D_j = d_j, 1 \le j \le k) \\ &\qquad = P(U^{\mathscr{R},F(T)}_{i:d_{\bullet k}:n} \le F(x_i), 1 \le i \le d_{\bullet k}, D_j = d_j, 1 \le j \le k) \end{aligned}$$

provided that $\mathbf{d}_k \in \mathfrak{D}$ with $d_{\bullet k} \ge 1$ where

$$\mathfrak{D} = \{\mathbf{a}_k \in \mathbb{N}_0^k : 0 \le a_i \le [n - a_{\bullet i-1} - R^0_{\bullet i-1}]_+, i = 1, \ldots, k\}. \qquad (4.2)$$

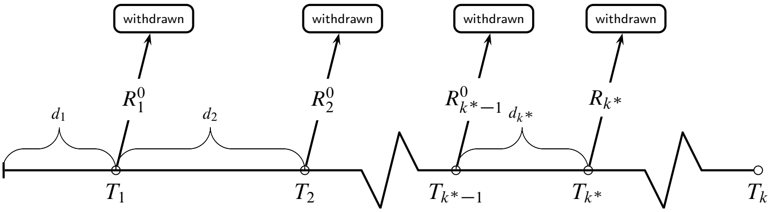

Fig. 4.1 Generation process of progressively Type-I censored order statistics with numbers of observed failures

A formal proof of this identity can be carried out along the lines of the corresponding result for progressive Type-II censoring established in Balakrishnan and Dembińska [96, 97] (see also the proof of Theorem 2.1.1). This tells us that it is sufficient to study uniform distribution. We get the following result which is similar to Theorem 2.1.1 for the Type-II censoring setup.

Theorem 4.1.1. Denote by $F^{\leftarrow}$ the quantile function of the baseline cumulative distribution function F. Then,

$$\big(X_{1:M:n}^{\mathscr{R},T},\ldots,X_{M:M:n}^{\mathscr{R},T}\big) \stackrel{\mathrm{d}}{=} \big(F^{\leftarrow}(U_{1:M:n}^{\mathscr{R},F(T)}),\ldots,F^{\leftarrow}(U_{M:M:n}^{\mathscr{R},F(T)})\big).$$

The subsequent discussion makes use of the ideas of Iliopoulos and Balakrishnan [469] who established conditional independence of the blocks of random variables, given $(D_1,\ldots,D_k) = \mathbf{d}_k \in \mathfrak{D}$, where $\mathfrak{D}$ is defined in (4.2). We adapt their arguments to establish the joint distribution of progressively Type-I censored order statistics.

For $\mathbf{d}_k \in \mathfrak{D}$, denote by T_{k^*} the censoring time point specified by the index

$$k^* = k^*(\mathbf{d}_k) = \max\{1 \le \ell \le k | n - d_{\bullet\ell-1} - R_{\bullet\ell-1}^0 > 0\}.$$

By definition it follows that either T_{k^*} is the final censoring time or the last failure time has been observed in $(T_{k^*-1}, T_{k^*}]$. Notice that this time is fixed given $\mathbf{d}_k$ and that the definition of k^* implies $d_j = 0$ for $k^* < j \le k$ and $\mathbf{d}_k \in \mathfrak{D}$. Based on Figs. 1.6 and 1.8, the situation is depicted in Fig. 4.1.

Before presenting a representation for the preceding probability, we establish the following theorem.

Theorem 4.1.2. Let $\mathbf{d}_k \in \mathfrak{D}$ with $d_{\bullet k} \ge 1$, and $U_1,\ldots,U_n$ be IID U(0,1)-random variables, and $0 = T_0 < T_1 < \cdots < T_k \le 1$.

Given $(D_1,\ldots,D_k) = \mathbf{d}_k$, we have $M = d_{\bullet k}$ and the progressively Type-I censored order statistics $U_{1:M:n}^{\mathscr{R},T},\ldots,U_{M:M:n}^{\mathscr{R},T}$ to be block independent, i.e., $U_{d_{\bullet j-1}+1:d_{\bullet k}:n}^{\mathscr{R},T},\ldots,U_{d_{\bullet j}:d_{\bullet k}:n}^{\mathscr{R},T}$, $1 \le j \le k^*$, are mutually independent with, for $d_j \ge 1$,

$$(U^{\mathscr{R},T}_{d_{\bullet j-1}+1:d_{\bullet k}:n}, \dots, U^{\mathscr{R},T}_{d_{\bullet j}:d_{\bullet k}:n}) \stackrel{\mathrm{d}}{=} (U^{(j)}_{1:d_j}, \dots, U^{(j)}_{d_j:d_j}),$$

where $U^{(j)}_{1:d_j}, \dots, U^{(j)}_{d_j:d_j}$ are order statistics from a uniform $\mathsf{U}(T_{j-1}, T_j)$-distribution, $1 \le j \le k^*$.

Proof. For brevity, let $d_{\bullet k} = m$. In order to prove the result, we consider the probability

$$q = P(U^{\mathscr{R},T}_{\ell:m:n} \le u_\ell, 1 \le \ell \le m, D_j = d_j, 1 \le j \le k) \tag{4.3}$$

for $0 \le u_1 \le \dots \le u_m \le 1$, where $\mathbf{d}_k \in \mathfrak{D}$ with $m = d_{\bullet k} \ge 1$. First, notice that we can replace k by k^* without loss of generality so that $n_\ell = n - d_{\bullet \ell-1} - R^0_{\bullet \ell-1} > 0$, $\ell = 1, \dots, k^*$. From the definition of $D_1, \dots, D_{k^*}$, it follows that

(1) exactly d_j ordered observations are located in the interval $[T_{j-1}, T_j)$, $1 \le j \le k^*$;
(2) R_j objects are progressively censored at time T_j, $1 \le j \le k^*$, where $R_j = R^0_j$, $1 \le j \le k^* - 1$ and $R_{k^*} = n - d_{\bullet k^*} - R^0_{\bullet k^*-1}$.

First, we assign a particular outcome of the progressively Type-I censored experiment to the progressively Type-I censored order statistics. Suppose, for $1 \le j \le k^*$,

(1) $U_{d_{\bullet j-1}+1} \le \dots \le U_{d_{\bullet j}}$ are observed in $[T_{j-1}, T_j)$ and
(2) $U_{m+R_{\bullet j-1}+1}, \dots, U_{m+R_{\bullet j}}$ are progressively censored at T_j.

This corresponds to the events

(1) $T_{j-1} \le U_{d_{\bullet j-1}+1} \le \dots \le U_{d_{\bullet j}} < T_j$ and
(2) $\min\{U_{m+R_{\bullet j-1}+1}, \dots, U_{m+R_{\bullet j}}\} \ge T_j$, $1 \le j \le k^*$.

Additionally, we have to take into account the events $\{U_\ell \le u_\ell\}$, $1 \le \ell \le m$. Using the independence of $U_1, \dots, U_n$, the corresponding probability is given by

$$\begin{aligned} p_{\mathrm{id}} &= P(U_\ell \le u_\ell, 1 \le \ell \le m, T_{j-1} \le U_{d_{\bullet j-1}+1} \le \dots \le U_{d_{\bullet j}} < T_j, \\ &\qquad \min\{U_{m+R_{\bullet j-1}+1}, \dots, U_{m+R_{\bullet j}}\} \ge T_j, 1 \le j \le k^*) \\ &= \prod_{j=1}^{k^*} P\big(U_\ell \le u_\ell, 1 \le \ell \le m, T_{j-1} \le U_{d_{\bullet j-1}+1} \le \dots \le U_{d_{\bullet j}} < T_j\big) \\ &\qquad \times \prod_{j=1}^{k^*} (1 - T_j)^{R_j}. \end{aligned}$$

Assuming a particular outcome represented by a permutation π of the indices $\{1, \dots, n\}$, we have $p_\pi = p_{\mathrm{id}}$ for any $\pi \in \mathfrak{S}_n$. Denoting by the event A_π the assignment of the permutation π to the unit indices, the probability in (4.3) can be written as

$$
\begin{aligned}
q &= \sum_{\pi\in\mathfrak{S}_n} P(U_{\ell:m:n}^{\mathscr{R},T} \le u_\ell, 1\le \ell\le m, D_j = d_j, 1\le j\le k^*|A_\pi)P(A_\pi)\\
&= \sum_{\pi\in\mathfrak{S}_n}\Big[\prod_{j=1}^{k^*} P\big(U_{\pi(\ell)}\le u_\ell, 1\le\ell\le m,\\
&\qquad T_{j-1}\le U_{\pi(d_{\bullet j-1}+1)}\le\cdots\le U_{\pi(d_{\bullet j})} < T_j\big)\Big]\Big(\prod_{j=1}^{k^*}(1-T_j)^{R_j}\Big)P(A_\pi)\\
&= \sum_{\pi\in\mathfrak{S}_n}\Big[\prod_{j=1}^{k^*} P\big(T_{j-1}\le U_{\ell:d_j}\le \min(u_{d_{\bullet j-1}+\ell}, T_j), 1\le \ell\le d_j\big)(1-T_j)^{R_j}\Big]\\
&\qquad \times P(A_\pi).
\end{aligned}
$$

The probability to choose a specific permutation π leading to the previous outcome is then given by

$$
P(A_\pi) = \frac{1}{n!}\prod_{j=1}^{k^*}\binom{n-d_{\bullet j-1}-R_{\bullet j-1}}{d_j}, \quad \pi\in\mathfrak{S}_n,
$$

which is also independent of π. Therefore,

$$
\begin{aligned}
q = \prod_{j=1}^{k^*}\Bigg[&\binom{n-d_{\bullet j-1}-R_{\bullet j-1}}{d_j}(1-T_j)^{R_j}\\
&\times P(T_{j-1}\le U_{\ell:d_j}\le\min(u_{d_{\bullet j-1}+\ell}, T_j), 1\le\ell\le d_j)\Bigg], \quad (4.4)
\end{aligned}
$$

where $U_{\ell:d_j}$, $1\le\ell\le d_j$, are uniform order statistics.

The probability $P(T_{j-1}\le U_{\ell:d_j}\le\min(u_{d_{\bullet j-1}+\ell}, T_j), 1\le\ell\le d_j)$ depends only on $u_{d_{\bullet j-1}+\ell}$ iff $T_{j-1}\le u_{d_{\bullet j-1}+\ell}\le T_j$ for $1\le\ell\le d_j, 1\le j\le k^*$. Differentiating (4.4) w.r.t. $u_1,\ldots,u_m$ yields the density function

$$
f_{\mathscr{R}}(\mathbf{u}_m) = \prod_{j=1}^{k^*}\binom{n-d_{\bullet j-1}-R_{\bullet j-1}}{d_j}(1-T_j)^{R_j}d_j!, \quad \mathbf{u}_m\in\mathfrak{T},
$$

where $\mathfrak{T} = \{\mathbf{u}_m | T_{j-1}\le u_{d_{\bullet j-1}+1}\le\cdots\le u_{d_{\bullet j}}\le T_j, 1\le j\le k^*\}$. Therefore, we get

$$q = \Big[\prod_{j=1}^{k^*} \binom{n - d_{\bullet j-1} - R_{\bullet j-1}}{d_j} (1-T_j)^{R_j} d_j!\Big] \int_{\mathfrak{T}} d\mathbf{u}_m.$$

In order to establish the conditional distribution $P^{U^{\mathscr{R},T}_{\ell:m:n}, 1\le \ell \le m | (D_1,\ldots,D_k)=\mathbf{d}_k}$, we need the joint probability mass function of $D_1, \ldots, D_k$. Choosing $u_1 = \cdots = u_m = 1$ in (4.4), we have to calculate the probability $P(T_{j-1} \le U_{\ell:d_j} \le T_j, 1 \le \ell \le d_j)$ which can be rewritten as

$$P(T_{j-1} \le U_\ell \le T_j, 1 \le \ell \le d_j) = (T_j - T_{j-1})^{d_j}.$$

Therefore, the probability mass function is given by

$$\begin{aligned} &P(D_1 = d_1, \ldots, D_k = d_k) \\ &\quad = \prod_{j=1}^{k^*} \binom{n - d_{\bullet j-1} - R_{\bullet j-1}}{d_j} (T_j - T_{j-1})^{d_j} (1-T_j)^{R_j}, \mathbf{d}_k \in \mathfrak{D}. \quad (4.5) \end{aligned}$$

From (4.3)–(4.5), we obtain for $(d_1, \ldots, d_k) \in \mathfrak{D}$ with $d_{\bullet k} \ge 1$

$$\begin{aligned} &P(U^{\mathscr{R},T}_{\ell:m:n} \le u_\ell, 1 \le \ell \le m | D_j = d_j, 1 \le j \le k) \\ &= \prod_{j=1}^{k^*} \frac{P(T_{j-1} \le U_{\ell:d_j} \le \min(u_{d_{\bullet j-1}+\ell}, T_j), 1 \le \ell \le d_j)}{(T_j - T_{j-1})^{d_j}} \\ &= \prod_{j=1}^{k^*} P(T_{j-1} \le U^{(j)}_{\ell:d_j} \le \min(u_{d_{\bullet j-1}+\ell}, T_j), 1 \le \ell \le d_j), \end{aligned}$$

where $U^{(j)}_{1:d_j}, \ldots, U^{(j)}_{d_j:d_j}$ are order statistics from a $\mathsf{U}(T_{j-1}, T_j)$-distribution, $1 \le j \le k^*$. This proves the theorem. □

Combining the preceding result with Theorem 4.1.1, we get directly the following theorem due to Iliopoulos and Balakrishnan [469].

Theorem 4.1.3. Let $\mathbf{d}_k \in \mathfrak{D}$ with $d_{\bullet k} \ge 1$ and let $X_1, \ldots, X_n$ be IID random variables with cumulative distribution function F. Let $-\infty = T_0 < T_1 < \cdots < T_k < \infty$. Given $(D_1, \ldots, D_k) = \mathbf{d}_k$, we have $M = d_{\bullet k}$ and the progressively Type-I censored order statistics $X^{\mathscr{R},T}_{1:M:n}, \ldots, X^{\mathscr{R},T}_{M:M:n}$ to be block independent, i.e., $X^{\mathscr{R},T}_{d_{\bullet j-1}+1:d_{\bullet k}:n}, \ldots, X^{\mathscr{R},T}_{d_{\bullet j}:d_{\bullet k}:n}$, $1 \le j \le k^*$, are mutually independent with, for $d_j \ge 1$,

$$(X^{\mathscr{R},T}_{d_{\bullet j-1}+1:d_{\bullet k}:n}, \ldots, X^{\mathscr{R},T}_{d_{\bullet j}:d_{\bullet k}:n}) \stackrel{\mathsf{d}}{=} (X^{(j)}_{1:d_j}, \ldots, X^{(j)}_{d_j:d_j}),$$

where $X^{(j)}_{1:d_j}, \dots, X^{(j)}_{d_j:d_j}$ are order statistics from the doubly truncated cumulative distribution function F with left and right truncation point T_{j-1} and T_j, respectively, $1 \le j \le k^*$.

Remark 4.1.4. Without loss of generality, we can assume that $0 < F(T_1) < \cdots < F(T_k) \le 1$. If this assumption is violated, we can reduce the number of censoring times by combining those T_ℓ with equal value $F(T_\ell)$. From a probabilistic point of view, failures in the interval $[T_{\ell-1}, T_\ell]$ with $F(T_{\ell-1}) = F(T_\ell)$ occur only with zero probability.

From (4.4), we obtain the following representation for the joint density function provided that F has a density function f.

Theorem 4.1.5. Let the effectively applied censoring numbers be defined by

$$R_j = \min(R_j^0, [n - d_{\bullet j} - R_{\bullet j-1}]_+), \quad 1 \le j \le k. \tag{4.6}$$

Then, the joint density function of $X^{\mathscr{R},T}_{1:M:n}, \dots, X^{\mathscr{R},T}_{M:M:n}, D_1, \dots, D_k$ is given by

$$f_{I,\mathscr{R}}(\mathbf{x}_m, \mathbf{d}_k) = \prod_{i=1}^{k} \binom{[n - d_{\bullet i-1} - R_{\bullet i-1}]_+}{d_i} d_i![1 - F(T_i)]^{R_i} \times \Big\{ \prod_{j=1}^{d_i} f(x_{d_{\bullet i-1}+j}) \mathbb{1}_{(T_{i-1},T_i]}(x_{d_{\bullet i-1}+j}) \Big\}, \tag{4.7}$$

for $\mathbf{d}_k \in \mathfrak{D}$ with $m = d_{\bullet k} \ge 1$ and $x_1 \le \cdots \le x_m$.

Remark 4.1.6. The case of common Type-I censoring is included in the above setting by choosing the censoring plan $\mathscr{R}^0 = (0^{*k-1})$. Then, (4.7) simplifies to

$$f_I(\mathbf{x}_m, \mathbf{d}_k) = \frac{n!}{(n-m)!}[1 - F(T_k)]^{n-m} \prod_{i=1}^{m} f(x_i) \mathbb{1}_{(-\infty,T_k]}(x_i)$$

(see, e.g., Wang and He [887]).

Suppressing the dependence on $\mathbf{d}_k = (d_1, \dots, d_k)$, the density function given in (4.7) can be written as

$$f_{I,\mathscr{R}}(\mathbf{x}_m) = C_I \Big[\prod_{i=1}^{m} f(x_i) \Big] \Big[\prod_{j=1}^{k} [1 - F(T_j)]^{R_j} \Big], \tag{4.8}$$

where $R_1, \dots, R_k$ denote the effectively applied censoring numbers. This formula is usually used in the literature and can be found, e.g., in Cohen [267]. Since the censoring times $T_1, \dots, T_k$ are fixed times, the progressively Type-I censored order statistics $X^{\mathscr{R},T}_{1:M:n}, \dots, X^{\mathscr{R},T}_{M:M:n}$ determine the random variables $D_1, \dots, D_k$ uniquely. This justifies (4.8) although it is not the density function of $X^{\mathscr{R},T}_{1:M:n}, \dots, X^{\mathscr{R},T}_{M:M:n}$ from a formal point of view. In particular, (4.8) is used to define the likelihood function in statistical inference based on a progressively Type-I censored sample.

Remark 4.1.7. It is interesting that expression (4.8) is quite similar to the density function of progressively Type-II censored order statistics $X^{\mathscr{R}}_{1:m:n}, \dots, X^{\mathscr{R}}_{m:m:n}$ given in (2.4), where k has to be replaced by m and T_ℓ by x_ℓ, $1 \le \ell \le m$, respectively. Moreover, the normalizing constant has to be chosen suitably. Subsequently, we denote by R_i the effectively applied censoring numbers.

4.2 Number of Observations

The random variables $D_1, \dots, D_k$ play a crucial role in the derivation of the distribution of a progressively Type-I censored sample. Moreover, in many inferential issues, it is important to have at least one observation. In this section, we investigate distributional properties of $D_1, \dots, D_k$. We get the following result which provides the joint distribution of $(D_1, \dots, D_k)$.

Corollary 4.2.1. Let $X^{\mathscr{R},T}_{1:M:n}, \dots, X^{\mathscr{R},T}_{M:M:n}$ be progressively Type-I censored order statistics based on F and censoring scheme $\mathscr{R}$. Then, $(D_1, \dots, D_k)$ has the joint probability mass function

$$\begin{aligned}&P(D_1 = d_1, \dots, D_k = d_k)\\ &= \prod_{i=1}^{k^*} \binom{n - d_{\bullet i-1} - R_{\bullet i-1}}{d_i} [F(T_i) - F(T_{i-1})]^{d_i} [1 - F(T_i)]^{R_i},\end{aligned}$$

where $R_i = R_i^0$, $1 \le i \le k^* - 1$, and $R_{k^*} = n - d_{\bullet k^*} - R_{\bullet k^*-1}$.

Using the effectively applied censoring numbers as defined in (4.6), we get the following alternative representation which can be found in Iliopoulos and Balakrishnan [469]:

$$\begin{aligned}&P(D_1 = d_1, \dots, D_k = d_k)\\ &= \prod_{i=1}^{k} \binom{[n - d_{\bullet i-1} - R_{\bullet i-1}]_+}{d_i} [F(T_i) - F(T_{i-1})]^{d_i} [1 - F(T_i)]^{R_i}.\end{aligned}$$

Here, we use the convention $\binom{0}{j} = 1$ when $j = 0$, and 0 otherwise. Notice that $\mathfrak{D}$ given in (4.2) is the support of $(D_1, \ldots, D_k)$. In particular, we have $P(D_1 = d_1, \ldots, D_k = d_k) = 0$ when $d_j > 0$ for some $k^* < j \leq k$. Let $j \leq k^*$. Then, the marginal distribution of $(D_1, \ldots, D_j)$ is given by

$$\begin{aligned} &P(D_1 = d_1, \ldots, D_j = d_j) \\ &= \prod_{i=1}^{j} \binom{n - d_{\bullet i-1} - R_{\bullet i-1}}{d_i} [F(T_i) - F(T_{i-1})]^{d_i} [1 - F(T_i)]^{R^*_{i,j}}, \end{aligned}$$

where $R^*_{i,j} = R^0_i$, $1 \leq i \leq j-1$ and $R^*_{j,j} = n - d_{\bullet j} - R^0_{\bullet j-1}$. Therefore, the conditional distribution of D_j, given $(D_1, \ldots, D_{j-1}) = \mathbf{d}_{j-1}$, $1 \leq j \leq k^*$, is given by

$$\begin{aligned} &P(D_j = d | D_1 = d_1, \ldots, D_{j-1} = d_{j-1}) \\ &= \binom{n_j}{d} \Big[\frac{F(T_j) - F(T_{j-1})}{1 - F(T_{j-1})} \Big]^d \Big[\frac{1 - F(T_j)}{1 - F(T_{j-1})} \Big]^{n_j - d}, \quad d \in \{0, \ldots, n_j\}, \end{aligned} \tag{4.9}$$

where $n_j = n - d_{\bullet j-1} - R^0_{\bullet j-1}$. The expression in (4.9) is readily seen to be the probability mass function of a binomial distribution with probability of success $\frac{F(T_j) - F(T_{j-1})}{1 - F(T_{j-1})}$. Obviously this conditional probability depends only on d and $d_{\bullet j-1}$. This motivates the following theorem which follows directly from (4.9). The particular case of an exponential distribution is studied in Balakrishnan et al. [150].

Theorem 4.2.2. Let $D_{\bullet 0} = 0$ and $p_j = \frac{F(T_j) - F(T_{j-1})}{1 - F(T_{j-1})}$. $1 \leq j \leq k$. Then, $(D_{\bullet j})_{0 \leq j \leq k}$ forms a Markov chain with transition probabilities

$$P(D_{\bullet j} = s | D_{\bullet j-1} = t) = \binom{n - t - R_{\bullet j-1}}{s - t} p_j^{s-t} (1 - p_j)^{n - s - R_{\bullet j-1}}$$

with $s = t, \ldots, n - R_{\bullet j-1}$ if $n - t - R_{\bullet j-1} \geq 0$, and

$$P(D_{\bullet j} = t | D_{\bullet j-1} = t) = 1 \quad \text{if } n - t - R_{\bullet j-1} < 0.$$

For inferential issues, it is important to have information about the total number of observations. From (4.1), we know that, for given $\mathscr{R}^0$ and censoring times $T_1 < \cdots < T_k$, the probability P_O of no observation is given by

$$P_O = P(D_{\bullet k} = 0) = [1 - F(T_k)]^{n - R^0_{\bullet k-1}} \prod_{i=1}^{k-1} [1 - F(T_i)]^{R^0_i}$$

(see also Corollary 4.2.1). Hence, we get for a fixed censoring plan and increasing sample size

$$\lim_{n\to\infty} P(D_{\bullet k} = 0) = 0 \text{ or } \lim_{n\to\infty} P(D_{\bullet k} \geq 1) = \lim_{n\to\infty}(1 - P(D_{\bullet k} = 0)) = 1$$

provided that $F(T_k) > 0$. This is a natural assumption which should hold. This shows that, in large samples, we can expect to observe failures with high probability. But, in the small sample case, the problem of no observation is still present unless the test duration is chosen extremely large:

$$\lim_{T_k\to\infty} P(D_{\bullet k} \geq 1) = \lim_{T_k\to\infty}(1 - P(D_{\bullet k} = 0)) = 1.$$

Normally, this is definitely not desirable from a practical viewpoint, and consequently inference has to be developed conditionally on the event $D_{\bullet k} \geq 1$ or, equivalently, on $(D_1, \ldots, D_k) \neq (0^{*k})$.

Chapter 5
Progressive Hybrid Censoring: Distributions and Properties

The models of progressive hybrid censoring have been introduced in Sect. 1.1.3. In the following, we will present some details on the distributions as well as related properties. A short review on this topic has been provided recently by Balakrishnan and Kundu [108, Sect. 7].

5.1 Type-I Progressive Hybrid Censoring

In Type-I progressive hybrid censoring, the lifetime experiment is stopped when either m complete failures have been observed or when the threshold time T has been reached. Let $D = \sum_{i=1}^{m} \mathbb{1}_{(-\infty,T]}(X_{i:m:n})$ denote the total number of observed failures. As in Cramer and Balakrishnan [292], we perceive the data with possibly less than m observed failure times as a sample of size m by adding the censoring time in the required number. For a progressively Type-II censored sample $X_{1:m:n}, \dots, X_{m:m:n}$ with censoring scheme $\mathscr{R}$, Type-I progressively hybrid censored order statistics $X_1^{(\mathsf{I})}, \dots, X_m^{(\mathsf{I})}$ are defined via

$$X_j^{(\mathsf{I})} = \min(X_{j:m:n}, T), \quad 1 \le j \le m. \tag{5.1}$$

From this construction, it is evident that the sample may include both observed failure times and censoring times. Conditionally on $D = d$, $d \in \{0, \dots, m\}$, we have

$$X_1^{(\mathsf{I})}, \dots, X_m^{(\mathsf{I})} | (D = d) \stackrel{\mathsf{d}}{=} X_{1:m:n}, \dots, X_{d:m:n}, T^{*m-d}.$$

For $d = 0$, the experiment has been terminated before observing the first failure, and the sample is given by T^{*m}.

The probability mass function of D is important in the following analysis. It has been presented in Lemma 2.5.4. Cramer and Balakrishnan [292] calculated the probabilities $P(X_j^{(\mathsf{I})} \le t_j, 1 \le j \le m, D = d)$. First,

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_5,

$$P(X_j^{(\mathrm{I})} \leq t_j, 1 \leq j \leq m, D = 0) = (1 - F(T))^n \mathbb{1}_{[T,\infty)}(t_1),$$
$$P(X_j^{(\mathrm{I})} \leq t_j, 1 \leq j \leq m, D = m) = F_{1,\ldots,m:m:n}(\mathbf{t}_{m-1}, t_m \wedge T).$$

For $d \in \{1, \ldots, m-1\}$, they found

$$\begin{aligned} P(X_j^{(\mathrm{I})} &\leq t_j, 1 \leq j \leq m, D = d) \\ &= \mathbb{1}_{[T,\infty)}(\min_{d+1 \leq j \leq m} t_j) C_d (1 - F(T))^{\gamma_{d+1}} F_{1,\ldots,d:d:n-\gamma_{d+1}}^{\mathscr{R}_{\rhd d}}(\mathbf{t}_{d-1}, t_d \wedge T), \end{aligned}$$

where $\mathscr{R}_{\rhd d} = (R_1, \ldots, R_d)$ is a right truncated censoring scheme and C_d is given after (5.2). Introducing the notation $F_{1,\ldots,0:0:0} \equiv 1$, this expression is valid for $d = 0$ and for $d = m$ with $\gamma_{m+1} = 0$ and $\min_{m+1 \leq j \leq m} t_j = \infty$ as well. This yields the result

$$P(X_j^{(\mathrm{I})} \leq t_j, 1 \leq j \leq m) = \sum_{d=0}^{m} P(X_j^{(\mathrm{I})} \leq t_j, 1 \leq j \leq m, D = d)$$
$$= \sum_{d=0}^{m} \mathbb{1}_{[T,\infty)}(\min_{d+1 \leq j \leq m} t_j) C_d (1 - F(T))^{\gamma_{d+1}} F_{1,\ldots,d:d:n-\gamma_{d+1}}^{\mathscr{R}_{\rhd d}}(\mathbf{t}_{d-1}, t_d \wedge T), \quad (5.2)$$

where $C_d = \prod_{j=1}^{d} \gamma_j / (\gamma_j - \gamma_{d+1})$, $1 \leq d \leq m$, $C_0 = 1$.

This joint cumulative distribution function is not absolutely continuous w.r.t. the Lebesgue measure. As for progressively Type-I/II censored order statistics, Cramer and Balakrishnan [292] established a quantile representation similar to those in Theorems 4.1.1 and 2.1.1.

Theorem 5.1.1. Let $X_j^{(\mathrm{I})}$, $1 \leq j \leq m$, be Type-I progressively hybrid censored order statistics from a continuous cumulative distribution function F with time threshold T. Then,

$$X_j^{(\mathrm{I})}, \quad 1 \leq j \leq m \quad \stackrel{\mathrm{d}}{=} \quad F^{\leftarrow}\big(U_j^{(\mathrm{I})}\big), \quad 1 \leq j \leq m,$$

where $U_j^{(\mathrm{I})}$, $1 \leq j \leq m$, are Type-I progressively hybrid censored order statistics from a uniform distribution with time threshold $F(T)$.

Given an absolutely continuous cumulative distribution function F, the conditional cumulative distribution function $F^{X_j^{(\mathrm{I})}, 1 \leq j \leq m | D = d}$ has a density function w.r.t. the product measure $\lambda^d \otimes \bigotimes_{j=1}^{m-d} \varepsilon_T$, where λ^d denotes the d-dimensional Lebesgue measure and ε_T is a one-point distribution in T. The corresponding conditional density function is given by

$$\begin{aligned} f^{X_j^{(\mathrm{I})}, 1 \leq j \leq m | D = d}&(\mathbf{t}_d, T^{*(m-d)}) \\ &= \frac{C_d}{P(D = d)} (1 - F(T))^{\gamma_{d+1}} f_{1,\ldots,d:d:n-\gamma_{d+1}}^{\mathscr{R}_d}(\mathbf{t}_d), \quad t_1 \leq \cdots \leq t_d \leq T \end{aligned}$$

(see Childs et al. [260]). From (2.41), Cramer and Balakrishnan [292] derived for $t_1 \le \cdots \le t_d \le T$ and $1 \le d \le m-1$, the identity

$$F^{X_j^{(\mathrm{I})},1\le j\le m|D=d}(\mathbf{t}_d, T^{*(m-d)}) = F^{\mathscr{R}}_{1,\dots,d:m:n}(\mathbf{t}_d|X_{d+1:m:n} = T),$$
$$t_1 \le \cdots \le t_d \le T.$$

Interpreting $F^{\mathscr{R}}_{1,\dots,m:m:n}(\cdot|X_{m+1:m:n} = T)$ as the truncated cumulative distribution function $F^{\mathscr{R}}_{1,\dots,m:m:n}/F_{m:m:n}(T)$, this relation is also true for $D = m$. Hence, (5.2) reads alternatively as

$$P(X_j^{(\mathrm{I})} \le t_j, 1 \le j \le m)$$
$$= \sum_{d=0}^{m} P(D=d)\mathbb{1}_{[T,\infty)}(\min_{d+1\le j\le m} t_j)F^{\mathscr{R}}_{1,\dots,d:m:n}(\mathbf{t}_d|X_{d+1:m:n} = T).$$

Conditional on $D = d$, this yields the density functions

$$f^{X_j^{(\mathrm{I})},1\le j\le d|D=d} = f^{\mathscr{R}}_{1,\dots,d:m:n}(\cdot|X_{d+1:m:n} = T), \quad 1 \le d \le m-1, \tag{5.3}$$
$$f^{X_j^{(\mathrm{I})},1\le j\le m|D=m} = \frac{f^{\mathscr{R}}_{1,\dots,d:m:n}}{F_{m:m:n}(T)}.$$

Therefore, the conditional density function of Type-I progressively hybrid censored order statistics, given $D = d$, can be interpreted as the conditional density function of the first d progressively Type-II censored order statistics, given $X_{d+1:m:n} = T$. These expressions will be used to derive the conditional density function of spacings for an exponential distribution.

The one-dimensional marginal cumulative distribution functions can be obtained directly from (5.1), i.e., from the relation $X_j^{(\mathrm{I})} = X_{j:m:n} \wedge T$:

$$F^{X_j^{(\mathrm{I})}}(t) = \mathbb{1}_{[T,\infty)}(t)(1 - F^{X_{j:m:n}}(T)) + F^{X_{j:m:n}}(t \wedge T), \quad t \in \mathbb{R}.$$

Obviously, the cumulative distribution function has a jump of height $1 - F^{X_{j:m:n}}(T)$ at $t = T$.

5.1.1 Spacings for Exponential Distribution

Spacings play an important role in the analysis of exponential progressively Type-II censored order statistics. Surprisingly, they have not been studied for Type-I progressively hybrid censored exponential data until Cramer and Balakrishnan [292] came up with the following results based on (5.3). Let $Z_1^{(\mathrm{I})}, \dots, Z_m^{(\mathrm{I})}$ be a Type-I progressively hybrid censored sample from an $\mathsf{Exp}(\mu, \vartheta)$ distribution, and suppose

$0 < d < m$. The normalized spacings of the first d random variables $Z_1^{(\mathrm{I})}, \ldots, Z_d^{(\mathrm{I})}$ are defined as

$$W_j^{(\mathrm{I})} = \gamma_j (Z_j^{(\mathrm{I})} - Z_{j-1}^{(\mathrm{I})}), \quad 1 \le j \le d,$$

where $Z_0^{(\mathrm{I})} = \mu$. Obviously, conditional on $D = d$, $Z_j^{(I)} = Z_{j:m:n}$ and $W_j^{(\mathrm{I})}$, $1 \le j \le d$, denote the normalized spacings of exponential progressively Type-II censored order statistics. Then, using the identity

$$\sum_{j=1}^{d} (R_j + 1)(z_j - \mu) + \gamma_{d+1}(T - \mu) = \sum_{j=1}^{d} \gamma_j (z_j - z_{j-1}) + \gamma_{d+1}(T - z_d) \quad (5.4)$$

and the representation of the conditional density function

$$f^{Z_j^{(\mathrm{I})}, 1 \le j \le d | D=d}(\mathbf{z}_d)$$
$$= \frac{\prod_{j=1}^{d+1} \gamma_j}{\vartheta^{d+1} f_{d+1:m:n}(T)} \exp\Big\{ -\frac{1}{\vartheta} \Big[\sum_{j=1}^{d} \gamma_j (z_j - z_{j-1}) + \gamma_{d+1}(T - z_d) \Big] \Big\},$$
$$\mu \le z_1 \le \cdots \le z_d \le T,$$

the density transformation formula, and the identity $Z_d^{(\mathrm{I})} = \mu + \sum_{j=1}^{d} W_j^{(\mathrm{I})} / \gamma_j$, Cramer and Balakrishnan [292] established the (conditional) joint density function of the spacings as

$$f^{W_j^{(\mathrm{I})}, 1 \le j \le d | D=d}(\mathbf{w}_d)$$
$$= \frac{\gamma_{d+1} \mathrm{e}^{-\gamma_{d+1}(T-\mu)/\vartheta}}{\vartheta f_{d+1:m:n}(T)} \Big[\prod_{j=1}^{d} \frac{1}{\vartheta} \exp\Big\{ -\Big(1 - \frac{\gamma_{d+1}}{\gamma_j}\Big) \frac{w_j}{\vartheta} \Big\} \Big], \quad \mathbf{w}_d \in \mathcal{W}_d(T), \quad (5.5)$$

with support

$$\mathcal{W}_d(T) = \Big\{ \mathbf{w}_d | w_j \ge 0, 1 \le j \le d, \sum_{j=1}^{d} \frac{w_j}{\gamma_j} \le T - \mu \Big\}.$$

Although the density function in (5.5) has a product form similar to that in case of the progressive Type-II censoring in Theorem 2.3.2, the spacings $W_1^{(\mathrm{I})}, \ldots, W_j^{(\mathrm{I})}$ are not (conditionally) independent. This is due to the fact that, for a finite T, the support of the density function is restricted to a simplex by the condition $\sum_{j=1}^{d} \frac{w_j}{\gamma_j} \le T - \mu$. In particular, this result yields the identity

$$\int_{\mathcal{W}_d(T)} \prod_{j=1}^{d} \frac{1}{\vartheta} \exp\Big\{ -\Big(1 - \frac{\gamma_{d+1}}{\gamma_j}\Big) \frac{w_j}{\vartheta} \Big\} d\mathbf{w}_d = \frac{\vartheta}{\gamma_{d+1}} f_{d+1:m:n}(T) \mathrm{e}^{\gamma_{d+1}(T-\mu)/\vartheta}.$$

For $D = m$, analogous results hold. From the joint density function of $Z_1^{(\mathrm{I})}, \ldots, Z_m^{(\mathrm{I})}$, given $D = m$,

$$f^{Z_j^{(\mathrm{I})}, 1 \le j \le m | D=m}(\mathbf{z}_m) = \frac{\prod_{j=1}^m \gamma_j}{\vartheta^m F_{m:m:n}(T)} \exp\Big\{ -\frac{1}{\vartheta} \sum_{j=1}^m \gamma_j (z_j - z_{j-1}) \Big\},$$

$$\mu \le z_1 \le \cdots \le z_m \le T,$$

we get, as before, by the density transformation formula, the conditional density function as

$$f^{W_j^{(\mathrm{I})}, 1 \le j \le m | D=m}(\mathbf{w}_m) = \frac{1}{F_{m:m:n}(T)} \Big[\prod_{j=1}^m \frac{1}{\vartheta} \exp\Big\{ -\frac{w_j}{\vartheta} \Big\} \Big], \ \mathbf{w}_m \in \mathcal{W}_m(T).$$

Finally, these results can be used to establish representations for the marginal density function $f^{W_j^{(\mathrm{I})} | D=d}$ and the corresponding cumulative distribution function of the spacings. A proof is given in Cramer and Balakrishnan [292].

Theorem 5.1.2. For $1 \le k \le d \le m$, let $f_{d:m:n}^{(k)}$ denote the marginal density function and $F_{d:m:n}^{(k)}$ denote the marginal cumulative distribution function of the dth exponential progressively Type-II censored order statistic with censoring scheme $(R_1, \ldots, R_{k-2}, R_{k-1} + R_k + 1, R_{k+1}, \ldots, R_m)$ (or, equivalently, with parameters $\gamma_1, \ldots, \gamma_{k-1}, \gamma_{k+1}, \ldots, \gamma_m$). Then, for $1 \le d \le k < m$, the spacing $W_k^{(\mathrm{I})}$ has a conditional density function given by

$$f^{W_k^{(\mathrm{I})} | D=d}(w) = \frac{f_{d:m:n}^{(k)}(T - w/\gamma_k) \mathrm{e}^{-w/\vartheta}}{\vartheta f_{d+1:m:n}(T)}, \quad 0 \le w \le \gamma_k (T - \mu).$$

The corresponding cumulative distribution function is given by

$$F^{W_k^{(\mathrm{I})} | D=d}(t) = 1 - \frac{f_{d+1:m:n}(T - t/\gamma_k) \mathrm{e}^{-t/\vartheta}}{f_{d+1:m:n}(T)}, \quad 0 \le t \le \gamma_k (T - \mu).$$

For $D = m$, the spacing $W_k^{(\mathrm{I})}$ has a conditional density function

$$f^{W_k^{(\mathrm{I})} | D=m}(w) = \frac{F_{m-1:m-1:n}^{(k)}(T - w/\gamma_k) \mathrm{e}^{-w/\vartheta}}{\vartheta F_{m:m:n}(T)}, \quad 0 \le w \le \gamma_k (T - \mu), \quad (5.6)$$

and the corresponding distribution function is

$$F^{W_k^{(\mathrm{I})} | D=m}(t) = 1 - \frac{F_{m:m:n}(T - t/\gamma_k) \mathrm{e}^{-t/\vartheta}}{F_{m:m:n}(T)}, \quad 0 \le t \le \gamma_k (T - \mu).$$

Remark 5.1.3. As pointed out by Cramer and Balakrishnan [292], the structure does not simplify for usual order statistics. However, the density functions given in Theorem 5.1.2 can be interpreted as density functions of progressively Type-II censored order statistics with a one-step censoring scheme with the progressive censoring taking place at the $(k-1)$th censoring step. For cumulative distribution functions, they obtained, for $1 \le d \le n-1$,

$$F^{W_k^{(\mathrm{I})}|D=d}(t) = 1 - \frac{f_{d+1:n}(T - t/(n-k+1))\mathrm{e}^{-t/\vartheta}}{f_{d+1:n}(T)},$$
$$0 \le t \le (n-k+1)(T-\mu), 1 \le k \le d-1,$$

and, for $d = n$,

$$\begin{aligned} F^{W_k^{(\mathrm{I})}|D=n}(t) &= 1 - \frac{F_{n:n}(T - t/(n-k+1))\mathrm{e}^{-t/\vartheta}}{F_{n:n}(T)} \\ &= 1 - \Big(\frac{1-\mathrm{e}^{-(T-\mu-t/(n-k+1))/\vartheta}}{1-\mathrm{e}^{-(T-\mu)/\vartheta}}\Big)^n \mathrm{e}^{-t/\vartheta}, \\ &\qquad 0 \le t \le (n-k+1)(T-\mu). \end{aligned}$$

5.1.2 Distributions of Total Time on Test and Related Statistics

The distribution of the spacings can be used to determine the distributions of the total time on test statistic and a modified total time on test statistic. First, consider an Exp(ϑ)-distribution and define $Z_0^{(\mathrm{I})} = 0$. Given $D = d$, the total time on test statistic is

$$S_d = \sum_{j=1}^{d}(R_j+1)Z_j^{(\mathrm{I})} + \gamma_{d+1}T = \sum_{j=1}^{d}\Big(1 - \frac{\gamma_{d+1}}{\gamma_j}\Big)W_j^{(\mathrm{I})} + \gamma_{d+1}T, \quad 1 \le d \le m, \tag{5.7}$$

with support $[\gamma_{d+1}T, nT]$. As before, the cases $1 \le d \le m-1$ and $d = m$ have to be handled separately. Cramer and Balakrishnan [292] found for $1 \le d \le m-1$

$$f^{W_j^{(\mathrm{I})},1\le j\le d-1,S_d|D=d}(\mathbf{w}_{d-1}, s) = \frac{\gamma_{d+1}\gamma_d}{(\gamma_d - \gamma_{d+1})\vartheta^{d+1} f_{d+1:m:n}(T)} \exp\Big\{-\frac{s}{\vartheta}\Big\}, \tag{5.8}$$

and for $d = m$,

$$f^{W_j^{(\mathrm{I})},1\le j\le m-1,S_m|D=m}(\mathbf{w}_{m-1}, s) = \frac{1}{\vartheta^m F_{m:m:n}(T)} \exp\Big\{-\frac{s}{\vartheta}\Big\}. \tag{5.9}$$

Let $d \in \{2, \ldots, m\}$. The support $\mathcal{M}_{d-1}^{(s)}$ of the density functions in (5.8) and (5.9) can be written as

$$\mathcal{M}_{d-1}^{(s)} = \mathcal{S}_{d-1}(\alpha_1/s^*, \ldots, \alpha_{d-1}/s^*) \cap \mathcal{H}_{d-1}^{(s)}$$

with the simplex $\mathcal{S}_{d-1}(\alpha_1/s^*, \ldots, \alpha_{d-1}/s^*)$ and the half-space

$$\mathcal{H}_{d-1}^{(s)} = \Big\{\mathbf{w}_{d-1} | s - \gamma_d T \le \sum_{j=1}^{d-1} \beta_j w_j\Big\}.$$

Using results of Gerber [395] and Cho and Cho [263], an explicit formula for the volume can be established. As shown by Cramer and Balakrishnan [292], the volume of $\mathcal{M}_{d-1}^{(s)}$ can be written as the univariate B-spline B_{d-1} of degree $d-1$ with knots $\gamma_{d+1}T < \cdots < \gamma_1 T$ defined as

$$B_{d-1}(s|\gamma_{d+1}T, \ldots, \gamma_1 T) = \frac{(-1)^d}{T^d} d \cdot \sum_{i=1}^{d+1} a_{i,d+1}[\gamma_i T - s]_+^{d-1},$$
$$s \in [\gamma_{d+1}T, \gamma_1 T]. \quad (5.10)$$

Further details on B-splines can be found in, e.g., de Boor [330]. Their connection to statistics and order statistics is extensively discussed in Dahmen and Micchelli [319]. Now, the preceding relations yield the identity

$$\text{Volume}(\mathcal{M}_{d-1}^{(s)}) = \Big(\frac{T^d(\gamma_d - \gamma_{d+1})}{d!} \prod_{j=1}^{d-1} \gamma_j\Big) B_{d-1}(s|\gamma_{d+1}T, \ldots, \gamma_1 T),$$
$$s \in [\gamma_{d+1}T, \gamma_1 T].$$

Notice that the support of the B-spline is given by $[\gamma_{d+1}T, \gamma_1 T]$ because $B_{d-1}(s|\gamma_{d+1}T, \ldots, \gamma_1 T) = 0$ for $s \notin [\gamma_{d+1}T, \gamma_1 T]$. Finally, this yields the following representations of the density functions.

Theorem 5.1.4. Let $1 \le d \le m-1$. The density function $f^{S_d|D=d}$ is given by

$$f^{S_d|D=d}(s) = \frac{T^d \prod_{j=1}^{d+1} \gamma_j}{d! \vartheta^{d+1} f_{d+1:m:n}(T)} B_{d-1}(s|\gamma_{d+1}T, \ldots, \gamma_1 T) \mathrm{e}^{-s/\vartheta},$$
$$0 \le s \le nT. \quad (5.11)$$

For $d = m$, the density function $f^{S_m|D=m}$ is given by

$$f^{S_m|D=m}(s) = \frac{T^m \prod_{j=1}^{m} \gamma_j}{m! \vartheta^m F_{m:m:n}(T)} B_{m-1}(s|\gamma_{m+1}T, \ldots, \gamma_1 T) \mathrm{e}^{-s/\vartheta}, \; 0 \le s \le nT.$$

Similarly, Cramer and Balakrishnan [292] derived the density function of a modified total time on test statistic defined by

$$V_d = \sum_{j=2}^{d} \Big(1 - \frac{\gamma_{d+1}}{\gamma_j}\Big) W_j^{(\mathrm{I})} + \gamma_{d+1}\Big(T - \mu - \frac{W_1^{(\mathrm{I})}}{n}\Big), \tag{5.12}$$

where $d \in \{1, \ldots, m\}$. For $1 < d < m$, the joint density function of $W_1^{(\mathrm{I})}$ and V_d, conditionally on $D = d$, is given by

$$f^{W_j^{(\mathrm{I})}, 1 \le j \le d-1, V_d | D=d}(\mathbf{w}_{d-1}, v) = \frac{\gamma_{d+1}\gamma_d}{(\gamma_d - \gamma_{d+1})\vartheta^{d+1} f_{d+1:m:n}(T)} \exp\Big\{-\frac{w_1 + v}{\vartheta}\Big\}. \tag{5.13}$$

Integrating out the variables $w_2, \ldots, w_{d-1}$ in (5.13) results in the conditional bivariate density $f^{W_1^{(\mathrm{I})}, V_d | D=d}$ as

$$\begin{aligned} f^{W_1^{(\mathrm{I})}, V_d | D=d}(w, v) &= \frac{(T - \mu - \frac{w}{n})^{d-1} \prod_{j=2}^{d+1} \gamma_j}{(d-1)! \vartheta^{d+1} f_{d+1:m:n}(T)} \\ &\times B_{d-2}\Big(v | \gamma_{d+1}\Big(T - \mu - \frac{w}{n}\Big), \ldots, \gamma_2\Big(T - \mu - \frac{w}{n}\Big)\Big) \exp\Big\{-\frac{w + v}{\vartheta}\Big\}, \\ &\qquad 0 \le w \le n(T - \mu), 0 \le v \le \gamma_2(T - \mu). \end{aligned} \tag{5.14}$$

The marginal density function $f^{W_1^{(\mathrm{I})} | D=d}$ is given in (5.6). The density function $f^{V_d | D=d}$ results by integrating out w. This integration can be carried out explicitly. It reduces to the calculation of the integrals

$$\int_0^{n(T-\mu)} \Big[\gamma_i\Big(T - \mu - \frac{w}{n}\Big) - v\Big]_+^{d-2} \exp\Big\{-\frac{w}{\vartheta}\Big\} dw, \quad i = 2, \ldots, d+1.$$

Notice that (5.14) has the same structure as the expression (5.11) obtained in the scale setting.

5.1.3 Moment Generating Function

The conditional moment generating function has been utilized in the derivation of the distribution of estimators such as the maximum likelihood estimators by various authors. In particular, it leads directly to expressions for moments of the estimators. For these purposes, we establish an expression for the moment generating function of S_d given in (5.7). First, consider the case when $1 \le d \le m - 1$. Then, for appropriately chosen $w \in \mathbb{R}$ such that the expectation exists,

$$E(\mathrm{e}^{wS_d}|D=d) = \frac{\left(\prod_{j=1}^{d}\gamma_j\right)\overline{F}(T)^{\gamma_{d+1}}}{P(D=d)}$$
$$\times \int_0^T \int_0^{x_{d-1}} \cdots \int_0^{x_2} \exp\left\{w\sum_{j=1}^{d}(R_j+1)x_j + w\gamma_{d+1}T\right\}\prod_{j=1}^{d} f(x_j)\overline{F}^{R_j}(x_j)\,d\mathbf{x}_d.$$

Introducing the notation $1/\vartheta^* = 1/\vartheta - w$ and denoting by F_* the cumulative distribution function of an $\mathsf{Exp}(\vartheta^*)$-distribution, we can rewrite this expression as

$$\frac{\prod_{j=1}^{d}\gamma_j\overline{F}(T)^{\gamma_{d+1}}\mathrm{e}^{w\gamma_{d+1}T/\vartheta}}{P(D=d)}\left(\frac{\vartheta^*}{\vartheta}\right)^d\int_0^T\int_0^{x_{d-1}}\cdots\int_0^{x_2}\prod_{j=1}^{d} f_*(x_j)\overline{F}_*^{R_j}(x_j)\,d\mathbf{x}_d.$$

By Lemma 2.4.8, this yields

$$\begin{aligned}E(\mathrm{e}^{wS_d}|D=d) &= \frac{\prod_{j=1}^{d}\gamma_j\overline{F}(T)^{\gamma_{d+1}}\mathrm{e}^{w\gamma_{d+1}T/\vartheta}}{P(D=d)(1-w\vartheta)^d}\sum_{j=1}^{d+1}a_{j,d+1}^{(k)}\overline{F}_*(T)^{\gamma_j-\gamma_{d+1}}\\ &= \frac{\prod_{j=1}^{d}\gamma_j}{P(D=d)(1-w\vartheta)^d}\sum_{j=1}^{d+1}a_{j,d+1}^{(k)}\overline{F}_*(T)^{\gamma_j}\\ &= \frac{F_{d:m:n}((1-w\vartheta)T) - F_{d+1:m:n}((1-w\vartheta)T)}{P(D=d)(1-w\vartheta)^d}.\end{aligned}$$

An analogous result shows that

$$E(\mathrm{e}^{wS_m}|D=m) = \frac{F_{m:m:n}((1-w\vartheta)T)}{P(D=m)(1-w\vartheta)^m}.$$

Remark 5.1.5. Introducing a random variable D_w defined as

$$D_w = \sum_{i=1}^{m}\mathbb{1}_{(-\infty,(1-w\vartheta)T]}(X_{i:m:n}),\quad w<\frac{1}{\vartheta},$$

the conditional moment generating function can be written as

$$E(\mathrm{e}^{wS_d}|D=d) = \frac{P(D_w=d)}{P(D_0=d)}\cdot\frac{1}{(1-w\vartheta)^d},\quad d=1,\ldots,m.$$

Moreover, this expression shows that $w<\frac{1}{\vartheta}$ must hold in order to ensure existence of the moment generating function. In particular, for any $\vartheta>0$, the moment generating function is defined on an interval $(-\infty,\varepsilon)$ with $\varepsilon>0$.

Differentiating the conditional moment generating function w.r.t. w yields, for $1 \le d \le m-1$, the terms

$$\begin{aligned}
&\frac{\partial}{\partial w} E(\mathrm{e}^{wS_d}|D=d) \\
&= d\vartheta \frac{F_{d:m:n}((1-w\vartheta)T) - F_{d+1:m:n}((1-w\vartheta)T)}{P(D=d)(1-w\vartheta)^{d+1}} \\
&\quad - \vartheta T \frac{f_{d:m:n}((1-w\vartheta)T) - f_{d+1:m:n}((1-w\vartheta)T)}{P(D=d)(1-w\vartheta)^{d}} \\
&= \frac{d\vartheta E(\mathrm{e}^{wS_d}|D=d)}{(1-w\vartheta)} - \vartheta T \frac{f_{d:m:n}((1-w\vartheta)T) - f_{d+1:m:n}((1-w\vartheta)T)}{P(D=d)(1-w\vartheta)^{d}}
\end{aligned}$$

and

$$\frac{\partial}{\partial w} E(\mathrm{e}^{wS_m}|D=m) = \frac{m\vartheta E(\mathrm{e}^{wS_m}|D=m)}{(1-w\vartheta)} - \vartheta T \frac{f_{m:m:n}((1-w\vartheta)T)}{P(D=m)(1-w\vartheta)^{m}}.$$

Similarly,

$$\begin{aligned}
&\frac{\partial^2}{\partial w^2} E(\mathrm{e}^{wS_d}|D=d) \\
&= \frac{d(d+1)\vartheta^2 E(\mathrm{e}^{wS_d}|D=d)}{(1-w\vartheta)^2} \\
&\quad - 2d\vartheta^2 T \frac{f_{d:m:n}((1-w\vartheta)T) - f_{d+1:m:n}((1-w\vartheta)T)}{P(D=d)(1-w\vartheta)^{d+1}} \\
&\quad + \vartheta^2 T^2 \frac{f'_{d:m:n}((1-w\vartheta)T) - f'_{d+1:m:n}((1-w\vartheta)T)}{P(D=d)(1-w\vartheta)^{d}}
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2}{\partial w^2} E(\mathrm{e}^{wS_m}|D=m) &= \frac{m(m+1)\vartheta^2 E(\mathrm{e}^{wS_m}|D=m)}{(1-w\vartheta)^2} \\
&\quad - 2\vartheta^2 T \frac{f_{m:m:n}((1-w\vartheta)T)}{P(D=m)(1-w\vartheta)^{m+1}} + \vartheta^2 T^2 \frac{f'_{m:m:n}((1-w\vartheta)T)}{P(D=m)(1-w\vartheta)^{m}}.
\end{aligned}$$

Evaluating these expressions at $w=0$ and using the fact that $E(S_0|D=0)=nT$, we get the mean total time on test as

$$ES_D = \sum_{d=0}^{m} E(S_d | D = d) P(D = d) = \vartheta ED.$$

Hence, we have proven a Wald-type equation for S_D.

Remark 5.1.6. Notice that

$$ED = \sum_{d=1}^{m} F_{d:m:n}(T).$$

Moreover, for S_D^2, we get

$$ES_D^2 = \vartheta^2 E(D(D+1)) - 2\vartheta^2 T \sum_{d=1}^{m} f_{d:m:n}(T).$$

Notice that for Type-I hybrid censoring, $D \sim \mathsf{bin}(n, F(T))$. Hence, in this case, we arrive at

$$ES_D = \vartheta n(1 - \mathrm{e}^{-T/\vartheta})$$

showing that $\frac{1}{n} S_D$ is an unbiased estimator of $\vartheta(1 - \mathrm{e}^{-T/\vartheta})$.

Similarly, we can obtain an expression for the expectation of $\frac{1}{D} S_D$, conditional on $D \geq 1$. In this case, the result is

$$\begin{aligned}
E_\vartheta\Big(\frac{1}{D} S_D \Big| D \geq 1\Big) &= \frac{1}{P(D \geq 1)} \sum_{d=1}^{m} \frac{1}{d} E_\vartheta(S_d | D = d) P(D = d) \\
&= \vartheta - \frac{\vartheta T}{P(D \geq 1)} \Big(\sum_{d=1}^{m-1} \frac{f_{d:m:n}(T) - f_{d+1:m:n}(T)}{d} + \frac{f_{m:m:n}(T)}{m} \Big) \\
&= \vartheta - \frac{\vartheta T}{P(D \geq 1)} \Big(f_{1:m:n}(T) + \sum_{d=2}^{m} \Big[\frac{1}{d} - \frac{1}{d-1} \Big] f_{d:m:n}(T) \Big) \\
&= \vartheta - \frac{nT\mathrm{e}^{-nT/\vartheta}}{(1 - \mathrm{e}^{-T/\vartheta})^n} + \frac{\vartheta T}{(1 - \mathrm{e}^{-T/\vartheta})^n} \sum_{d=2}^{m} \frac{1}{d(d-1)} f_{d:m:n}(T). \qquad (5.15)
\end{aligned}$$

The mean squared error of $\frac{1}{D} S_D$, conditional on $D \geq 1$, is given by

$$\begin{aligned}
\mathsf{MSE}_\vartheta\Big(\frac{1}{D} S_D \Big| D \geq 1\Big) &= \vartheta^2 E_\vartheta\Big(\frac{1}{D} \Big| D \geq 1\Big) \\
&\quad + \frac{nT^2\mathrm{e}^{-nT/\vartheta}}{(1 - \mathrm{e}^{-T/\vartheta})^n} - \frac{\vartheta^2 T^2}{(1 - \mathrm{e}^{-T/\vartheta})^n} \sum_{d=2}^{m} \frac{2d-1}{d^2(d-1)^2} f'_{d:m:n}(T). \qquad (5.16)
\end{aligned}$$

5.2 Type-II Progressive Hybrid Censoring

Childs et al. [260] (see also Kundu and Joarder [561]) proposed an alternative hybrid censoring procedure called Type-II progressive hybrid censoring. Given a (fixed) threshold time T, the life test terminates at $T_2^* = \max\{X_{m:m:n}, T\}$. This approach guarantees that the life test yields at least the observation of m failure times. Given the progressively Type-II censored sample $X_{1:m:n}, \dots, X_{m:m:n}$ with an initially planned censoring scheme $\mathscr{R} = (R_1, \dots, R_m)$, the right censoring at time $X_{m:m:n}$ is not carried out. We just continue to observe the failure times after $X_{m:m:n}$ until we either arrive at T or the maximum in the progressively Type-II censored sample

$$X_{1:m+R_m:n}, \dots, X_{m:m+R_m:n}, X_{m+1:m+R_m:n}, \dots, X_{m+R_m:m+R_m:n} \tag{5.17}$$

is observed. Notice that this sample can be viewed as progressively Type-II censored order statistics with censoring scheme $\mathscr{R}^* = (R_1, \dots, R_{m-1}, 0^{R_m+1})$. In this sense, we define $\gamma_j = \sum_{i=j}^m (R_i + 1)$, $j = 1, \dots, m-1$, $\gamma_j = m + R_m - j + 1$, $j = m, \dots, m + R_m$. As in the case of Type-I progressive hybrid censoring, we introduce the random counter D as

$$D = \sum_{i=1}^{m+R_m} \mathbb{1}_{(-\infty,T]}(X_{i:m+R_m:n}), \tag{5.18}$$

having support $\{0, \dots, m + R_m\}$. In particular, the distribution of D is given by

$$\begin{aligned}
P(D < m) &= P(X_{m:m+R_m:n} > T) = \overline{F}_{m:m+R_m:n}(T) = \overline{F}_{m:m:n}(T),\\
P(D = d) &= \frac{1 - F(T)}{\gamma_{d+1} f(T)} f_{d+1:m+R_m:n}(T), \quad m \le d \le m + R_m - 1,\\
P(D = m + R_m) &= P(X_{m+R_m:m+R_m:n} \le T) = F_{m+R_m:m+R_m:n}(T).
\end{aligned}$$

This yields the Type-II progressively hybrid censored sample

$$X_j^{(\mathsf{II})} = X_{j:m+R_m:n}, \quad 1 \le j \le D^*,$$

with random sample size $D^* = \max\{m, D\}$. Conditional on $D^* = d$ with $d \in \{m, \dots, m + R_m\}$, we have

$$X_1^{(\mathsf{II})}, \dots, X_D^{(\mathsf{II})} | (D^* = d) \stackrel{\mathsf{d}}{=} X_{1:m+R_m:n}, \dots, X_{d:m+R_m:n}.$$

Then, as pointed out in Cramer et al. [315], we get for $\mathbf{t}_m \in \mathbb{R}^m$ and $d \in \{m, \dots, m + R_m\}$

$$P(X_j^{(\mathsf{II})} \le t_j, 1 \le j \le d, D = d)$$
$$= F_{1,\ldots,d:m:n}(\mathbf{t}_{d-1}, t_d \wedge T) - F_{1,\ldots,d+1:m:n}(\mathbf{t}_{d-1}, t_d \wedge T, T)$$

and

$$P(X_j^{(\mathsf{II})} \le t_j, 1 \le j \le m, D < m)$$
$$= F_{1,\ldots,m:m:n}(\mathbf{t}_m) - F_{1,\ldots,m:m:n}(\mathbf{t}_{m-1}, t_m \wedge T).$$

This yields the conditional density functions

$$f^{X_j^{(\mathsf{II})},1\le j\le d|D=d}(\mathbf{t}_d) = f^{\mathscr{R}^*}_{1,\ldots,d:m+R_m:n}(\mathbf{t}_d | X^{\mathscr{R}^*}_{d+1:m+R_m:n} = T),$$
$$d = m, \ldots, m + R_m - 1,$$
$$f^{X_j^{(\mathsf{II})},1\le j\le m+R_m|D=m+R_m}(\mathbf{t}_{m+R_m})$$
$$= \frac{f^{\mathscr{R}^*}_{1,\ldots,m+R_m:m+R_m:n}(\mathbf{t}_{m+R_m})\mathbb{1}_{(-\infty,T]}(t_{m+R_m})}{F_{m+R_m:m+R_m:n}(T)}.$$

For $D < m$, the density function

$$f^{X_j^{(\mathsf{II})},1\le j\le m|D<m}(\mathbf{t}_m) = \frac{f^{\mathscr{R}}_{1,\ldots,m:m:n}(\mathbf{t}_m)\mathbb{1}_{[T,\infty)}(t_m)}{1 - F_{m:m:n}(T)}$$

results, which belongs to a left truncated distribution.

5.2.1 Exponential Distributions

As in the case of Type-I progressive hybrid censoring, Cramer et al. [315] established the distributions of the spacings for a Type-II progressively hybrid censored sample $Z_1^{(\mathsf{II})}, \ldots, Z_d^{(\mathsf{II})}$ from an $\mathsf{Exp}(\mu, \vartheta)$-distribution with $m \le d \le m + R_m$. Let

$$W_j^{(\mathsf{II})} = \gamma_j (Z_j^{(\mathsf{II})} - Z_{j-1}^{(\mathsf{II})}), \quad 1 \le j \le d,$$

be the normalized spacings of the first d random variables $Z_1^{(\mathsf{II})}, \ldots, Z_d^{(\mathsf{II})}$, where $Z_0^{(\mathsf{II})} = \mu$.

Then, proceeding as in Cramer and Balakrishnan [292], Cramer et al. [315] got for $m \le d \le m + R_m - 1$ and $\gamma_{m+R_m+1} \equiv 0$ the (conditional) joint density function of the spacings as

$$f^{W_j^{(\mathrm{II})},1\le j\le d|D=d}(\mathbf{w}_d)$$
$$= \frac{\gamma_{d+1}\mathrm{e}^{-\gamma_{d+1}(T-\mu)/\vartheta}}{\vartheta f_{d+1:m:n}(T)}\Big[\prod_{j=1}^{d}\frac{1}{\vartheta}\exp\Big\{-\Big(1-\frac{\gamma_{d+1}}{\gamma_j}\Big)\frac{w_j}{\vartheta}\Big\}\Big], \quad \mathbf{w}_d\in\mathcal{W}_d^{\le}(T),$$

with support

$$\mathcal{W}_d^{\le}(T)=\Big\{\mathbf{w}_d|w_j\ge 0, 1\le j\le d, \sum_{j=1}^{d}\frac{w_j}{\gamma_j}\le T-\mu\Big\}.$$

For $D=m+R_m$, they found

$$f^{W_j^{(\mathrm{II})},1\le j\le m+R_m|D=m+R_m}(\mathbf{w}_{m+R_m})$$
$$=\frac{1}{F_{m+R_m:m+R_m:n}(T)}\Big[\prod_{j=1}^{m+R_m}\frac{1}{\vartheta}\exp\Big\{-\frac{w_j}{\vartheta}\Big\}\Big], \quad \mathbf{w}_{m+R_m}\in\mathcal{W}_{m+R_m}^{\le}(T).$$

For $D<m$, the representation

$$f^{W_j^{(\mathrm{II})},1\le j\le m|D<m}(\mathbf{w}_m)=\frac{1}{1-F_{m:m:n}(T)}\Big[\prod_{j=1}^{m}\frac{1}{\vartheta}\exp\Big\{-\frac{w_j}{\vartheta}\Big\}\Big],$$
$$\mathbf{w}_m\in\mathcal{W}_m^{>}(T),$$

holds, where

$$\mathcal{W}_m^{>}(T)=\Big\{\mathbf{w}_m|w_j\ge 0, 1\le j\le m, \sum_{j=1}^{m}\frac{w_j}{\gamma_j}>T-\mu\Big\}.$$

From these expressions, the marginal cumulative distribution functions of the spacings result readily. They are given in the following theorem due to Cramer et al. [315].

Theorem 5.2.1. For $t\ge 0$,

$$F^{W_k^{(\mathrm{II})}|D<m}(t)=1-\frac{(1-F_{m:m:n}(T-t/\gamma_k))\mathrm{e}^{-t/\vartheta}}{1-F_{m:m:n}(T)}, \quad 1\le k\le m,$$

and, for $0 \le t \le \gamma_k(T-\mu)$,

$$F^{W_k^{(\mathrm{II})}|D=d}(t) = 1 - \frac{f_{d+1:m+R_m:n}(T - t/\gamma_k)\mathrm{e}^{-t/\vartheta}}{f_{d+1:m+R_m:n}(T)}, 1 \le k \le d,$$

$$m \le d < m + R_m,$$

$$F^{W_k^{(\mathrm{II})}|D=m+R_m}(t) = 1 - \frac{F_{m+R_m:m+R_m:n}(T - t/\gamma_k)\mathrm{e}^{-t/\vartheta}}{F_{m+R_m:m+R_m:n}(T)},$$

$$1 \le k \le m + R_m.$$

It is interesting to note that the spacings have a bounded support except for the case $D < m$. In order to derive the distribution of the total time on test, given $D = d \in \{m, \ldots, m + R_m\}$, the expressions established in Cramer and Balakrishnan [292] can be directly used (see Theorem 5.1.4). For $m \le d \le m + R_m$, the random variable

$$S_d = \sum_{j=1}^{d}(R_j + 1)Z_j^{(\mathrm{II})} + \gamma_{d+1}T = \sum_{j=1}^{d}\Big(1 - \frac{\gamma_{d+1}}{\gamma_j}\Big)W_j^{(\mathrm{II})} + \gamma_{d+1}T$$

has support $[\gamma_{d+1}T, nT]$. Given $D < m$,

$$S_m = \sum_{j=1}^{m}(R_j + 1)Z_j^{(\mathrm{II})} = \sum_{j=1}^{m} W_j^{(\mathrm{II})}.$$

Proceeding as in Cramer and Balakrishnan [292] and using the density transformation formula, Cramer et al. [315] obtained the joint density function

$$f^{W_j^{(\mathrm{II})},1\le j\le m-1,S_m|D<m}(\mathbf{w}_{m-1}, s) = \frac{1}{\vartheta^m(1 - F_{m:m:n}(T))}\exp\Big\{-\frac{s}{\vartheta}\Big\}.$$

Then, the conditional density functions of the total time on test have the following expression.

Theorem 5.2.2. Let $m \le d \le m + R_m - 1$. The density function $f^{S_d|D=d}$ is given by

$$f^{S_d|D=d}(s) = \frac{T^d\prod_{j=1}^{d+1}\gamma_j}{d!\vartheta^{d+1}f_{d+1:m+R_m:n}(T)}B_{d-1}(s|\gamma_{d+1}T, \ldots, \gamma_1 T)\mathrm{e}^{-s/\vartheta},$$

$$0 \le s \le nT.$$

For $d = m + R_m$, we have

$$f^{S_{m+R_m}|D=m+R_m}(s) = \frac{T^{m+R_m}\prod_{j=1}^{m}\gamma_j}{(m+R_m)!\vartheta^{m+R_m}F_{m+R_m:m+R_m:n}(T)} \times B_{m+R_m-1}(s|0,\gamma_{m+R_m}T,\ldots,\gamma_1 T)\mathrm{e}^{-s/\vartheta}, \quad 0 \le s \le nT.$$

Furthermore,

$$f^{S_m|D<m}(s) = \frac{s^{m-1}\mathrm{e}^{-s/\vartheta}}{(m-1)!\vartheta^m(1-F_{m:m:n}(T))} - \frac{T^m\prod_{j=1}^{m}\gamma_j}{m!\vartheta^m(1-F_{m:m:n}(T))}B_{m-1}(s|0,\gamma_m T,\ldots,\gamma_1 T)\mathrm{e}^{-s/\vartheta}, \quad s \ge 0.$$

This results in the following expressions for f^{S_D}.

Theorem 5.2.3. The density function of S_D is given by

$$f^{S_D}(s) = \frac{s^{m-1}\mathrm{e}^{-s/\vartheta}}{(m-1)!\vartheta^m} - \frac{T^m\prod_{j=1}^{m}\gamma_j}{m!\vartheta^m}B_{m-1}(s|0,\gamma_m T,\ldots,\gamma_1 T)\mathrm{e}^{-s/\vartheta} + \sum_{d=m}^{m+R_m}\frac{T^d\prod_{j=1}^{d}\gamma_j}{d!\vartheta^d}B_{d-1}(s|\gamma_{d+1}T,\ldots,\gamma_1 T)\mathrm{e}^{-s/\vartheta}, \quad s \ge 0.$$

Results for the distribution of the modified total time on test V_D defined in (5.12) for Type-II hybrid censored data have also been derived by Cramer et al. [315].

5.3 Generalized Progressive Hybrid Censoring

Both Type-I and Type-II hybrid censoring have some drawbacks as pointed out in Chandrasekar et al. [245]. For instance, Type-I hybrid censoring may result in few failures before the termination time T whereas Type-II hybrid censoring may take a long time to observe the minimum number m of failures. In order to overcome these weaknesses, Chandrasekar et al. [245] proposed two versions of generalized hybrid censoring. Their approach has been adopted by Górny and Cramer [406] to progressively censored data. For brevity, we present only the data situation. For details on distributions and inferential results, we refer to Górny and Cramer [406].

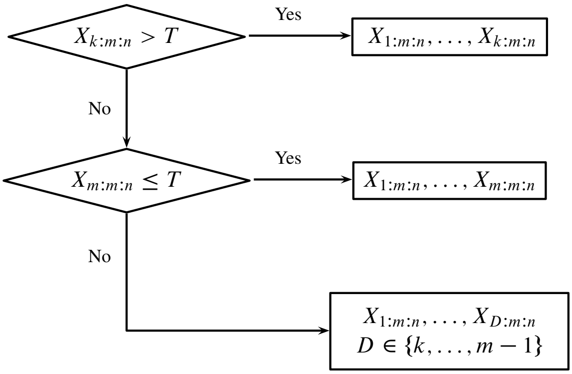

Fig. 5.1 Resulting samples in generalized progressive Type-I hybrid censoring with $1 \le k < m$

Generalized Progressive Type-I Hybrid Censoring

Let $T \in (0, \infty)$ and $k \in \{1, \ldots, m-1\}$. The *generalized progressive Type-I hybrid censoring scheme* aims to stop the life test at $X_{k:m:n}$ provided that the kth failure $X_{k:m:n}$ occurs after time T. If the kth failure is observed before reaching the threshold time T, the experiment will be terminated at $\min\{T, X_{m:m:n}\}$. The corresponding stopping time can be expressed as

$$T_{\mathsf{I}}^* = \min\big\{\max\{T, X_{k:m:n}\}, X_{k:m:n}\big\},\ k < m \text{ and } T \in (0, \infty).$$

Introducing $D = \sum_{j=1}^{m} \mathbb{1}_{(-\infty,T]}(X_{j:m:n})$, the possibly occurring sample situations are depicted in Fig. 5.1. Notice that this procedure ensures at least k observations but the experiment may also exceed the desired maximum experimental time T in case $X_{k:m:n}$ exceeds T. The experimental time is bounded as

$$X_{k:m:n} \le T_{\mathsf{I}}^* \le \max\{T, X_{k:m:n}\}.$$

Generalized Progressive Type-II Hybrid Censoring

Using the progressively censored data given in (5.17) (see also Cramer et al. [315]), a generalization of Type-II progressive hybrid censoring is introduced as follows. Let $m \in \{1, \ldots, n\}$ be fixed and $T_1, T_2 \in (0, \infty)$ with $T_1 < T_2$. Then, the following situations are considered:

(1) Assuming that the mth failure has been observed before T_1, the experiment will be stopped at T_1;
(2) If the mth failure is between the threshold times T_1 and T_2, the experiment is terminated at the mth failure time $X_{m:m:n}$;

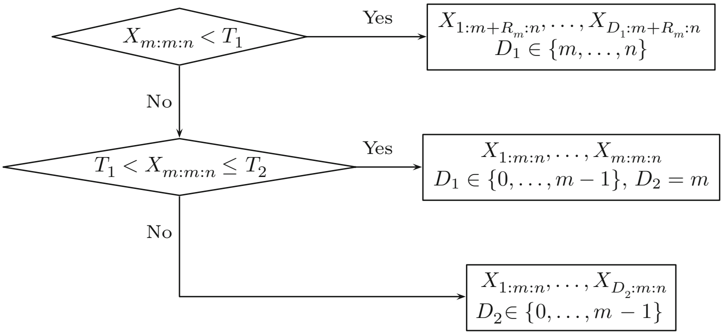

Fig. 5.2 Resulting samples in generalized progressive Type-II hybrid censoring with $T_1 < T_2$

(3) If $X_{m:m:n}$ exceeds T_2, then the experiment is terminated at T_2. This guarantees that the experiment will not last longer than T_2.

This censoring model will be called *generalized progressive Type-II hybrid censoring*.

Notice that in scenario (1), more than m failures may be observed. On the other hand, it may happen that no failure is observed when scenario (3) is present. The stopping time T_{II}^* is defined by

$$T_{\text{II}}^* = \max\big\{\min\{X_{m:m+R_m:n}, T_2\}, T_1\big\}, \text{ with } T_1, T_2 \in (0, \infty) \text{ and } T_1 < T_2.$$

Introducing discrete random variables D_1 and D_2 as

$$D_1 = \sum_{j=1}^{m+R_m} \mathbb{1}_{(-\infty,T_1]}(X_{j:m+R_m:n}) \text{ and } D_2 = \sum_{j=1}^{m+R_m} \mathbb{1}_{(-\infty,T_2]}(X_{j:m+R_m:n}),$$

the possible sampling situations are as in Fig. 5.2. This shows that the experimental time is bounded as

$$T_1 \le T_{\text{II}}^* \le T_2.$$

Finally, it has to be mentioned that the probability of observing no failure before T_2, i.e., $P(X_{m:m:n} > T_2) = \overline{F}_{m:m:n}(T_2)$, may be positive.

Chapter 6
Adaptive Progressive Type-II Censoring and Related Models

A crucial assumption in the design of the progressively censored experiment is that the censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$ is prefixed. However, although this assumption is assumed in the standard model, it may not be satisfied in real-life experiments since the experimenter may change the censoring numbers during the experiment (for whatever reasons). Therefore, it is desirable to have a model that takes into account such an adaptive process. Such a model has been proposed by Ng et al. [690] who introduced a (prefixed) threshold parameter $T > 0$ as a control parameter in their life-testing experiment. Given some prefixed censoring scheme $\mathscr{S} = (S_1, \ldots, S_m)$, this scheme is adapted after step $j^* = \max\{j : X_{j:m:n} < T\}$ such that no further censoring is carried out until the mth failure time has been observed. Hence, the censoring scheme is changed at the progressive censoring step $j^* + 1$, i.e., at the first observed failure time exceeding the threshold T. The effectively applied censoring scheme is

$$\mathscr{S}^* = (S_1, \ldots, S_{j^*}, 0^{*m-j^*-1}, n - m - \sum_{i=1}^{j^*} S_i). \tag{6.1}$$

Therefore, as long as the failures occur before time T, the initially planned progressive censoring scheme is employed. After passing time T, no items are withdrawn at all except for the last failure time when all remaining surviving units are removed.

This approach illustrates how an experimenter can control the experiment. If the interest is in getting observations early, then the experimenter will remove less units (or even none). If larger observed failure times are preferred, then the experimenter will remove more units at the beginning of the experiment. Therefore, a more flexible handling of the selection of the censoring scheme is preferable. Cramer and Iliopoulos [294] introduced a very general and flexible model where the next applied censoring number R_j may depend on both the previous numbers $R_1, \ldots, R_{j-1}$ and the observed failure times $x_{1:m:n}, \ldots, x_{j:m:n}$. They proposed a construction of

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_6,

the adaptive model using a stochastic kernel approach. For prefixed progressive censoring schemes, Beutner [192] has worked out this procedure.

6.1 General Model of Adaptive Progressive Type-II Censoring

The adaptive construction process is based on a life-testing experiment with n identical units having lifetimes $X_1, \ldots, X_n$. As before, it is desired to observe exactly m failures. The crucial change in the progressive censoring procedure is given by the fact that the future censoring numbers may depend on both past censoring numbers and observed failure times. Before introducing the construction process, we recall a property of progressively Type-II censored order statistics which will become important in the following definition. From the generation process of progressively Type-II censored order statistics given in Procedure 1.1.3, it follows that the jth progressively Type-II censored order statistic depends only on the first part of the (prefixed) censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$. In particular, we have

$$X_{1:m:n}^{\mathscr{R}} = X_{1:n}, \quad X_{j:m:n}^{\mathscr{R}} = X_{j:m:n}^{(R_1,\ldots,R_{j-1})}, \quad j = 2, \ldots, m. \tag{6.2}$$

This property is also evident from the representation of the joint density function $f^{X_{1:m:n}^{\mathscr{R}},\ldots,X_{j:m:n}^{\mathscr{R}}}$ given in Corollary 2.1.3. Since the adaption process works sequentially, we will use the representation in (6.2) with the incomplete censoring scheme to define adaptive progressively Type-II censored order statistics.

The generation process of adaptive progressively Type-II censored order statistics $Y_{(1)}, Y_{(2)}, \ldots, Y_{(m)}$ can be described as follows:

① The first failure time is the minimum of the random variables: $Y_{(1)} = X_{1:n} = X_{1:m:n}$;
② At time $Y_{(1)}$, a random number $R_1^\star$ of surviving units is removed from the experiment, where $R_1^\star$ has a distribution depending in some way on $Y_{(1)} = y_1$. Its support is given by $\{0, 1, \ldots, n-m\}$;
③ The minimum lifetime of the remaining $n-R_1^\star-1$ units is observed and denoted by $Y_{(2)} = X_{2:m:n}^{(R_1^\star)}$;
④ Then, a random number of $R_2^\star$ surviving items is withdrawn from the experiment. $R_2^\star$ is assumed to have a distribution depending in some way on $(Y_{(1)}, Y_{(2)}) = (y_1, y_2)$ and $R_1^\star = R_1$. Its support is given by the set $\{0, 1, \ldots, n-m-R_1\}$;
⑤ Continuing this process, we observe the vector of failure times

$$(Y_{(1)}, Y_{(2)}, \ldots, Y_{(m)}) = (X_{1:m:n}, X_{2:m:n}^{(R_1^\star)}, \ldots, X_{m:m:n}^{(R_1^\star,\ldots,R_{m-1}^\star)}).$$

Notice that the conditional distribution of the number of removed items $R_j^\star$ at censoring step j depends on $(Y_{(1)}, \ldots, Y_{(j)}) = (y_1, \ldots, y_j)$ and $(R_1^\star, \ldots, R_{j-1}^\star) = (R_1, \ldots, R_{j-1})$ with support $\{0, \ldots, n - m - \sum_{i=1}^{j-1} R_i\}$;

⑥ At the time of the mth failure $Y_{(m)}$, all remaining surviving units are removed.

For $k = 1, \ldots, m-1$, we define the sets of possible censoring sequences up to the censoring step k by

$$\mathscr{C}_{m,n}^k = \Big\{(R_1, \ldots, R_k) \in \mathbb{N}_0^k \;\Big|\; \sum_{i=1}^k R_i \le n - m\Big\}.$$

By analogy with the nonrandom censoring procedure, we define the (random) numbers $\Gamma_1, \ldots, \Gamma_m$ of units remaining in the experiment before a censoring step by

$$\Gamma_1 = n, \qquad \Gamma_j = n - j + 1 - \sum_{i=1}^{j-1} R_i^\star, \quad j = 2, \ldots, m.$$

The adaptive progressively censored data is given by the random vectors $\mathbf{Y}_{(m)}$ and $\mathbf{R}_m^\star$. If no confusion is possible, we will use for brevity the notation $Y_{(1)}, \ldots, Y_{(k)}$ for adaptive progressively Type-II censored order statistics throughout this chapter. If necessary, a notation as in (6.2) will be used to emphasize the dependence on a particular censoring scheme:

$$\mathbf{Y}_{(1)} = X_{1:m:n}, \quad \mathbf{Y}_{(k)} = X_{k:m:n}^{\mathbf{R}_{k-1}^\star}, \quad k = 2, \ldots, m.$$

This illustrates that the distribution of $\mathbf{Y}_{(k)}$ depends only on the first $k-1$ components of the random censoring scheme $\mathbf{R}_m^\star$.

Distributional Assumptions

In order to propose a probabilistic model for adaptive progressive Type-II censoring, Cramer and Iliopoulos [294] imposed the following assumptions on the considered distributions:

(i) The lifetimes $X_1, \ldots, X_n$ are supposed to be independent and have an absolutely continuous cumulative distribution function F with density function f;

(ii) The conditional distribution of $Y_{(j)}$, given $\mathbf{Y}_{(j-1)} = \mathbf{y}_{j-1}$ and $\mathbf{R}_{j-1}^\star = \mathbf{R}_{j-1}$, depends only on y_{j-1} and $\mathbf{R}_{j-1}$ and is the same as the distribution of the minimum in a sample of size $\gamma_j = n - j + 1 - \sum_{i=1}^{j-1} R_j$ from the left truncated cumulative distribution function $G_{y_{j-1}} = \frac{F(\cdot)-F(y_{j-1})}{1-F(y_{j-1})}$. Thus, for $j = 2, \ldots, m$,

$$Y_{(j)}|\{\mathbf{Y}_{(j-1)} = \mathbf{y}_{j-1}, \mathbf{R}^\star_{j-1} = \mathbf{R}_{j-1}\} \stackrel{\mathsf{d}}{=} Y_{(j)}|\{Y_{(j-1)} = y_{j-1}, \mathbf{R}^\star_{j-1} = \mathbf{R}_{j-1}\}.$$

Therefore, for $y_{j-1} \leq y_j$ with $F(y_{j-1}) < 1$, this conditional distribution has density function (cf. (2.35) and Theorem 2.5.2)

$$f_j(y_j|y_{j-1}, \mathbf{R}_{j-1}) = \gamma_j \frac{f(y_j)}{1 - F(y_{j-1})} \left\{ \frac{1 - F(y_j)}{1 - F(y_{j-1})} \right\}^{\gamma_j - 1};$$

(iii) Finally, the conditional probability mass function of $R^\star_1$, given $\mathbf{Y}_{(1)} = \mathbf{y}_1$, is denoted by $g_1(\cdot|\mathbf{y}_1)$. $g_j(\cdot|\mathbf{y}_j, \mathbf{R}_{j-1})$ denotes the probability mass function of $R^\star_j$, given $\mathbf{Y}_{(j)} = \mathbf{y}_j$ and $\mathbf{R}^\star_{j-1} = \mathbf{R}_{j-1}$, $j = 2, \ldots, m-1$. According to the construction process, the probability mass function of $R^\star_m$, given $\mathbf{Y}_{(m)} = \mathbf{y}_m$ and $\mathbf{R}^\star_{m-1} = \mathbf{R}_{m-1}$, is a one-point distribution, namely,

$$g_m(R_m|\mathbf{y}_m, \mathbf{R}_{m-1}) = \mathbb{1}_{\mathscr{C}^m_{m,n}}(\mathbf{R}_m).$$

This condition ensures that the censoring scheme is admissible.

Since the first observation is the minimum of the lifetimes, the density function of $Y_{(1)}$ is given by

$$f_1(y_1) = nf(y_1)\{1 - F(y_1)\}^{n-1}, \quad y_1 \in \mathbb{R}.$$

Then, by the assumptions given above, the joint density function $f^{\mathbf{Y}_{(k)}, \mathbf{R}^\star_k}$ of $\mathbf{Y}_{(k)}$, $\mathbf{R}^\star_k$ can be expressed as

$$f^{\mathbf{Y}_{(k)}, \mathbf{R}^\star_k}(\mathbf{y}_k, \mathbf{R}_k) = f^*_k(\mathbf{y}_k|\mathbf{R}_{k-1}) \cdot g^*_k(\mathbf{R}_k|\mathbf{y}_k), \quad k = 1, \ldots, m, \tag{6.3}$$

where

$$f^*_k(\mathbf{y}_k|\mathbf{R}_{k-1}) = \left(\prod_{j=1}^{k-1} \left[\gamma_j f(y_j)\{1 - F(y_j)\}^{R_j} \right] \right) \\ \times \gamma_k f(y_k)\{1 - F(y_k)\}^{\gamma_k - 1} \mathbb{1}_{\mathbb{R}^k_\leq}(\mathbf{y}_k), \tag{6.4}$$

$$g^*_k(\mathbf{R}_k|\mathbf{y}_k) = \prod_{j=1}^{k} g_j(R_j|\mathbf{y}_j, \mathbf{R}_{j-1}) \mathbb{1}_{\mathscr{C}^k_{m,n}}(\mathbf{R}_k).$$

$f^*_1(\cdot|\mathbf{R}_0)$ is defined as f_1. Cramer and Iliopoulos [294] observed that

$$\int_{\mathbb{R}^k_\leq} f^*_k(\mathbf{y}_k|\mathbf{R}_{k-1}) d\mathbf{y}_k = 1$$

for all $\mathbf{R}_{k-1} \in \mathscr{C}_{m,n}^{k-1}$ although $f_k^*(\cdot|\mathbf{R}_{k-1})$ is in general *not* the conditional distribution of $\mathbf{Y}_{(k)}$, given $\mathbf{R}_{k-1}^\star = \mathbf{R}_{k-1}$. Thus, for any $\mathbf{R}_{k-1} \in \mathscr{C}_{m,n}^{k-1}$, $f_k^*(\cdot|\mathbf{R}_{k-1})$ is a proper density function on $\mathbb{R}^k$. In particular, $f_k^*(\cdot|\mathbf{R}_{k-1})$ is exactly the marginal density function of the first k progressively Type-II censored order statistics $X_{1:m:n}^{\mathscr{R}}, \ldots, X_{k:m:n}^{\mathscr{R}}$ from the population distribution function F with censoring scheme $\mathscr{R} = (\mathbf{R}_{k-1}, R_k, \ldots, R_m) \in \mathscr{C}_{m,n}^m$. Similarly,

$$\sum_{\mathbf{R}_k} g_k^*(\mathbf{R}_k|\mathbf{y}_k) = 1$$

for all $\mathbf{y}_k \in \mathbb{R}_\leq^k$ although $g_k^*(\cdot|\mathbf{y}_k)$ is *not* the conditional distribution of $\mathbf{R}_k^\star$, given $\mathbf{Y}_{(k)} = \mathbf{y}_k$. Hence, $g_k^*(\cdot|\mathbf{y}_k)$ is a probability mass function on the set $\mathscr{C}_{m,n}^k$ for any $\mathbf{y}_k \in \mathbb{R}_\leq^k$.

These results lead directly to the following properties established by Cramer and Iliopoulos [294] which we state without proofs:

(i) For any $k = 1, \ldots, m-1$, the marginal distribution of $(\mathbf{Y}_{(k)}, \mathbf{R}_{k-1}^\star)$ has the density function $f^{\mathbf{Y}_{(k)}, \mathbf{R}_{k-1}^\star}$ given in (6.3);

(ii) The conditional distribution of $(Y_{(k+1)}, \ldots, Y_{(m)}, R_{k+1}^\star, \ldots, R_m^\star)$, given $(\mathbf{Y}_{(k)}, \mathbf{R}_k^\star) = (\mathbf{y}_k, \mathbf{R}_k)$, is the same as that of $(\mathbf{Y}_{(m-k)}, \mathbf{R}_{m-k}^\star)$ but with cumulative distribution function F left truncated at y_k, $\gamma_{k+1} = n-k-\sum_{i=1}^k R_i$ in the place of n, and with probability mass function g_j's depending also on $\mathbf{y}_k, \mathbf{R}_k$;

(iii) Let $U_1, \ldots, U_m$ be independent uniform random variables and let $R_1^\star, \ldots,$ $R_m^\star$ be nonnegative integer valued random variables with conditional joint probability mass function

$$f^{\mathbf{R}_m^\star|\mathbf{U}_m}(\mathbf{R}_m|\mathbf{u}_m) = g_0^*(\mathbf{R}_m|\mathbf{u}_m) = \prod_{j=1}^m g_{0,j}(R_j|\mathbf{u}_j, \mathbf{R}_{j-1}) \mathbb{1}_{\mathscr{C}_{m,n}^m}(\mathbf{R}_m),$$

for some (conditional) probability mass functions $g_{0,1}, \ldots, g_{0,m}$. Then, the following stochastic representation of adaptive progressively Type-II censored order statistics holds.

Let F be a strictly increasing and absolutely continuous cumulative distribution function. For $j = 1, \ldots, m$, define the random variable $Y_{(j)}$ by

$$Y_{(j)} = F^{\leftarrow}\big(1 - \textstyle\prod_{i=1}^j U_i^{1/\Gamma_i}\big).$$

For $\mathbf{R}_m \in \mathscr{C}_{m,n}^m$ and $u_1, \ldots, u_m \in [0,1]$, let $y_j = F^{\leftarrow}\big(1 - \prod_{i=1}^j u_i^{1/\gamma_i}\big)$, $1 \leq j \leq m$, and

$$g_j(R_j|\mathbf{y}_j, \mathbf{R}_{j-1}) = g_{0,j}\big(R_j\big|\mathbf{u}_j, \mathbf{R}_{j-1}\big), \quad j = 1, \ldots, m,$$

where $\gamma_j = n - j + 1 - \sum_{i=1}^{j-1} R_i$, $1 \le j \le m$. Then, the joint density function of $\mathbf{Y}_{(m)}$ and $\mathbf{R}_m^\star$ is given by (6.3);

(iv) Suppose there is some $k \ge 0$ such that for any $j = 1, 2, \ldots$, the conditional distribution of Γ_{j+1}, given $\boldsymbol{\Gamma}_j = \boldsymbol{\gamma}_j, \mathbf{Y}_{(j)} = \mathbf{y}_j$, depends only on $y_j, \ldots, y_{j-k}, \gamma_j, \ldots, \gamma_{j-k}$.
Then, the sequence of random vectors $(\Gamma_{j-k}, \ldots, \Gamma_j, Y_{(j-k)}, \ldots, Y_{(j)})$, $j = k+1, k+2, \ldots$, forms a Markov chain.

An important property is summarized in the following theorem. It provides a simple tool to calculate the distribution of certain statistics.

Theorem 6.1.1 (Cramer and Iliopoulos [294]). Let $\mathcal{T} : (0,\infty)^m \times [0,\infty)^m \longrightarrow \mathbb{R}^m$ be a measurable function satisfying the following conditions:

(i) The mapping $\mathcal{V} : (0,\infty)^m \times [0,\infty)^m \longrightarrow \mathbb{R}^m \times [0,\infty)^m$ defined by

$$(\mathbf{y}_m, \mathbf{R}_m) \mapsto (\mathbf{t}_m, \mathbf{R}_m) = (\mathcal{T}(\mathbf{y}_m, \mathbf{R}_m), \mathbf{R}_m)$$

is bijective and differentiable such that its Jacobian matrix is regular;

(ii) Given a fixed progressive censoring scheme $\mathbf{R}_m = \mathscr{R} \in \mathscr{C}_{m,n}^m$, the distribution of $\mathcal{T}(\mathbf{X}^{\mathscr{R}}, \mathscr{R})$ does not depend on $\mathscr{R}$, where $\mathbf{X}^{\mathscr{R}}$ is the vector of the corresponding progressively Type-II censored order statistics with population distribution function F;

(iii) For any $\mathbf{R}_m \in \mathscr{C}_{m,n}^m$ and $j = 1, \ldots, m-1$, the first j components of the inverse transform

$$\mathbf{y}_m = \left(\mathcal{V}^{-1}(\mathbf{t}_m, \mathbf{R}_m)\right)_1$$

do not depend on $R_{j-1}, \ldots, R_m$.

Then, $\mathcal{T}(\mathbf{Y}_{(m)}, \mathbf{R}_m^\star)$ has the same distribution as $\mathcal{T}(\mathbf{X}^{\mathscr{R}}, \mathscr{R})$ for any fixed censoring scheme $\mathbf{R}_m = \mathscr{R}$.

The preceding theorem leads directly to the following property of spacings which is an extension of Theorem 2.3.2 to the adaptive model. In order to get the independence of the spacings in the adaptive censoring model, the normalization is important because Γ_j and $Y_{(j)}$ are not independent in general.

Theorem 6.1.2 (Cramer and Iliopoulos [294]). Let $Y_{(1)}, \ldots, Y_{(m)}$ be adaptively progressively Type-II censored order statistics generated from a two-parameter exponential distribution $\mathrm{Exp}(\mu, \vartheta)$. Then, the normalized spacings

$$D_j = \Gamma_j (Y_{(j)} - Y_{(j-1)}), \quad j = 1, \ldots, m, \tag{6.5}$$

where $Y_{(0)} = \mu$, are independent $\mathrm{Exp}(\vartheta)$ distributed random variables. Moreover,

$$S = \sum_{i=1}^{m} (1 + R_i^\star) Y_{(i)} \sim \Gamma(\vartheta, m).$$

Proof. Suppose $Z_i \sim \mathsf{Exp}(\mu, \vartheta)$ are IID random variables with $\mu \in \mathbb{R}, \vartheta > 0$, $1 \le i \le n$. Then, according to Theorem 2.3.2, for any (fixed) censoring scheme $\mathbf{R}_m = (R_1, \ldots, R_m)$, the normalized spacings $D_j = \gamma_j (Z_{j:m:n} - Z_{j-1:m:n})$, $j = 1, \ldots, m$, where $Z_{0:m:n} = \mu$, are independent $\mathsf{Exp}(\vartheta)$ distributed random variables. Their sum

$$\sum_{i=1}^{m} D_i = \sum_{i=1}^{m} (1 + R_i) X_{i:m:n}$$

has a gamma distribution $\Gamma(\vartheta, m)$. Since the joint distribution of the normalized spacings is independent of the censoring scheme $\mathbf{R}_m$, Theorem 6.1.1 shows that the spacings in (6.5) defined via the adaptively progressively Type-II censored order statistics are independent $\mathsf{Exp}(\vartheta)$ distributed random variables. This also yields the distribution of S to be a $\Gamma(\vartheta, m)$ distribution and thus invariant w.r.t. the adaptive censoring procedure. □

In particular, Theorem 6.1.2 shows that $Y_{(1)}$ and $Y_{(2)}, \ldots, Y_{(m)}$ are independent. This result is also true in the nonadaptive case. It is important for properties of the maximum likelihood estimators in the two-parameter exponential case. Moreover, Theorem 6.1.2 yields the representation

$$Y_{(j)} = \mu + \vartheta \sum_{i=1}^{j} \frac{1}{\Gamma_j} Z_j^*, \quad j = 1, \ldots, m,$$

where $Z_1^*, \ldots, Z_m^*$ are IID standard exponential random variables [cf. (2.13)]. Notice that Γ_j and Z_j^* are generally not independent.

A similar result can be established for normalized ratios of adaptively progressively Type-II censored order statistics from generalized Pareto distributions.

Example 6.1.3. For $X_j \sim \mathsf{Pareto}(\alpha)$, $1 \le j \le n$, the random variables

$$Y_{(1)}^{\Gamma_1 \alpha} \quad \text{and} \quad \left(\frac{Y_{(j)}}{Y_{(j-1)}}\right)^{\Gamma_j \alpha}, \quad j = 2, \ldots, m,$$

are independent with Pareto(1) distributions (cf. Corollary 2.3.14). For reflected power function distributions RPower(β), we have

$$(1 - Y_{(1)})^{\Gamma_1 \beta} \quad \text{and} \quad \left(\frac{1 - Y_{(j)}}{1 - Y_{(j-1)}}\right)^{\Gamma_j \beta}, \quad j = 2, \ldots, m,$$

to be independent uniform random variables (cf. Corollary 2.3.11). For $\beta = 1$, it includes the uniform distribution as population distribution.

6.2 Particular Models

The general approach of adaptive progressive Type-II censoring covers some particular models that have been discussed in the literature. Subsequently, we illustrate this connection by presenting the corresponding density functions and probability mass functions as given above.

6.2.1 Nonadaptive Type-II Progressive Censoring

It is clear from the construction of the adaptive censoring model that nonadaptive Type-II progressive censoring can be embedded into this more general approach by choosing one-point distributions for the censoring scheme which are independent of the observed failure times. In particular, let $\mathscr{S} = (S_1, \dots, S_m) \in \mathscr{C}^m_{m,n}$ be a censoring scheme and let g_j be the probability mass function of a one-point distribution in S_j, i.e.,

$$g_j(\cdot|\mathbf{y}_j, \mathbf{R}_{j-1}) = \mathbb{1}_{\{S_j\}}(\cdot), \quad 1 \le j \le m.$$

Then, $g^*_m(\cdot|\mathbf{y}_m) = \mathbb{1}_{\{\mathscr{S}\}}(\cdot)$ is the probability mass function of a one-point distribution in the censoring scheme $\mathscr{S}$. Thus, the joint density function in (6.3) becomes

$$f^{\mathbf{Y}_{(m)},\mathbf{R}^\star_m}(\mathbf{y}_m, \mathbf{R}_m) = \prod_{j=1}^m \left[\eta_j f(y_j)\{1 - F(y_j)\}^{S_j}\right] I_{\mathbb{R}^m_\le}(\mathbf{y}_m)\mathbb{1}_{\{\mathscr{S}\}}(\mathbf{R}_m)$$

with $\eta_j = \sum_{i=j}^m (S_i + 1)$, $1 \le j \le m$. Hence, $f^{\mathbf{Y}_{(m)}} = f^{\mathbf{X}^{\mathscr{S}}}$.

6.2.2 Ng–Kundu–Chan Model

The Ng–Kundu–Chan model introduced in Ng et al. [690] has been explained at the beginning of this chapter. For a (prefixed) threshold parameter $T > 0$ and a prefixed censoring scheme $\mathscr{S} = (S_1, \dots, S_m)$, the original scheme $\mathscr{S}$ is changed when the first observed failure time exceeds the threshold T. With $j^* = \max\{j : X^{\mathscr{S}}_{j:m:n} < T\}$, the adaptive censoring scheme is given in (6.1). The resulting adaptive censoring scheme is given by the censoring plan

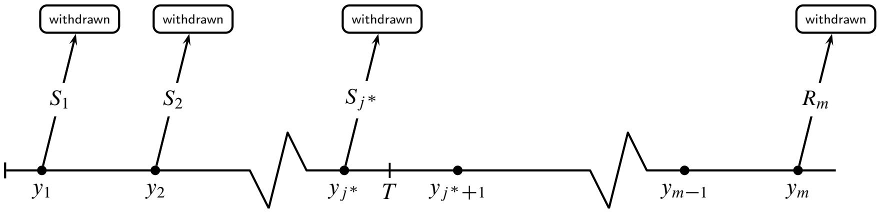

Fig. 6.1 Generation process of adaptive progressively Type-II censored order statistics in the Ng–Kundu–Chan model when $X_{m:m:n} > T$

$$(S_1, \ldots, S_{j*}, 0^{*m-j^*-1}, n - m - \sum_{i=1}^{j*} S_i) \text{ provided } X^{\mathscr{S}}_{j*:m:n} < T \leq X^{\mathscr{S}}_{j*+1:m:n}.$$

Hence, as long as the failures occur before time T, the initially planned censoring scheme is employed. However, a change of the censoring plan is made only if $j^* \leq m - 2$. Otherwise, the originally planned censoring scheme is employed. Furthermore, the first $j^* + 1$ observations are identical in both progressive censoring models for $j^* \leq m - 2$ so that

$$(X^{\mathscr{S}}_{1:m:n}, \ldots, X^{\mathscr{S}}_{j*+1:m:n}) = (Y_{(1)}, \ldots, Y_{(j*+1)}).$$

Figure 6.1 depicts the generation procedure of adaptive progressively Type-II censored order statistics in the Ng–Kundu–Chan model. After exceeding T, the censoring numbers except for R_m are defined to be zero. This means that, after T, the next successive $m - j^* - 1$ failures are observed. The experiment is terminated when m failures have been observed. If $T > X_{m-1:m:n}$, no adaption is carried out.

Choosing the probability mass functions as $(1 \leq j \leq m - 1)$

$$\begin{aligned} g_j(R_j|\mathbf{y}_j, \mathbf{R}_{j-1}) &= g_j(R_j|y_j) \\ &= \mathbb{1}_{\{S_j\}}(R_j)\mathbb{1}_{(-\infty,T)}(y_j) + \mathbb{1}_{\{0\}}(R_j)\mathbb{1}_{[T,\infty)}(y_j), \end{aligned}$$

the model can be seen as a particular case of adaptive progressive Type-II censoring. It turns out that the adaptive process takes only into account the observed failure times.

6.2.3 *Flexible Progressive Censoring*

Bairamov and Parsi [80] proposed a modification of progressive censoring method called flexible progressive censoring for IID lifetimes $X_1, \ldots, X_n$. We present the method in a slightly more general way w.r.t. the restrictions imposed on the censoring numbers. Given thresholds $T_1 \le \cdots \le T_m$ and censoring numbers $S_1, \ldots, S_{m-1}, S_1^*, \ldots, S_{m-1}^*$ with $n - \sum_{j=1}^{m-1} \max\{S_j, S_j^*\} \ge m$, the procedure works as follows:

① The first observation is given by $Y_{(1)} = X_{1:m:n}$. Then,

$$R_1^\star = S_1^* + (S_1 - S_1^*)\mathbb{1}_{(-\infty,T_1]}(X_{1:m:n})$$

units are removed from the experiment;

② For $j \in \{2, \ldots, m-1\}$, $Y_{(j)} = X_{j:m:n}^{\mathbf{R}_{j-1}^\star}$ denotes the next failure time after removing $R_{j-1}^\star$ units in the $(j-1)$th censoring step. Then,

$$R_j^\star = S_j^* + (S_j - S_j^*)\mathbb{1}_{(-\infty,T_j]}(X_{j:m:n}^{\mathbf{R}_{j-1}^\star})$$

units are randomly withdrawn from the life-testing experiment;

③ After observing the mth failure time, the remaining $n - \sum_{j=1}^{m-1}(R_j^\star + 1) - 1$ items are censored.

Choosing the probability mass functions as $(1 \le j \le m-1)$

$$\begin{aligned} g_j(R_j|\mathbf{y}_j, \mathbf{R}_{j-1}) &= g_j(R_j|y_j) \\ &= \mathbb{1}_{\{S_j\}}(R_j)\mathbb{1}_{(-\infty,T_j]}(y_j) + \mathbb{1}_{\{S_j^*\}}(R_j)\mathbb{1}_{(T_j,\infty)}(y_j), \end{aligned}$$

the model can be seen as a particular case of adaptive progressive Type-II censoring. On the other hand, it may be viewed as a generalization of the Ng–Kundu–Chan model (here $T_j = T$ and $S_j^* = 0$, $j = 1, \ldots, m$). A generalization of this model has been proposed by Kinaci [532].

6.2.4 *Progressive Censoring with Random Removals*

Progressive censoring with random removals has been introduced by Yuen and Tse [936]. Here, the censoring numbers are chosen according to some probability distribution on the set of possible censoring numbers. The support of the distributions is chosen so that a total of m observations is guaranteed. Moreover, the selection of the censoring scheme is assumed to be independent of the lifetimes $X_1, \ldots, X_n$. Therefore, this sampling scheme can be seen as a two-step procedure:

① A random censoring scheme $\mathbf{R}_m^\star$ is chosen according to some given discrete distribution on the set $\mathscr{C}_{m,n}^n$ of admissible censoring schemes;

② The progressive censoring procedure is carried out with the resulting censoring scheme of this random experiment.

This approach can be embedded into the model by Cramer and Iliopoulos [294] by choosing the conditional probability mass functions g_j to be independent of $\mathbf{y}_m$:

$$g_m^*(\mathbf{R}_m|\mathbf{y}_m) \equiv g_m^*(\mathbf{R}_m) = \prod_{j=1}^{m} g_j(R_j|\mathbf{R}_{j-1})\mathbb{1}_{\mathscr{C}_{m,n}^m}(\mathbf{R}_m).$$

Therefore, g_m^* depends only on the censoring scheme but not on the failure times. Obviously, g_m^* is the marginal probability mass function of $\mathbf{R}_m^\star$. Moreover, it follows directly from (6.4) that the distribution of $\mathbf{Y}_{(m)}$ is a mixture, with mixing probabilities $g^*(\mathbf{R}_m)$, given by

$$f^{\mathbf{Y}_{(m)}}(\mathbf{y}_m) = \sum_{\mathbf{R}_m \in \mathscr{C}_{m,n}^m} f_m^*(\mathbf{y}_m|\mathbf{R}_{m-1}) \cdot g^*(\mathbf{R}_m). \tag{6.6}$$

Chapter 7
Moments of Progressively Type-II Censored Order Statistics

7.1 General Distributions

7.1.1 Representations for Moments

Moments of progressively Type-II censored order statistics can be calculated by standard approaches. If the population distribution has a density function f, then, for $k \in \mathbb{N}_0$,

$$EX_{r:m:n}^k = \int_{-\infty}^{\infty} t^k f^{X_{r:m:n}}(t)dt,$$

where $f^{X_{r:m:n}}$ is as in (2.28). From the quantile representation Theorem 2.1.1, we obtain the following important, general integral representation of moments. It relates the moments of a progressively Type-II censored order statistic to the density function of a uniform progressively Type-II censored order statistic. The expression is very useful in the derivation of bounds (see Sect. 7.5).

Theorem 7.1.1. Let $X_{r:m:n}$ be a progressively Type-II censored order statistic from a cumulative distribution function F, $1 \le r \le m$, and $k \ge 0$. Then,

$$EX_{r:m:n}^k = \int_0^1 (F^{\leftarrow}(t))^k f^{U_{r:m:n}}(t)\, dt$$

provided the moment exists.

Using the above relation for a progressively Type-II censored order statistic $X_{r:m:n}$, we can apply (2.28) to obtain

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_7,

$$EX_{r:m:n}^k = \Big(\prod_{i=1}^r \gamma_i\Big) \sum_{j=1}^r \frac{1}{\gamma_j} a_{j,r} \int_0^\infty (F^{\leftarrow}(t))^k \gamma_j (1-t)^{\gamma_j - 1} dt$$

$$= \Big(\prod_{i=1}^r \gamma_i\Big) \sum_{j=1}^r \frac{1}{\gamma_j} a_{j,r} EX_{1:\gamma_j}^k, \tag{7.1}$$

provided the involved moments exist. This representation proves that the expectation of a progressively Type-II censored order statistic can be expressed in terms of expectations of minima. The existence problem is discussed in Sect. 7.1.2.

Alternatively, the expectation of a random variable $X \sim F$ can be calculated from the general formula (see David and Nagaraja [327, p. 38])

$$EX = \int_0^\infty [1 - F(t)]dt - \int_{-\infty}^0 F(t)dt.$$

It should be noted that a similar expression for product moments can be obtained. In particular, using (2.29), we obtain for $k_1 \le k_2$ the expression

$$EX_{k_1:m:n} X_{k_2:m:n}$$

$$= \Big(\prod_{j=1}^{k_2} \gamma_j\Big) \sum_{i=1}^{k_1} a_{i,k_1} \sum_{j=k_1+1}^{k_2} a_{j,k_2}^{(k_1)} \frac{1}{\gamma_j(\gamma_i - \gamma_j)} E[X_{1:\gamma_j} X_{1:\gamma_i - \gamma_j}].$$

This shows that it is sufficient to compute the product moments of two minima from the same sample in order to compute the product moment of two progressively Type-II censored order statistics.

7.1.2 Existence of Moments

The existence of some positive moment of a generalized order statistic can be ensured by the existence of a higher moment for the underlying distribution. The following results are taken from Cramer et al. [310].

Theorem 7.1.2. Let X be a random variable with distribution function F, and let $X_{r:m:n}$ be a progressively Type-II censored order statistic from F. If $E|X|^\beta < \infty$ for some $\beta > 0$, then $E|X_{r:m:n}|^\alpha < \infty$ for all $0 < \alpha \le \beta$.

Even for order statistics, the above conditions are unsatisfactory, since in the case of an underlying Cauchy distribution, the expectation EX does not exist, whereas the expected values of the corresponding order statistics are finite except for the first and the last one. Sen [789] presented a well-known theorem on this topic, which has been extended to m-generalized order statistics in Kamps [498, Theorem II.1.2.2]. Cramer et al. [310] established an extended version of Sen's theorem without requiring any restriction on the model parameters. For progressively Type-II censored order statistics, their result reads as follows.

Theorem 7.1.3. Let X have a continuous distribution function F, and let the progressively Type-II censored order statistic $X_{r:m:n}$ be based on F. If $E|X|^\beta < \infty$ for some $\beta > 0$, then $E|X_{r:m:n}|^\alpha < \infty$ for all $\alpha > \beta$, which satisfy the condition

$$\frac{\alpha}{\beta} \le r \le \gamma_r + r - \frac{\alpha}{\beta} \quad \text{or, equivalently,} \quad \alpha \le \beta \cdot \min\{r, \gamma_r\}. \tag{7.2}$$

Choosing $m = n$ and $\mathscr{R} = (0^{*m})$, the following result due to Sen [789] is contained as a special case.

Corollary 7.1.4 (Sen [789]). Let X have a continuous distribution function F, and let the order statistic $X_{r:n}$, $1 \le r \le n$, be based on F. If $E|X|^\beta < \infty$ for some $\beta > 0$, then $E|X_{r:n}|^\alpha < \infty$ for all $\alpha > \beta$, which satisfy the condition

$$\frac{\alpha}{\beta} \le r \le n + 1 - \frac{\alpha}{\beta} \quad \text{or, equivalently,} \quad \alpha \le \beta \cdot \min\{r, n - r + 1\}.$$

Example 7.1.5. For a standard Cauchy$(0, 1)$-distributed random variable X, $E|X|^\beta < \infty$ for $0 < \beta < 1$. Choosing $\beta = \frac{1}{2}$, $\alpha = 1$, we find from (7.2) that the first moment of a progressively Type-II censored order statistic $X_{r:m:n}$ exists if $2 \le \min\{r, \gamma_r\}$. This implies that the first moment exists for $2 \le r \le m - 1$. The first moment of $X_{m:m:n}$ does exist if $R_m \ge 1$. Similar considerations with $\alpha = 2$ lead to conditions for second moments.

Subsequently, we derive bounds for expectations of functions of progressively Type-II censored order statistics. These are based on inequalities for the joint density function $f^{\mathbf{X}^{\mathscr{R}}}$ given in (2.4). Further results on bounds for progressively Type-II censored order statistics are presented in Sect. 7.5.

First, we establish an upper bound for moments on progressively Type-II censored order statistics in terms of moments of order statistics. The results are based on (7.3) with $x_1 \le \cdots \le x_m$:

$$\begin{aligned} f^{X^{\mathscr{R}}_{1:m:n},\ldots,X^{\mathscr{R}}_{m:m:n}}(\mathbf{x}_m) &= c(\mathscr{R}) \prod_{j=1}^{m} f(x_j)\overline{F}(x_j)^{R_j} \\ &\le \frac{c(\mathscr{R})}{m!} f^{X_{1:m},\ldots,X_{m:m}}(\mathbf{x}_m) \quad \text{a.s.} \end{aligned} \tag{7.3}$$

for an arbitrary censoring scheme $\mathscr{R}$. This fundamental inequality applies to every selection of (progressively Type-II censored) order statistics as well. Given a selection $1 \le r_1 < \cdots < r_s \le n$, $s \in \mathbb{N}$, integration of (7.3) yields

$$f^{X^{\mathscr{R}}_{r_1:m:n},\ldots,X^{\mathscr{R}}_{r_s:m:n}} \leq \frac{c(\mathscr{R})}{m!} f^{X_{r_1:m},\ldots,X_{r_s:m}} \quad \text{a.s.}$$

Hence, the preceding result leads directly to a simple criterion for the existence of moments of progressively Type-II censored order statistics. The result can be extended to arbitrary product moments of progressively Type-II censored order statistics.

Theorem 7.1.6. Let g be a nonnegative function. Then

$$Eg(X^{\mathscr{R}}_{r:m:n}) \leq \frac{c(\mathscr{R})}{m!} Eg(X_{r:m}), \quad 1 \leq r \leq m.$$

In particular, this yields $E|X^{\mathscr{R}}_{r:m:n}|^k \leq \frac{c(\mathscr{R})}{m!} E|X_{r:m}|^k$ so that the existence of the kth moment of the rth order statistic implies the existence of the respective moment of $X^{\mathscr{R}}_{r:m:n}$.

Utilizing this conclusion, Sen's [789] result given in Corollary 7.1.4 can be directly applied to progressively Type-II censored order statistics. However, it should be noted that Theorem 7.1.3 is weaker since the upper bound in (7.2) is $m + 1 + \sum_{j=r}^{m} R_j$ and, thus, usually is larger. The difference is given by $\sum_{j=r}^{m} R_j$.

Corollary 7.1.7. Let $E|X|^{\beta}$ exist for some $\beta > 0$. Then, $E(X^{\mathscr{R}}_{r:m:n})^k$ exists for all r satisfying $\lfloor \frac{k}{\beta} \rfloor \leq r \leq m + 1 - \lfloor \frac{k}{\beta} \rfloor$.

It has to be noted that the existence of the kth moment of the rth progressively Type-II censored order statistic does not generally imply the existence of the kth moment of the rth order statistic.

Example 7.1.8. Consider a Pareto(q)-distribution with $q > 0$. Since the kth moment of the rth progressively Type-II censored order statistic from a Pareto distribution is given by $E(X^{\mathscr{R}}_{r:m:n})^k = \prod_{j=1}^{r} \frac{\gamma_j}{\gamma_j - k/q}$ (see Theorem 7.2.5 and Cramer and Kamps [300, p. 334 in terms of generalized order statistics]), we obtain the existence condition

$$\gamma_r = \sum_{i=r}^{m} (R_i + 1) > k/q. \tag{7.4}$$

Since order statistics correspond to the censoring scheme (0^{*m}), the existence condition in this case is $m + r - 1 > k/q$ where the left-hand side is strictly smaller than that of (7.4) provided that $R_i \geq 1$ for some $i \in \{1, \ldots, m\}$.

A slight generalization of the above inequalities can be obtained by the following idea. Let $\mathscr{R} = (R_1, \ldots, R_m)$ be a given censoring scheme and $R_{\min} = \min_{1 \leq i \leq n} R_i$ and $R_{\max} = \max_{1 \leq i \leq n} R_i$ be the minimum and maximum censoring numbers, respectively. This yields the bounds

$$\frac{c(\mathscr{R})}{m!} f^{Y_{1:m},\dots,Y_{m:m}} \le f^{X^{\mathscr{R}}_{1:m:n},\dots,X^{\mathscr{R}}_{m:m:n}} \le \frac{c(\mathscr{R})}{m!} f^{V_{1:m},\dots,V_{m:m}} \quad \text{a.s.},$$

where $Y_{j:m}$ and $V_{j:m}$, $1 \le j \le m$, denote order statistics based on the cumulative distribution functions $1-\overline{F}^{R_{\max}}$ and $1-\overline{F}^{R_{\min}}$, respectively. Hence, the corresponding bounds for the expectation of $g(X^{\mathscr{R}}_{r:m:n})$, where g is a nonnegative function, become

$$\frac{c(\mathscr{R})}{m!} Eg(Y_{r:m}) \le Eg(X^{\mathscr{R}}_{r:m:n}) \le \frac{c(\mathscr{R})}{m!} Eg(V_{r:m}), \quad 1 \le r \le m.$$

Now, we consider the existence of moments of progressively Type-II censored order statistics w.r.t. some partial ordering $\preccurlyeq$ of the censoring schemes.

Definition 7.1.9. Let $\mathscr{R}, \mathscr{S} \in \mathbb{N}_0^m$ be two censoring schemes with $\sum_{i=1}^m R_i = \sum_{i=1}^m S_i$. Then

$$\mathscr{S} \preccurlyeq \mathscr{R} \iff \sum_{i=1}^k S_i \le \sum_{i=1}^k R_i, \quad k = 1,\dots,m-1.$$

It is easy to see that $\preccurlyeq$ defines a partial ordering on the set of all admissible censoring schemes $\mathscr{C}^m_{m,n}$ given in (1.1) with a fixed censoring number $m \in \mathbb{N}_0$. Its application leads to the following inequality.

Proposition 7.1.1 (Cramer [284]). *Let $\mathscr{R}, \mathscr{S} \in \mathscr{C}^m_{m,n}$ be two censoring schemes with $\mathscr{S} \preccurlyeq \mathscr{R}$. Then,*

$$f^{X^{\mathscr{S}}_{1:m:n},\dots,X^{\mathscr{S}}_{m:m:n}} \le \frac{c(\mathscr{S})}{c(\mathscr{R})} f^{X^{\mathscr{R}}_{1:m:n},\dots,X^{\mathscr{R}}_{m:m:n}}. \tag{7.5}$$

From the preceding result, we can deduce the following corollary.

Corollary 7.1.10. Let $\mathscr{R}, \mathscr{S} \in \mathscr{C}^m_{m,n}$ be two censoring schemes with $\mathscr{S} \preccurlyeq \mathscr{R}$ and g be a nonnegative function. Then,

$$Eg(X^{\mathscr{S}}_{m:m:n}) \le \frac{c(\mathscr{S})}{c(\mathscr{R})} Eg(X^{\mathscr{R}}_{m:m:n}).$$

In particular, this yields $E|X^{\mathscr{S}}_{m:m:n}|^k \le \frac{c(\mathscr{S})}{c(\mathscr{R})} E|X^{\mathscr{R}}_{m:m:n}|^k$ so that the existence of the kth moment of $X^{\mathscr{R}}_{m:m:n}$ implies the existence of the respective moment of $X^{\mathscr{S}}_{m:m:n}$.

It turns out that $\mathscr{C}^m_{m,n}$ has a minimal and a maximal element w.r.t. $\preccurlyeq$. Corollary 7.1.10 in connection with Lemma 7.1.11 leads to the upper and lower bounds given in Corollary 7.1.12.

Lemma 7.1.11. Let $\mathscr{R} \in \mathscr{C}^m_{m,n}$ be an arbitrary censoring scheme. Then, $\mathscr{O}_1 = (n-m, 0^{*m-1})$ is a maximal element of $\mathscr{C}^m_{m,n}$, i.e., $\mathscr{R} \preccurlyeq \mathscr{O}_1$, and $\mathscr{O}_m = (0^{*m-1}, n-m)$ is a minimal element of $\mathscr{C}^m_{m,n}$, i.e., $\mathscr{O}_m \preccurlyeq \mathscr{R}$.

Corollary 7.1.12. Let $\mathscr{R} \in \mathscr{C}^m_{m,n}$ be a censoring scheme and g be a nonnegative function. Then,

$$\frac{(n-m)!c(\mathscr{R})}{n!} Eg(X^{\mathscr{O}_m}_{m:m:n}) \le Eg(X^{\mathscr{R}}_{m:m:n}) \le \frac{c(\mathscr{R})}{n(m-1)!} Eg(X^{\mathscr{O}_1}_{m:m:n}).$$

In particular, this yields

$$\frac{(n-m)!c(\mathscr{R})}{n!} E|X^{\mathscr{O}_m}_{m:m:n}|^k \le E|X^{\mathscr{R}}_{m:m:n}|^k \le \frac{c(\mathscr{R})}{n(m-1)!} E|X^{\mathscr{O}_1}_{m:m:n}|^k.$$

It is directly seen from the joint density function (2.4) that the progressively Type-II censored order statistics $X^{\mathscr{O}_m}_{1:m:n}, \dots, X^{\mathscr{O}_m}_{m:m:n}$ have the same distribution as the first m order statistics in a sample of n, i.e., $X_{1:n}, \dots, X_{m:n}$. In view of this result, Corollary 7.1.13 leads (in some sense) to a converse result to Theorem 7.1.6.

Corollary 7.1.13. Let $\mathscr{R} \in \mathscr{C}^m_{m,n}$ be a censoring scheme and g be a nonnegative function. Then,

$$Eg(X_{m:n}) \le \frac{n!}{(n-m)!c(\mathscr{R})} Eg(X^{\mathscr{R}}_{m:m:n}).$$

In particular, this yields $E|X_{m:n}|^k \le \frac{n!}{(n-m)!c(\mathscr{R})} E|X^{\mathscr{R}}_{m:m:n}|^k$ which means that the existence of the kth moment of $X^{\mathscr{R}}_{m:m:n}$ yields the existence of the kth moment of the order statistic $X_{m:n}$.

Remark 7.1.14. Using the same technique as above, similar bounds can be established for progressively Type-II censored order statistics $X^{\mathscr{R}}_{r:m:n}$ with $1 \le r \le m$. Since Theorem 2.4.1 yields for $r \in \{1, \dots, m\}$

$$\left(X^{\mathscr{R}}_{1:m:n}, \dots, X^{\mathscr{R}}_{r:m:n}\right) \stackrel{\mathsf{d}}{=} \left(X^{\mathscr{R}_r}_{1:r:n}, \dots, X^{\mathscr{R}_r}_{r:r:n}\right)$$

and $\gamma_j(\mathscr{R}) = \gamma_j(\mathscr{R}_r)$, $1 \le j \le r$, we have a similar bound as in (7.5). In particular,

$$f^{X^{\mathscr{S}}_{1:m:n},\dots,X^{\mathscr{S}}_{r:m:n}} \le \Big(\prod_{j=1}^r \frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})}\Big) f^{X^{\mathscr{R}}_{1:m:n},\dots,X^{\mathscr{R}}_{r:m:n}}, \quad 1 \le r \le m.$$

For instance, the corresponding bounds in Corollary 7.1.10 become

$$Eg(X^{\mathscr{S}}_{r:m:n}) \le \Big(\prod_{j=1}^r \frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})}\Big) Eg(X^{\mathscr{R}}_{r:m:n}), \quad 1 \le r \le m.$$

In Corollary 7.1.13, the bounds are given by

$$Eg(X_{r:n}) \leq \frac{n!}{(n-r)! \prod_{j=1}^{r} \gamma_j(\mathscr{R})} Eg(X_{r:m:n}^{\mathscr{R}}), \quad 1 \leq r \leq m.$$

7.2 Moments for Particular Distributions

For some particular distributions, explicit expressions for (product) moments are available. In this section, we present such expressions for selected distributions. They are used, for instance, in developing optimal linear estimation (see Chap. 11).

7.2.1 Exponential Distribution

The following expressions for expectations, variances, and covariances result directly from (2.13).

Theorem 7.2.1. Let $Z_{1:m:n}, \ldots, Z_{m:m:n}$ be progressively Type-II censored order statistics based on an exponential distribution $\text{Exp}(\mu, \vartheta)$. Then:

(i) $EZ_{r:m:n} = \mu + \sum_{j=1}^{r} \frac{1}{\gamma_j}$, $1 \leq r \leq m$;
(ii) $\text{Var}(Z_{r:m:n}) = \sum_{j=1}^{r} \frac{1}{\gamma_j^2}$, $1 \leq r \leq m$;
(iii) $\text{Cov}(Z_{r:m:n}, Z_{s:m:n}) = \text{Var}(Z_{r:m:n}) = \sum_{j=1}^{r} \frac{1}{\gamma_j^2}$, $1 \leq r \leq s \leq m$.

Introducing the notation $a_r = \frac{1}{\gamma_r^2}$, $a_{\bullet r} = \sum_{j=1}^{r} a_j = \sum_{j=1}^{r} \frac{1}{\gamma_j^2}$, $1 \leq r \leq m$, Theorem 7.2.1 yields the covariance matrix

$$\text{Cov}(\mathbf{Z}^{\mathscr{R}}) = \Sigma_{\mathbf{Z}}^{\mathscr{R}} = \begin{pmatrix} a_{\bullet 1} & a_{\bullet 1} & a_{\bullet 1} & \cdots & a_{\bullet 1} \\ a_{\bullet 1} & a_{\bullet 2} & a_{\bullet 2} & \cdots & a_{\bullet 2} \\ a_{\bullet 1} & a_{\bullet 2} & a_{\bullet 2} & \cdots & a_{\bullet 3} \\ \vdots & \vdots & & \ddots & \\ a_{\bullet 1} & a_{\bullet 2} & a_{\bullet 3} & \cdots & a_{\bullet m} \end{pmatrix}. \tag{7.6}$$

The special structure of this matrix will be of interest in the derivation of best linear estimators since the inverse of this matrix can be calculated explicitly. The inverse of the matrix in (7.6) is given by (see Graybill [413, p. 187])

$$\Sigma_{\mathbf{Z}^{\mathscr{R}}}^{-1} = \begin{pmatrix} \gamma_1^2+\gamma_2^2 & -\gamma_2^2 & 0 & \cdots & \cdots & 0 \\ -\gamma_2^2 & \gamma_2^2+\gamma_3^2 & -\gamma_3^2 & 0 & \cdots & 0 \\ 0 & -\gamma_3^2 & \gamma_3^2+\gamma_4^2 & -\gamma_4^2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & -\gamma_{m-1}^2 & \gamma_{m-1}^2+\gamma_m^2 & -\gamma_m^2 \\ 0 & \cdots & \cdots & 0 & -\gamma_m^2 & \gamma_m^2 \end{pmatrix}. \tag{7.7}$$

7.2.2 Weibull Distributions

Expressions for moments of Weibull progressively Type-II censored order statistics can be directly taken from Kamps and Cramer [503] who have considered sequential order statistics based on two-parameter $\mathsf{Weibull}(\vartheta,\beta)$-distributions. The kth moment, $k\in\mathbb{N}_0$, is given by

$$EX_{i:m:n}^k = \vartheta^{k/\beta} c_{i-1}\Gamma(k/p+1)\sum_{j=1}^{i} a_{j,i}\gamma_j^{-(k/\beta+1)}.$$

Product moments can be derived from Cramer and Kamps [298]. In particular, we get for $1\le i<j\le m$

$$EX_{i:m:n}X_{j:m:n} = c_{j-1}\beta^2\vartheta^2\sum_{\ell=i+1}^{j} a_{\ell,j}^{(i)}\sum_{\nu=1}^{i} a_{\nu,i}\phi_\beta(\gamma_\ell,\gamma_\nu), \tag{7.8}$$

where

$$\phi_\beta(s,t) = \int_0^\infty\int_0^y (xy)^\beta\exp\{-sy^\beta-(t-s)x^\beta\}\,dx\,dy, \quad s,t>0. \tag{7.9}$$

For $t>s>0$, Lieblein [594] obtained a simple representation of ϕ in terms of the incomplete beta function by means of a differential equation approach, i.e.,

$$\phi_\beta(s,t) = \beta^{-2}(s(t-s))^{-(1+1/\beta)}\Gamma(2+2/\beta)\mathsf{B}_{1-s/t}(1+1/\beta,1+1/\beta),$$

where $\mathsf{B}_x(u,v)=\int_0^x z^{u-1}(1-z)^{v-1}dz$ denotes the incomplete beta function; see Balakrishnan and Cohen [92] for details. Notice that $\gamma_\ell<\gamma_\nu$ for $1\le\nu\le i, i+1\le \ell\le m$ because $\gamma_1>\cdots>\gamma_m$ and $i<j$. Therefore, $\gamma_\ell<\gamma_\nu$ holds in (7.8), and we can apply (7.9) to find an explicit representation of the product moment in (7.8).

7.2.3 Reflected Power Distribution

The following expressions for expectations, variances, and covariances result directly from the representations in Corollary 2.3.10.

Theorem 7.2.2. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics based on an RPower(q)-distribution. Then:

(i) $EX_{r:m:n} = 1 - \prod_{j=1}^{r} \frac{q\gamma_j}{1+q\gamma_j}$, $1 \le r \le m$;
(ii) $E(X_{r:m:n}^2) = 1 - 2\prod_{j=1}^{r} \frac{q\gamma_j}{1+q\gamma_j} + \prod_{j=1}^{r} \frac{q\gamma_j}{2+q\gamma_j}$, $1 \le r \le m$;
(iii) $E(X_{r:m:n}^k) = \sum_{\nu=0}^{k}(-1)^\nu \binom{k}{\nu} \prod_{j=1}^{r} \frac{q\gamma_j}{\nu+q\gamma_j}$, $1 \le r \le m$;
(iv) $\mathrm{Var}(X_{r:m:n}) = \prod_{j=1}^{r} \frac{q\gamma_j}{2+q\gamma_j} - \prod_{j=1}^{r} \frac{q^2\gamma_j^2}{(1+q\gamma_j)^2}$, $1 \le r \le m$;
(v) $\mathrm{Cov}(X_{r:m:n}, X_{s:m:n}) = \Big(\prod_{j=r+1}^{s} \frac{q\gamma_j}{1+q\gamma_j}\Big)\mathrm{Var}(X_{r:m:n})$, $1 \le r \le s \le m$.

Proof. The representation for $E(X_{r:m:n}^k)$ follows from Corollary 2.3.10 using the binomial theorem and taking expectations of every summand. The covariance follows from the identity

$$\mathrm{Cov}(X_{r:m:n}, X_{s:m:n}) = \mathrm{Cov}\Big(\prod_{j=1}^{r} U_j^{1/(q\gamma_j)}, \prod_{j=1}^{s} U_j^{1/(q\gamma_j)}\Big)$$

$$= E\Big(\prod_{j=1}^{r} U_j^{2/(q\gamma_j)} \prod_{j=r+1}^{s} U_j^{1/(q\gamma_j)}\Big) - E\Big(\prod_{j=1}^{r} U_j^{1/(q\gamma_j)}\Big) \cdot E\Big(\prod_{j=1}^{s} U_j^{1/(q\gamma_j)}\Big)$$

$$= E\Big(\prod_{j=r+1}^{s} U_j^{1/(q\gamma_j)}\Big)\Big[E\Big(\prod_{j=1}^{r} U_j^{2/(q\gamma_j)}\Big) - E^2\Big(\prod_{j=1}^{r} U_j^{1/(q\gamma_j)}\Big)\Big] \tag{7.10}$$

$$= \Big(\prod_{j=r+1}^{s} \frac{q\gamma_j}{1+q\gamma_j}\Big)\mathrm{Var}(X_{r:m:n}). \tag{7.11}$$

In particular, the second factor in (7.10) yields the expression for the variance of $X_{r:m:n}$. ◻

From Burkschat et al. [236] (see also Balakrishnan and Aggarwala [86]), we find with the notation $\alpha_\nu = 1 - b_\nu, a_\nu = \frac{d_\nu}{e_\nu}, b_\nu = \frac{c_\nu}{d_\nu}$ and

$$c_\nu = \prod_{j=1}^{\nu} \gamma_j, d_\nu = \prod_{j=1}^{\nu}(\gamma_j + 1/q), e_\nu = \prod_{j=1}^{\nu}(\gamma_j + 2/q), \quad 1 \le \nu \le m, \tag{7.12}$$

the representations

$$EX_{r:m:n} = \alpha_r, \quad \mathrm{Cov}(X_{r:m:n}, X_{s:m:n}) = (a_r - b_r)b_s.$$

From Theorem 7.2.1, we get the following covariance matrix

$$\mathrm{Cov}(\mathbf{X}^{\mathscr{R}}) = \Sigma_{\mathbf{X}^{\mathscr{R}}}$$
$$= \begin{pmatrix} (a_1-b_1)b_1 & (a_1-b_1)b_2 & (a_1-b_1)b_3 & \cdots & (a_1-b_1)b_m \\ (a_1-b_1)b_2 & (a_2-b_2)b_2 & (a_2-b_2)b_3 & \cdots & (a_2-b_2)b_m \\ (a_1-b_1)b_3 & (a_2-b_2)b_3 & (a_3-b_3)b_3 & \cdots & (a_3-b_3)b_m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (a_1-b_1)b_m & (a_2-b_2)b_m & (a_3-b_3)b_m & \cdots & (a_m-b_m)b_m \end{pmatrix}. \quad (7.13)$$

The special structure of this matrix will be of interest in the derivation of best linear estimators. The inverse is a symmetric tridiagonal matrix (cf. Graybill [413, p. 198] or Arnold et al. [58, pp. 174–175]) with

$$\begin{aligned} (\Sigma^{-1}_{\mathbf{X}^{\mathscr{R}}})_{ii} &= (\gamma_i q + 1)^2 \frac{e_i}{c_i} + \gamma_{i+1} q^2 \frac{e_{i+1}}{c_i}, \quad 1 \le i \le m-1, \\ (\Sigma^{-1}_{\mathbf{X}^{\mathscr{R}}})_{mm} &= (\gamma_m q + 1)^2 \frac{e_m}{c_m}, \\ (\Sigma^{-1}_{\mathbf{X}^{\mathscr{R}}})_{i,i+1} &= -(\gamma_{i+1} q + 1) \frac{e_{i+1}}{c_i}, \quad 1 \le i \le m-1. \end{aligned} \quad (7.14)$$

7.2.4 Uniform Distribution

Since the uniform distribution is of particular interest in many problems, we present the moments of uniform progressively Type-II censored order statistics subsequently even though they can be seen as particular case of progressively Type-II censored order statistics from a reflected power function distribution ($q = 1$).

Theorem 7.2.3. Let $U_{1:m:n}, \dots, U_{m:m:n}$ be progressively Type-II censored order statistics based on a $\mathrm{U}(0,1)$-distribution. Then:

(i) $EU_{r:m:n} = 1 - \prod_{j=1}^r \frac{\gamma_j}{1+\gamma_j}$, $1 \le r \le m$;
(ii) $E(U_{r:m:n}^2) = 1 - 2\prod_{j=1}^r \frac{\gamma_j}{1+\gamma_j} + \prod_{j=1}^r \frac{\gamma_j}{2+\gamma_j}$, $1 \le r \le m$;
(iii) $E(U_{r:m:n}^k) = \sum_{\nu=0}^k \binom{k}{\nu}(-1)^\nu \prod_{j=1}^r \frac{\gamma_j}{\nu+\gamma_j}$, $1 \le r \le m$;
(iv) $\mathrm{Var}(U_{r:m:n}) = \prod_{j=1}^r \frac{\gamma_j}{2+\gamma_j} - \prod_{j=1}^r \frac{\gamma_j^2}{(1+\gamma_j)^2}$, $1 \le r \le m$;
(v) $\mathrm{Cov}(U_{r:m:n}, U_{s:m:n}) = \left(\prod_{j=r+1}^s \frac{\gamma_j}{1+\gamma_j}\right) \mathrm{Var}(U_{r:m:n})$, $1 \le r \le s \le m$.

The inverse $\Sigma_{\mathbf{U}\mathscr{R}}^{-1}$ of the covariance matrix [see (7.13)] is a symmetric tridiagonal matrix given by

$$\begin{aligned}(\Sigma_{\mathbf{U}\mathscr{R}}^{-1})_{ii} &= (\gamma_i+1)^2\frac{e_i}{c_i} + \gamma_{i+1}\frac{e_{i+1}}{c_i}, \quad 1 \le i \le m-1\\ (\Sigma_{\mathbf{U}\mathscr{R}}^{-1})_{mm} &= (\gamma_m+1)^2\frac{e_m}{c_m}\\ (\Sigma_{\mathbf{U}\mathscr{R}}^{-1})_{i,i+1} &= -(\gamma_{i+1}+1)\frac{e_{i+1}}{c_i}, \quad 1 \le i \le m-1,\end{aligned} \tag{7.15}$$

where

$$c_\nu = \prod_{j=1}^{\nu}\gamma_j, \quad e_\nu = \prod_{j=1}^{\nu}(\gamma_j+2), \quad 1 \le \nu \le m.$$

Remark 7.2.4. For order statistics, i.e., $m = n$ and $\mathscr{R} = (0^{*m})$, the above expressions reduce to $EU_{j:n} = \frac{j}{n+1}$, $\mathrm{Var}(U_{j:n}) = \frac{j(n-j+1)}{(n+1)^2(n+2)}$, and $\mathrm{Cov}(U_{i:n}, U_{j:n}) = \frac{i(n-j+1)}{(n+1)^2(n+1)}$, $1 \le i \le j \le n$, which are well known for order statistics from a standard uniform distribution (see, for example, David and Johnson [324], David and Nagaraja [327, pp. 35/36], and Arnold et al. [58]).

7.2.5 Pareto Distribution

The following expressions for expectations, variances, and covariances result directly from the representations in Corollary 2.3.13. In order to ensure the existence of the moment $EX_{r:m:n}^k$, we have to place some conditions on r, k and the parameter of the underlying Pareto distribution.

Theorem 7.2.5. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics based on a Pareto(α)-distribution. Then:

(i) $EX_{r:m:n} = \prod_{j=1}^{r}\frac{\alpha\gamma_j}{\alpha\gamma_j-1}$, provided $\alpha\gamma_r > 1$, $1 \le r \le m$;

(ii) $E(X_{r:m:n}^k) = \prod_{j=1}^{r}\frac{\alpha\gamma_j}{\alpha\gamma_j-k}$, provided $\alpha\gamma_r > k$, $1 \le r \le m$;

(iii) $\mathrm{Var}(X_{r:m:n}) = \prod_{j=1}^{r}\frac{\alpha\gamma_j}{2+\alpha\gamma_j} - \prod_{j=1}^{r}\frac{\alpha^2\gamma_j^2}{(1+\alpha\gamma_j)^2}$, provided $\alpha\gamma_r > 2$, $1 \le r \le m$;

(iv) $\mathrm{Cov}(X_{r:m:n}, X_{s:m:n}) = \left(\prod_{j=r+1}^{s}\frac{\alpha\gamma_j}{\alpha\gamma_j-1}\right)\mathrm{Var}(X_{r:m:n})$, provided $\alpha\gamma_s > 2$, $1 \le r \le s \le m$.

Proof. First, we notice that the parameters γ_j are decreasingly ordered: $\gamma_1 > \ldots > \gamma_m$. Then, the representation for $E(X_{r:m:n}^k)$ follows directly from Corollary 2.3.13 replacing β by $-\alpha$. As in the case of the reflected power distribution [see (7.10) and (7.11)], the covariance is given by

$$\begin{aligned}
&\mathsf{Cov}(X_{r:m:n}, X_{s:m:n}) \\
&= E\Big(\prod_{j=r+1}^{s} U_j^{-1/(\alpha\gamma_j)}\Big)\left[E\Big(\prod_{j=1}^{r} U_j^{-2/(\alpha\gamma_j)}\Big) - E^2\Big(\prod_{j=1}^{r} U_j^{-1/(\alpha\gamma_j)}\Big)\right] \\
&= \Big(\prod_{j=r+1}^{s} \frac{\alpha\gamma_j}{\alpha\gamma_j - 1}\Big)\mathsf{Var}(X_{r:m:n}).
\end{aligned}$$

This proves the result. □

Using the definitions in (7.12) with β replaced by $-\alpha$, we obtain from Burkschat et al. [236] the same structure of both the covariance matrix in (7.13) and its inverse in (7.14).

7.2.6 Lomax Distribution

Using the results for the Pareto distribution, we find the following representation for the moments of a Lomax distribution with parameter $\alpha > 0$. The results can be proved using the quantile function $F^{\leftarrow}(t) = (1-t)^{-1/\alpha} - 1$, $t \in (0,1)$, and Theorem 7.2.5.

Corollary 7.2.6. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics based on a Lomax(α)-distribution. Then:

(i) $EX_{r:m:n} = \prod_{j=1}^{r} \frac{\alpha\gamma_j}{\alpha\gamma_j - 1} - 1$, provided $\alpha\gamma_r > 1$, $1 \le r \le m$;

(ii) $\mathsf{Var}(X_{r:m:n}) = \prod_{j=1}^{r} \frac{\alpha\gamma_j}{2+\alpha\gamma_j} - \prod_{j=1}^{r} \frac{\alpha^2\gamma_j^2}{(1+\alpha\gamma_j)^2}$, provided $\alpha\gamma_r > 2$, $1 \le r \le m$;

(iii) $\mathsf{Cov}(X_{r:m:n}, X_{s:m:n}) = \Big(\prod_{j=r+1}^{s} \frac{\alpha\gamma_j}{\alpha\gamma_j - 1}\Big)\mathsf{Var}(X_{r:m:n})$, provided $\alpha\gamma_s > 2$, $1 \le r \le s \le m$.

7.2.7 Extreme Value Distribution

In order to compute the Fisher information, Dahmen et al. [320] have calculated the first and second moments of progressively Type-II censored order statistics from a standard extreme value distribution. Denoting by γ Euler's constant, they obtained the following results ($1 \le s \le m$):

$$\begin{aligned}
EY_{s:m:n} &= -c_{s-1}\sum_{i=1}^{s}\frac{a_{i,s}}{\gamma_i}(\gamma + \log\gamma_i), \\
EY_{s:m:n}^2 &= \frac{1}{6}\pi^2 + c_{s-1}\sum_{i=1}^{s}\frac{a_{i,s}}{\gamma_i}\big(\gamma + \log\gamma_i\big)^2.
\end{aligned} \tag{7.16}$$

Using (2.26), the mean simplifies to

$$EY_{s:m:n} = -\gamma - c_{s-1} \sum_{i=1}^{s} \frac{a_{i,s}}{\gamma_i} \log \gamma_i \,.$$

The second-order moment can be simplified in a similar manner. An expression for higher-order moments involving integrals is available in Mann [636].

7.3 Recurrence Relations for Moments

In this section, we assume that the moments always exist. Subsequently, we use the notation $\mu^{\mathscr{R}}_{j:m:n} = EX^{\mathscr{R}}_{j:m:n}$ and $\mu^{\mathscr{R}}_{i,j:m:n} = E(X^{\mathscr{R}}_{i:m:n}X^{\mathscr{R}}_{j:m:n})$. For higher-order moments, we write $\mu^{\otimes k\mathscr{R}}_{j:m:n} = E(X^{\mathscr{R}}_{j:m:n})^k$, $k \in \mathbb{N}_0$. The censoring scheme $\mathscr{R}$ is suppressed if it is obvious.

7.3.1 General Results

From the basic relations (2.46) and (2.47) in Sect. 2.6, we find directly the following relations for single and product moments.

Corollary 7.3.1. For $1 \le r < m-1$, we have

$$\mu^{\mathscr{R}}_{r+1:m:n} = \frac{\gamma_{r+1}}{\gamma_1 - \gamma_{r+1}} \mu^{\mathscr{R}}_{r:m:n} - \frac{\gamma_1}{\gamma_1 - \gamma_{r+1}} \mu^{(R_2,\ldots,R_m)}_{r:m-1:n-R_1-1}.$$

For $1 \le r < s < m-1$, we have

$$\mu^{\mathscr{R}}_{r+1,s+1:m:n} = \frac{\gamma_{r+1} - \gamma_{s+1}}{\gamma_1 - \gamma_{r+1}} \mu^{\mathscr{R}}_{r,s+1:m:n} + \frac{\gamma_{s+1}}{\gamma_1 - \gamma_{r+1}} \mu^{\mathscr{R}}_{r,s:m:n} - \frac{\gamma_1}{\gamma_1 - \gamma_{r+1}} \mu^{(R_2,\ldots,R_m)}_{r,s:m-1:n-R_1-1}.$$

7.3.2 Results for Particular Distributions

Recurrence relations for many distributions have been established in the literature (see, e.g., Balakrishnan and Aggarwala [86]). Here, we present only a selection of them. For order statistics, a huge number of results is available. For further reading, one may refer to the reviews by Arnold and Balakrishnan [51] and Balakrishnan and Sultan [126] as well as further articles in Balakrishnan and Rao [116, 117].

Exponential Distribution

The simplest available relations for moments of exponential progressively Type-II censored order statistics are established in Balakrishnan et al. [130]. In particular, we have the following results for the kth moment $\mu_{j:m:n}^{\otimes k \mathscr{R}} = E(Z_{j:m:n}^{\mathscr{R}})^k$ of a progressively Type-II censored order statistic $Z_{j:m:n}^{\mathscr{R}}$.

Theorem 7.3.2. Let $k \in \mathbb{N}_0$. Then, the following relations hold for the kth moments of exponential progressively Type-II censored order statistics:

(i) For $1 \le m \le n$,

$$\mu_{1:m:n}^{\otimes k+1 \mathscr{R}} = \mu_{1:1:n}^{\otimes k+1 (\gamma_1 - 1)} = \frac{k+1}{n} \mu_{1:1:n}^{\otimes k (n-1)} = \frac{(k+1)!}{n^{k+1}};$$

(ii) For $2 \le j \le m$,

$$\mu_{j:m:n}^{\otimes k+1} = \mu_{j:m:n}^{\otimes k+1} = \frac{k+1}{\gamma_j} \mu_{j:m:n}^{\otimes k} + \mu_{j-1:m:n}^{\otimes k+1}. \tag{7.17}$$

Proof. The first part is a direct consequence of the relation $Z_{1:m:n}^{\mathscr{R}} \sim \mathsf{Exp}(\gamma_1)$. The recurrence relation for $j > 1$ follows from the following identity with IID standard exponential random variables $Z_1, \dots, Z_j$:

$$\begin{aligned}
\mu_{j:m:n}^{\otimes k+1} &= E\Big(\sum_{\nu=1}^{j} \frac{1}{\gamma_\nu} Z_\nu\Big)^{k+1} = E\Big(Z_{j-1:m:n} + \frac{1}{\gamma_j} Z_j\Big)^{k+1} \\
&= \sum_{\nu=0}^{k+1} \binom{k+1}{\nu} \frac{1}{\gamma_j^\nu} E Z_j^\nu E Z_{j-1:m:n}^{k+1-\nu} \\
&= E Z_{j-1:m:n}^{k+1} + \sum_{\nu=1}^{k+1} \binom{k+1}{\nu} \frac{\nu!}{\gamma_j^\nu} E Z_{j-1:m:n}^{k+1-\nu} \\
&= \mu_{j-1:m:n}^{\otimes k+1} + \frac{k+1}{\gamma_j} \sum_{\nu=1}^{k+1} \binom{k}{\nu-1} \frac{(\nu-1)!}{\gamma_j^{\nu-1}} E Z_{j-1:m:n}^{k+1-\nu} \\
&= \mu_{j-1:m:n}^{\otimes k+1} + \frac{k+1}{\gamma_j} \sum_{\nu=0}^{k} \binom{k}{\nu} \frac{\nu!}{\gamma_j^{\nu}} E Z_{j-1:m:n}^{k-\nu} \\
&= \mu_{j-1:m:n}^{\otimes k+1} + \frac{k+1}{\gamma_j} \mu_{j:m:n}^{\otimes k}
\end{aligned}$$

This proves the relation. □

Remark 7.3.3. The proof can also be carried out using integral expressions given in Aggarwala and Balakrishnan [12] (see also Balakrishnan and Aggarwala [86, Theorem 4.2]). This approach results in the relation

$$\mu_{j:m:n}^{\otimes k+1\mathscr{R}} = \frac{1}{R_j+1}\Big[(k+1)\,\mu_{j:m:n}^{\otimes k\mathscr{R}} - \gamma_{j+1}\,\mu_{j:m-1:n}^{\otimes k+1(R_1,\dots,R_{j-1},R_j+R_{j+1}+1,R_{j+2},\dots,R_m)} - \gamma_j\,\mu_{j-1:m-1:n}^{\otimes k+1(R_1,\dots,R_{j-2},R_{j-1}+R_j+1,R_{j+1},\dots,R_m)}\Big]. \quad (7.18)$$

Applying the identity (see Theorem 2.4.1)

$$\mu_{j:m:n}^{\otimes k\mathscr{R}} = \mu_{j:j:n}^{\otimes k(R_1,\dots,R_{j-1},\gamma_j-1)},$$

(7.18) can be rewritten as relation (7.17) (see also Balakrishnan et al. [130]).

The proof of Aggarwala and Balakrishnan [12] proceeds by means of the differential equation

$$(f(t) =)\ F'(t) = 1 - F(t), \quad t > 0, \quad (7.19)$$

valid for the cumulative distribution function of the standard exponential distribution. Since the method is of general interest and can be applied to other distributions as well (by utilizing the corresponding differential equation), we present the derivation of (7.17) using this method. Let $\mathbf{x}_j = (x_1,\dots,x_j)$ with $x_1 \le \dots \le x_j$, $1 \le j \le m$. First, notice that, from (2.4), we can write the joint density function of $Z_{1:m:n},\dots,Z_{j:m:n}$ as

$$f(\mathbf{x}_j) = g_{\mathscr{R}}(\mathbf{x}_{j-1})\gamma_j f(x_j)(1-F(x_j))^{\gamma_j-1}\mathbb{1}_{(x_{j-1},\infty)}(x_j), \quad (7.20)$$

where $g_{\mathscr{R}}(\mathbf{x}_{j-1}) = \prod_{i=1}^{j-1}[\gamma_i f(x_i)(1-F(x_i))^{R_i}]$, $x_1 \le \dots \le x_{j-1}$. We then obtain

$$\mu_{j:m:n}^{\otimes k+1} = \int\int_{x_{j-1}}^{\infty} \gamma_j x_j^{k+1} f(x_j)(1-F(x_j))^{\gamma_j-1}dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1})d\mathbf{x}_{j-1}.$$

Integrating the inner integral by parts and using (7.19), we have

$$= \int x_{j-1}^{k+1}(1-F(x_{j-1}))^{\gamma_j} g_{\mathscr{R}}(\mathbf{x}_{j-1})d\mathbf{x}_{j-1} + (k+1)\int\int_{x_{j-1}}^{\infty} x_j^k(1-F(x_j))^{\gamma_j}\,dx_j\; g_{\mathscr{R}}(\mathbf{x}_{j-1})d\mathbf{x}_{j-1}$$

$$= \int x_{j-1}^{k+1} f^{Z_{1:m:n},\ldots,Z_{j-1:m:n}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1}$$

$$+ \frac{k+1}{\gamma_j} \int \int_{x_{j-1}}^{\infty} x_j^k \gamma_j f(x_j)(1-F(x_j))^{\gamma_j - 1} dx_j \, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1}$$

$$= \mu_{j-1:m:n}^{\otimes k+1} + \frac{k+1}{\gamma_j} \mu_{j:m:n}^{\otimes k}.$$

The corresponding result for product moments is given in Theorem 7.3.4

Theorem 7.3.4. Let $1 \le i < j \le m$, $m \le n$. Then,

$$\mu_{i,j:m:n} = \mu_{i,j-1:m:n} + \frac{1}{\gamma_j} \mu_{i:m:n}.$$

Proof. From the sum expression for exponential progressively Type-II censored order statistics, we conclude

$$\mu_{i,j:m:n} = E\big(Z_{i:m:n} Z_{j:m:n}\big) = E\big(Z_{i:m:n} Z_{j-1:m:n} + \frac{1}{\gamma_j} Z_{i:m:n} Z_j\big)$$

$$= E\big(Z_{i:m:n} Z_{j-1:m:n}\big) + \frac{1}{\gamma_j} E Z_{i:m:n} = \mu_{i,j-1:m:n} + \frac{1}{\gamma_j} \mu_{i:m:n}.$$

Hence, we arrive at the relation for product moments of exponential progressively Type-II censored order statistics. □

Remark 7.3.5. For order statistics, i.e., $R_i = 0$, $1 \le i \le m$, the preceding recurrence relations lead to the formulas derived by Joshi [486, 488] in the case of single and product moments, respectively.

Truncated Exponential Distribution

Results for truncated exponential distributions can be found, e.g., in Aggarwala and Balakrishnan [12], Balakrishnan and Aggarwala [86], and Balakrishnan et al. [130]. The density function of a doubly truncated exponential distribution is given by

$$f(t) = \frac{\mathrm{e}^{-t}}{P-Q}, \quad Q_1 \le t \le P_1,$$

where $Q = 1 - \mathrm{e}^{-Q_1} > 0$, $P = 1 - \mathrm{e}^{-P_1} > 0$. It satisfies the differential equation

$$f(t) = F'(t) = \frac{1-P}{P-Q} + 1 - F(t) = \delta + 1 - F(t), \quad Q_1 < t < P_1, \tag{7.21}$$

where $\delta = \frac{1-P}{P-Q}$. From (7.20), we obtain by integration by parts

$$\begin{aligned}
\mu_{j:m:n}^{\circledast k+1} &= \int \int_{x_{j-1}}^{P_1} \gamma_j x_j^{k+1} f(x_j)(1-F(x_j))^{\gamma_j -1} dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\
&= \int x_{j-1}^{k+1}(1-F(x_{j-1}))^{\gamma_j} g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\
&\qquad + (k+1) \int \int_{x_{j-1}}^{P_1} x_j^k (1-F(x_j))^{\gamma_j} dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\
&= \mu_{j-1:m:n}^{\circledast k+1} + (k+1) I_{j,k,n}^{\mathscr{R}}, \quad \text{say.}
\end{aligned} \tag{7.22}$$

Notice that this identity holds for any distribution. We will use (7.22) subsequently in all the examples.

Using the differential equation (7.21), $I_{j,k,n}^{\mathscr{R}}$ can be rewritten as

$$\begin{aligned}
I_{j,k,n}^{\mathscr{R}} &= \frac{1}{\gamma_j} \int \int_{x_{j-1}}^{P_1} x_j^k \gamma_j f(x_j)(1-F(x_j))^{\gamma_j -1} dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\
&\qquad - \delta \int \int_{x_{j-1}}^{P_1} x_j^k (1-F(x_j))^{\gamma_j -1} dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\
&= \frac{1}{\gamma_j} \mu_{j:m:n}^{\circledast k} \\
&\qquad - \delta \prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} \int \int_{x_{j-1}}^{P_1} x_j^k (1-F(x_j))^{\gamma_j -1} dx_j\, g_{\mathscr{R}^*}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\
&= \frac{1}{\gamma_j} \mu_{j:m:n}^{\circledast k} - \delta \prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} I_{j,k,n-1}^{\mathscr{R}^*},
\end{aligned}$$

where $\mathscr{R}^* \in \mathscr{C}_{m,n-1}^m$ denotes a censoring scheme such that the first $j-1$ components are the same as those of $\mathscr{R}$. From the preceding calculation, we conclude for any $k \in \mathbb{N}$, $\mathscr{R}$, and $n \in \mathbb{N}$ that

$$(k+1) I_{j,k,n}^{\mathscr{R}} = \mu_{j:m:n}^{\circledast k+1 \mathscr{R}} - \mu_{j-1:m:n}^{\circledast k+1 \mathscr{R}},$$

so that

$$(k+1) I_{j,k,n-1}^{\mathscr{R}^*} = \mu_{j:m:n-1}^{\circledast k+1 \mathscr{R}^*} - \mu_{j-1:m:n-1}^{\circledast k+1 \mathscr{R}^*}.$$

Therefore, we find

$$\mu_{j:m:n}^{\otimes k+1} = \mu_{j-1:m:n}^{\otimes k+1} + \frac{k+1}{\gamma_j}\,\mu_{j:m:n}^{\otimes k} - \frac{1-P}{P-Q}\Bigg(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigg)\big(\mu_{j:m:n-1}^{\otimes k+1\mathscr{R}^*} - \mu_{j-1:m:n-1}^{\otimes k+1\mathscr{R}^*}\big).$$

Notice that, for $\gamma_j = 1$ (which can only happen for $j = m$), we get

$$\begin{aligned} I_{j,k,n}^{\mathscr{R}} &= \frac{1}{\gamma_j}\,\mu_{j:m:n}^{\otimes k} - \delta\Bigg(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigg)\int\frac{1}{k+1}\Big[P_1^{k+1} - x_{j-1}^{k+1}\Big]g_{\mathscr{R}^*}(\mathbf{x}_{j-1})d\mathbf{x}_{j-1} \\ &= \frac{1}{\gamma_j}\,\mu_{j:m:n}^{\otimes k} - \frac{\delta}{k+1}\Bigg(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigg)\Big[P_1^{k+1} - \mu_{j-1:m:n-1}^{\otimes k+1\mathscr{R}^*}\Big] \end{aligned} \tag{7.23}$$

so that

$$\mu_{j:m:n}^{\otimes k+1} = \mu_{j-1:m:n}^{\otimes k+1} + \frac{k+1}{\gamma_j}\,\mu_{j:m:n}^{\otimes k} - \frac{1-P}{P-Q}\Bigg(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigg)\big(P_1^{k+1} - \mu_{j-1:m:n-1}^{\otimes k+1\mathscr{R}^*}\big).$$

Finally, we state that for $k \in \mathbb{N}_0$,

(i) $\mu_{1:1:1}^{\otimes k+1(0)} = (k+1)\,\mu_{1:1:1}^{\otimes k(0)} - \frac{1-P}{P-Q}P_1^{k+1} + \frac{1-Q}{P-Q}Q_1^{k+1}$;
(ii) For $n \geq 2$,

$$\mu_{1:m:n}^{\otimes k+1\mathscr{R}} = \mu_{1:1:n}^{\otimes k+1(n-1)} = \tfrac{(k+1)}{n}\,\mu_{1:1:n}^{\otimes k(n-1)} - \tfrac{1-P}{P-Q}\,\mu_{1:1:n-1}^{\otimes k+1(n-2)} + \tfrac{1-Q}{P-Q}\,Q_1^{k+1}.$$

Obviously, for $P \to 1$ and $Q \to 0$, the results converge to the relations for the exponential distribution given in Theorem 7.3.2.

The above expressions generalize relations obtained for moments of order statistics by Joshi [486, 487, 488] and Balakrishnan and Joshi [103].

Truncated Pareto Distributions

The density function of a doubly truncated Pareto distribution is given by

$$f(t) = \frac{qt^{-q-1}}{P-Q}, \quad 1 \leq Q_1 \leq t \leq P_1, q > 0,$$

where $P = 1 - P_1^{-q} > 0$, $Q = 1 - Q_1^{-q} > 0$, and $1 \le Q < P$. Obviously, the quantile function is given by

$$F^{\leftarrow}(t) = Q_1 F^{\leftarrow}_{P_1/Q_1}(t), \quad t \in (0,1),$$

where F_{P_1/Q_1} denotes the cumulative distribution function of a right truncated Pareto distribution with support $[1, P_1/Q_1]$. Thus,

$$EX^k_{j:m:n} = Q_1^k EY^k_{j:m:n},$$

where $Y^k_{j:m:n}$ is a progressively Type-II censored order statistic from a F_{P_1/Q_1}-population. This shows that it is sufficient to consider right truncated Pareto distributions. However, the following approach addresses both cases simultaneously. A cumulative distribution function of a doubly truncated Pareto distribution satisfies the differential equation

$$f(t) = F'(t) = \frac{1-P}{P-Q}\frac{q}{t} + [1-F(t)]\frac{q}{t} = \frac{\delta q}{t} + [1-F(t)]\frac{q}{t}, \quad Q_1 < t < P_1,$$

where $\delta = \frac{1-P}{P-Q}$, or, equivalently,

$$\frac{t}{q}f(t) - \delta = 1 - F(t), \quad Q_1 < t < P_1. \tag{7.24}$$

As in the exponential case, we obtain from (7.22)

$$\mu^{\otimes k}_{j:m:n} = \mu^{\otimes k}_{j-1:m:n} + kI^{\mathscr{R}}_{j,k-1,n}. \tag{7.25}$$

Now, we apply the differential equation (7.24) to obtain

$$\begin{aligned} I^{\mathscr{R}}_{j,k-1,n} &= \frac{1}{q\gamma_j}\int\int_{x_{j-1}}^{P_1} x_j^k \gamma_j f(x_j)(1-F(x_j))^{\gamma_j-1}dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1})d\mathbf{x}_{j-1} \\ &\qquad - \delta\int\int_{x_{j-1}}^{P_1} x_j^{k-1}(1-F(x_j))^{\gamma_j-1}dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1})d\mathbf{x}_{j-1} \\ &= \frac{1}{q\gamma_j}\mu^{\otimes k}_{j:m:n} - \delta\Bigl(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigr)I^{\mathscr{R}^*}_{j,k-1,n-1}. \end{aligned}$$

For $\gamma_j = 1$ (which is possible only for $j = m$), this equals [see (7.23)]

$$I^{\mathscr{R}}_{j,k-1,n} = \frac{1}{q\gamma_j}\mu^{\otimes k}_{j:m:n} - \delta\Bigl(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigr)\Bigl[P_1^k - \mu^{\otimes k\mathscr{R}^*}_{j-1:m:n-1}\Bigr].$$

From (7.25), we find

$$kI^{\mathscr{R}^*}_{j,k-1,n-1} = \mu^{\otimes k\mathscr{R}^*}_{j:m:n-1} - \mu^{\otimes k\mathscr{R}^*}_{j-1:m:n-1},$$

so that

$$\mu^{\otimes k}_{j:m:n} = \mu^{\otimes k}_{j-1:m:n} + \frac{k}{q\gamma_j}\,\mu^{\otimes k}_{j:m:n} - \delta\Bigl(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigr)(\mu^{\otimes k\mathscr{R}^*}_{j:m:n-1} - \mu^{\otimes k\mathscr{R}^*}_{j-1:m:n-1}).$$

Subtracting $\frac{k}{q\gamma_j}\,\mu^{\otimes k}_{j:m:n}$ and multiplying by $\frac{q\gamma_j}{q\gamma_j-k}$ (for $q\gamma_j \neq k$) lead to

$$\mu^{\otimes k}_{j:m:n} = \frac{q\gamma_j}{q\gamma_j-k}\left\{\mu^{\otimes k}_{j-1:m:n} - \frac{1-P}{P-Q}\Bigl(\prod_{i=1}^{j-1}\frac{\gamma_i}{\gamma_i-1}\Bigr)(\mu^{\otimes k\mathscr{R}^*}_{j:m:n-1} - \mu^{\otimes k\mathscr{R}^*}_{j-1:m:n-1})\right\}.$$

In this relation, kth moments of the truncated Pareto distributions are only involved. From the recurrence relation, we observe that $\mu^{\otimes k}_{j:m:n}$ cancels out for $q\gamma_j = k$ so that the relation cannot be used for the computation of $\mu^{\otimes k}_{j:m:n}$.

Finally, we find for $k \in \mathbb{N}_0$ and $n \geq 2$ (see Balakrishnan and Aggarwala [86, p. 57/58]),

$$\mu^{\otimes k(0)}_{1:1:1} = \begin{cases}\dfrac{q}{(k-q)(P-Q)}(P_1^{k-q} - Q_1^{k-q}), & k \neq q\\[2mm] \dfrac{q}{P-Q}(\log(P_1) - \log(Q_1)), & k = q\end{cases},$$

$$\mu^{\otimes k(n-1)}_{1:1:n} = \frac{nq}{k-nq}\left\{\frac{1-P}{P-Q}\,\mu^{\otimes k(n-2)}_{1:1:n} - \frac{1-Q}{P-Q}Q_1^k\right\}.$$

From Example 7.1.8, we know that the kth moment of $X_{j:m:n}$ from an untruncated Pareto distribution exists if $\gamma_j > \frac{k}{q}$ or, equivalently, $q\gamma_j > k$. Taking the limit $P \to 1$, we find the recurrence relation

$$\mu^{\otimes k}_{j:m:n} = \frac{q\gamma_j}{q\gamma_j-k}\,\mu^{\otimes k}_{j-1:m:n},$$

which is obvious from Theorem 7.2.5.

Truncated Power Distributions

The density function of a doubly truncated power distribution is given by

$$f(t) = \frac{qt^{q-1}}{P-Q}, \quad 0 \leq Q_1 \leq t \leq P_1 \leq 1, q > 0,$$

where $P = P_1^q > 0$, $Q = Q_1^q > 0$, and $Q < P$. Its cumulative distribution function is given by

$$F(t) = \frac{t^q - Q}{P - Q}, \quad t \in (Q_1, P_1).$$

It satisfies the differential equation

$$f(t) = F'(t) = \frac{P}{P-Q}\frac{q}{t} + [1 - F(t)]\frac{q}{t} = \frac{\delta q}{t} + [1 - F(t)]\frac{q}{t}, \quad Q_1 < t < P_1,$$

where $\delta = \frac{P}{P-Q}$, or, equivalently,

$$\delta - \frac{t}{q} f(t) = 1 - F(t), \quad Q_1 < t < P_1.$$

This differential equation is quite similar to the differential equation in the Pareto case [see (7.24)]. We get

$$I_{j,k-1,n}^{\mathscr{R}} = \delta \prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} I_{j,k-1,n-1}^{\mathscr{R}*} - \frac{1}{q\gamma_j}\, \mu_{j:m:n}^{\otimes k+1}.$$

Therefore, we obtain

$$\mu_{j:m:n}^{\otimes k} = \frac{q\gamma_j}{q\gamma_j + k} \left\{ \mu_{j-1:m:n}^{\otimes k} + \frac{P}{P-Q} \left(\prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} \right) (\mu_{j:m:n-1}^{\otimes k \mathscr{R}*} - \mu_{j-1:m:n-1}^{\otimes k \mathscr{R}*}) \right\}.$$

Finally, we have for $k \in \mathbb{N}_0$,

$$\mu_{1:1:1}^{\otimes k(0)} = \frac{q}{(k+q)(P-Q)} (P_1^{k+q} - Q_1^{k+q}).$$

Truncated Reflected Power Distributions

The density function of a doubly truncated reflected power distribution is given by

$$f(t) = \frac{q(1-t)^{q-1}}{P-Q}, \quad 0 \le Q_1 \le t \le P_1 \le 1, q > 0,$$

where $P = 1-(1-P_1)^q > 0$, $Q = 1-(1-Q_1)^q > 0$, and $Q < P$. Its cumulative distribution function is given by

$$F(t) = \frac{1-Q-(1-t)^q}{P-Q}, \quad t \in (Q_1, P_1).$$

It satisfies the differential equation

$$\frac{1-t}{q} f(t) = \frac{1-P}{P-Q} + 1 - F(t) = \delta + 1 - F(t), \quad Q_1 < t < P_1, \tag{7.26}$$

where $\delta = \frac{1-P}{P-Q}$. As above, we obtain from (7.22)

$$\mu_{j:m:n}^{\otimes k} = \mu_{j-1:m:n}^{\otimes k} + kI_{j,k-1,n}^{\mathscr{R}}.$$

Now, we apply the differential equation (7.26) to obtain

$$\begin{aligned} I_{j,k-1,n}^{\mathscr{R}} &= \frac{1}{q\gamma_j} \int \int_{x_{j-1}}^{P_1} (x_j^{k-1} - x_j^k)\gamma_j f(x_j)(1-F(x_j))^{\gamma_j-1} dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\ &\qquad - \delta \int \int_{x_{j-1}}^{P_1} x_j^{k-1}(1-F(x_j))^{\gamma_j-1} dx_j\, g_{\mathscr{R}}(\mathbf{x}_{j-1}) d\mathbf{x}_{j-1} \\ &= \frac{1}{q\gamma_j}(\mu_{j:m:n}^{\otimes k-1} - \mu_{j:m:n}^{\otimes k}) - \delta \prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} I_{j,k-1,n-1}^{\mathscr{R}*}, \end{aligned}$$

As in the Pareto case, we find

$$\begin{aligned} \mu_{j:m:n}^{\otimes k} = \mu_{j-1:m:n}^{\otimes k} &+ \frac{k}{q\gamma_j}(\mu_{j:m:n}^{\otimes k-1} - \mu_{j:m:n}^{\otimes k}) \\ &- \delta \prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} (\mu_{j:m:n-1}^{\otimes k \mathscr{R}*} - \mu_{j-1:m:n-1}^{\otimes k \mathscr{R}*}), \end{aligned}$$

or, equivalently,

$$\begin{aligned} \mu_{j:m:n}^{\otimes k} = \frac{q\gamma_j}{k+q\gamma_j} \Bigg\{ \mu_{j-1:m:n}^{\otimes k} &+ \frac{k}{q\gamma_j} \mu_{j:m:n}^{\otimes k-1} \\ &- \frac{1-P}{P-Q} \prod_{i=1}^{j-1} \frac{\gamma_i}{\gamma_i - 1} (\mu_{j:m:n-1}^{\otimes k \mathscr{R}*} - \mu_{j-1:m:n-1}^{\otimes k \mathscr{R}*}) \Bigg\}. \end{aligned}$$

For $P \to 1$, this tends to the relation for the (untruncated) reflected power function distribution with support [0, 1], i.e.,

$$\mu_{j:m:n}^{\otimes k} = \frac{q\gamma_j}{k+q\gamma_j}\left\{\mu_{j-1:m:n}^{\otimes k} + \frac{k}{q\gamma_j}\,\mu_{j:m:n}^{\otimes k-1}\right\}$$
$$= \frac{q\gamma_j}{k+q\gamma_j}\,\mu_{j-1:m:n}^{\otimes k} + \frac{k}{k+q\gamma_j}\,\mu_{j:m:n}^{\otimes k-1}.$$

This shows that $\mu_{j:m:n}^{\otimes k}$ is a convex combination of $\mu_{j-1:m:n}^{\otimes k}$ and $\mu_{j:m:n}^{\otimes k-1}$. Finally, we have for $k \in \mathbb{N}_0$ and $n \in \mathbb{N}$,

$$\mu_{1:1:1}^{\otimes k(0)} = \frac{1}{P-Q}\sum_{\nu=0}^{k}(-1)^{\nu}\binom{k}{\nu}\Big[(1-Q_1)^{q+k-\nu} - (1-P_1)^{q+k-\nu}\Big],$$
$$\mu_{1:1:1}^{\otimes 0(n-1)} = 1,$$
$$\mu_{1:1:1}^{\otimes k(n-1)} = \frac{qn}{k+qn}\left\{\frac{k}{qn}\,\mu_{1:1:1}^{\otimes k-1(n-1)} - \frac{1-P}{P-Q}\,\mu_{1:1:1}^{\otimes k(n-2)} + \frac{1-Q}{P-Q}\,Q_1^k\right\}.$$

Logistic and Related Distributions

Balakrishnan et al. [151] established recurrence relations for the logistic distribution with cumulative distribution function $F(t) = 1/(1+\mathrm{e}^{-t})$, $t \in \mathbb{R}$. First, a recurrence relation for the first progressively Type-II censored order statistic, the minimum, is given by

$$\mu_{1:m:n+1}^{\otimes k+1\mathscr{R}^*} = \mu_{1:m:n}^{\otimes k+1\mathscr{R}} - \frac{k+1}{n}\,\mu_{1:m:n}^{\otimes k\mathscr{R}},$$

where $\mathscr{R}^*$ is a censoring scheme with sample size $n+1$ (see Shah [795, 796] and Balakrishnan and Sultan [126]). Notice that the minimum is independent of the progressive censoring procedure. Denoting the censoring scheme by $\mathscr{R} = (R_1, \ldots, R_m)$, Balakrishnan et al. [151] presented the following relations for progressively Type-II censored order statistics. Let $\gamma_j(n) = n - \sum_{k=1}^{j-1}(R_k+1)$, $1 \le j \le m$, and $A(n, m-1) = \prod_{i=1}^{m}\gamma_i(n)$.

According to Remark 1.1.6, a progressively Type-II censored order statistic $X_{j:m:n}^{\mathscr{R}}$ (and thus its distribution) depends only on the first $j-1$ censoring numbers $R_1, \ldots, R_{j-1}$. Therefore, the sample size and the first $j-1$ censoring numbers determine the distribution and, thus, the moments of the distribution completely. Suppose we are interested in the moment $\mu_{j:m:n}^{\otimes k+1\mathscr{R}}$. Then, we have the identity $\mu_{j:m:n}^{\otimes k+1\mathscr{R}} = \mu_{j:m:n}^{\otimes k+1\mathscr{R}^*}$ where $\mathscr{R}^* \in \mathscr{C}_{m,n}^m$ is an arbitrary censoring plan with $\mathscr{R}_{\rhd j-1}^* = (R_1, \ldots, R_{j-1})$ [cf. (1.6)] and sample size n. For $\mu_{j:m:n+1}^{\otimes k+1\mathscr{R}^*}$, the censoring scheme $\mathscr{R}^* \in \mathscr{C}_{m,n+1}^m$ has sample size $n+1$, but $\mathscr{R}^*_{\triangleright j-1} = (R_1, \ldots, R_{j-1})$ still holds. Then, we can write the recurrence relation presented in Balakrishnan et al. [151, Theorem 2.2] as follows:

$$\mu_{j:m:n+1}^{\otimes k+1\mathscr{R}^*} = \frac{A(n+1, j-1)}{(R_j+2)A(n, j-1)}\Big[(R_j+1)\,\mu_{j:m:n}^{\otimes k+1\mathscr{R}^*} - (k+1)\,\mu_{j:m:n}^{\otimes k\mathscr{R}^*} + \gamma_j(n)\,\mu_{j:m:n}^{\otimes k+1\mathscr{R}^*} - \frac{A(n, j-1)}{A(n+1, j-1)}\,\mu_{j:m:n+1}^{\otimes k+1\mathscr{R}^*} - \gamma_{j-1}(n)\,\mu_{j-1:m:n}^{\otimes k+1\mathscr{R}^*} + \frac{A(n, j-1)}{A(n+1, j-2)}\,\mu_{j-1:m:n+1}^{\otimes k+1\mathscr{R}^*}\Big].$$

Noticing that $\mu_{j:m:n+1}^{\otimes k+1\mathscr{R}^*}$ is present on both sides of the equation, we can rearrange the equation and find the result

$$\mu_{j:m:n+1}^{\otimes k+1\mathscr{R}^*} = \frac{A(n+1, j-1)}{(R_j+1)A(n, j-1)}\Big[(R_j+1)\,\mu_{j:m:n}^{\otimes k+1\mathscr{R}^*} - (k+1)\,\mu_{j:m:n}^{\otimes k\mathscr{R}^*} + \gamma_j(n)\,\mu_{j:m:n}^{\otimes k+1\mathscr{R}^*} - \gamma_{j-1}(n)\,\mu_{j-1:m:n}^{\otimes k+1\mathscr{R}^*} + \frac{A(n, j-1)}{A(n+1, j-2)}\,\mu_{j-1:m:n+1}^{\otimes k+1\mathscr{R}^*}\Big].$$

Relations for product moments are also presented in Balakrishnan et al. [151].

Remark 7.3.6. Similar results for half-logistic distribution have been established by Balakrishnan and Saleh [119] and Saran and Pande [770]. Log-logistic and generalized half-logistic distributions are discussed in Balakrishnan and Saleh [120] and Balakrishnan and Saleh [121], respectively.

Doubly Truncated Burr Distributions

Doubly truncated Burr distributions with cumulative distribution function

$$F(t) = \frac{1-Q}{P-Q} - \frac{1}{P-Q}[1+\theta t^p]^{-\nu},$$

where $\nu, p, \theta > 0$ and $Q < P$ determine the proportion of truncation on the left and right, are considered in Saran and Pushkarna [771]. Relations for kth moments as well as for product moments are presented. Moreover, it is remarked that the results are valid for many particular cases including, e.g., Lomax, log-logistic, exponential, Rayleigh, and generalized Pareto distributions. Generalized Pareto distributions have also been considered by Mahmoud et al. [629].

7.4 Moments for Symmetric Distributions

In the literature on usual order statistics, many results are available when the baseline distribution is symmetric (see, for example, Arnold et al. [58] and Balakrishnan and Sultan [126]). In this section, we present results which enable us to compute the moments of progressively Type-II censored order statistics from an arbitrary symmetric distribution when the moments of progressively Type-II censored order statistics from the corresponding folded distribution are available. Balakrishnan and Aggarwala [86] introduced the idea of progressive Type-II left withdrawal scheme in order to arrive at this result (see p. 33). Here, we use another approach which leads to simpler formulas.

For convenience, we assume that the underlying distribution is symmetric about 0. Throughout this section, let f be the density function and F be the cumulative distribution function of the population distribution. Then, the corresponding folded distribution has density function g and cumulative distribution function G defined by

$$g(t) = 2f(t), \qquad G(t) = 2F(t) - 1, \quad t > 0. \tag{7.27}$$

Our results are based on the following result of Govindarajulu [410] established for usual order statistics.

Theorem 7.4.1 (Govindarajulu [410]). Let $X_{1:n}, \ldots, X_{n:n}$ be order statistics from a population with cumulative distribution function F symmetric about 0 and $Y_{1:n}, \ldots, Y_{n:n}$ be order statistics from the corresponding folded cumulative distribution function G as given in (7.27). Then, with $\mu_{r:n}^{\otimes k} = EX_{r:n}^k$ and $\nu_{r:n}^{\otimes k} = EY_{r:n}^k$, $r \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$,

$$2^n\, \mu_{r:n}^{\otimes k} = \sum_{\ell=0}^{r-1} \binom{n}{\ell} \nu_{r-\ell:n-\ell}^{\otimes k} + (-1)^k \sum_{\ell=r}^{n} \binom{n}{\ell} \nu_{\ell-r+1:\ell}^{\otimes k}. \tag{7.28}$$

From (7.1), we know that moments of progressively Type-II censored order statistics can be expressed in terms of moments of minima. Hence, we get the identity

$$\mu_{r:m:n}^{\otimes k} = \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} \frac{a_{j,r}}{\gamma_j} EX_{1:\gamma_j}^k.$$

Applying (7.28) to the kth moment of the minimum $X_{1:\gamma_j}$, we arrive at

$$\mu_{r:m:n}^{\otimes k} = \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} \frac{a_{j,r}}{\gamma_j 2^{\gamma_j}} \Big(\nu_{1:\gamma_j}^{\otimes k} + (-1)^k \sum_{\ell=1}^{\gamma_j} \binom{\gamma_j}{\ell} \nu_{\ell:\ell}^{\otimes k} \Big). \tag{7.29}$$

This expression shows that we need the kth moments of both minima and maxima of the folded distribution only.

A similar relation holds for product moments. Govindarajulu [410] established the following relation for product moments of order statistics ($1 \le i < j \le n$):

$$2^n \mu_{i,j:n} = \sum_{\ell=0}^{i-1} \binom{n}{\ell} \nu_{i-\ell,j-\ell:n-\ell} - \sum_{\ell=i}^{j-1} \binom{n}{\ell} \nu_{\ell-i+1:\ell} \nu_{j-\ell:n-\ell} + \sum_{\ell=j}^{n} \binom{n}{\ell} \nu_{\ell-j+1,\ell-i+1:\ell}. \quad (7.30)$$

Exploiting (2.29) for the uniform distribution, we get for $1 \le i < j \le m$ and $0 < s < t < 1$,

$$f^{U_{i:m:n},U_{j:m:n}}(s,t) = \Big(\prod_{p=1}^{j} \gamma_p\Big) \sum_{p=1}^{i} \sum_{k=i+1}^{j} a_{p,i} a_{k,j}^{(i)} (1-t)^{\gamma_k-1}(1-s)^{\gamma_p-\gamma_k-1}.$$

Notice that $\gamma_p \ge \gamma_k + 1$ for $p < k$. Then, expanding $(1-s)^{\gamma_p-\gamma_k-1} = \sum_{\rho=0}^{\gamma_p-\gamma_k-1} (-1)^\rho \binom{\gamma_p-\gamma_k-1}{\rho} s^\rho$ and defining $\xi_{pk} = \Big(\prod_{q=1}^{j} \gamma_q\Big) a_{p,i} a_{k,j}^{(i)}$, we get the identity

$$\begin{aligned} f^{U_{i:m:n},U_{j:m:n}}(s,t) &= \sum_{p=1}^{i} \sum_{k=i+1}^{j} \sum_{\rho=0}^{\gamma_p-\gamma_k-1} \xi_{pk} (-1)^\rho \binom{\gamma_p-\gamma_k-1}{\rho} s^\rho (1-t)^{\gamma_k-1} \\ &= \sum_{p=1}^{i} \sum_{k=i+1}^{j} \sum_{\rho=0}^{\gamma_p-\gamma_k-1} (-1)^\rho \frac{\xi_{pk}}{\gamma_k(\gamma_k+1)} \frac{\binom{\gamma_p-\gamma_k-1}{\rho}}{\binom{\gamma_k+\rho+1}{\rho}} f^{U_{\rho+1:\gamma_k+\rho+1},U_{\rho+2:\gamma_k+\rho+1}}(s,t) \\ &= \sum_{p=1}^{i} \sum_{k=i+1}^{j} \sum_{\rho=1}^{\gamma_p-\gamma_k} (-1)^{\rho-1} \frac{\xi_{pk}}{\gamma_k(\gamma_k+1)} \frac{\binom{\gamma_p-\gamma_k-1}{\rho-1}}{\binom{\gamma_k+\rho}{\rho-1}} f^{U_{\rho:\gamma_k+\rho},U_{\rho+1:\gamma_k+\rho}}(s,t). \end{aligned}$$

Hence, we get the following relation for product moments of progressively Type-II censored order statistics in terms of order statistics from the same distribution

$$\mu_{i,j:m:n} = \sum_{p=1}^{i} \sum_{k=i+1}^{j} \sum_{\rho=1}^{\gamma_p-\gamma_k} (-1)^{\rho-1} \frac{\xi_{pk}}{\gamma_k(\gamma_k+1)} \frac{\binom{\gamma_p-\gamma_k-1}{\rho-1}}{\binom{\gamma_k+\rho}{\rho-1}} E\big(X_{\rho:\gamma_k+\rho} X_{\rho+1:\gamma_k+\rho}\big). \quad (7.31)$$

Applying the relation (7.30) to $E\big(X_{\rho:\gamma_k+\rho}X_{\rho+1:\gamma_k+\rho}\big)$, we arrive at the desired representation. It should be noted that (7.30) simplifies because $i = \rho = j-1$. In this case, we have with n replaced by $\gamma_k + \rho$

$$2^{\gamma_k+\rho}\mu_{\rho,\rho+1:\gamma_k+\rho} = \sum_{\ell=0}^{\rho-1}\binom{\gamma_k+\rho}{\ell}\nu_{\rho-\ell,\rho+1-\ell:\gamma_k+\rho-\ell} - \binom{\gamma_k+\rho}{\rho}\nu_{1:\rho}\nu_{1:\gamma_k} + \sum_{\ell=\rho+1}^{\gamma_k+\rho}\binom{\gamma_k+\rho}{\ell}\nu_{\ell-\rho,\ell-\rho+1:\ell}. \tag{7.32}$$

The relations (7.29) and (7.31) can be directly applied to calculate moments of the standard Laplace distribution (for order statistics, see Govindarajulu [411] and Kotz et al. [546, pp. 61–62]). Here, the folded distribution is a standard exponential distribution. The identity (7.29) reads in this setting as, for $k = 1$,

$$\mu_{r:m:n}^{\otimes 1} = \Big(\prod_{i=1}^{r}\gamma_i\Big)\sum_{j=1}^{r}\frac{a_{j,r}}{\gamma_j 2^{\gamma_j}}\Big(\frac{1}{\gamma_j} - \sum_{\ell=1}^{\gamma_j}\binom{\gamma_j}{\ell}\sum_{p=1}^{\ell}\frac{1}{p}\Big),$$

and for $k = 2$,

$$\mu_{r:m:n}^{\otimes 2} = \Big(\prod_{i=1}^{r}\gamma_i\Big)\sum_{j=1}^{r}\frac{a_{j,r}}{\gamma_j 2^{\gamma_j}}\Big(\frac{1}{\gamma_j^2} + \sum_{\ell=1}^{\gamma_j}\binom{\gamma_j}{\ell}\Big[\sum_{p=1}^{\ell}\frac{1}{p^2} + \Big\{\sum_{p=1}^{\ell}\frac{1}{p}\Big\}^2\Big]\Big).$$

Hence, we have explicit formulas to calculate the means and variances of Laplace progressively Type-II censored order statistics. Since the (product) moments of exponential progressively Type-II censored order statistics are explicitly available (see Theorem 7.2.1), we can simplify (7.32) and arrive at

$$2^{\gamma_k+\rho}\mu_{\rho,\rho+1:\gamma_k+\rho} = \sum_{\ell=0}^{\rho-1}\binom{\gamma_k+\rho}{\ell}\Big[\nu_{\rho-\ell:\gamma_k+\rho-\ell}^{\otimes 2} + \frac{1}{\gamma_k}\nu_{\rho-\ell:\gamma_k+\rho-\ell}^{\otimes 1}\Big] - \binom{\gamma_k+\rho}{\rho}\frac{1}{\rho\gamma_k} + \sum_{\ell=\rho+1}^{\gamma_k+\rho}\binom{\gamma_k+\rho}{\ell}\Big[\nu_{\ell-\rho:\ell}^{\otimes 2} + \frac{1}{\rho}\nu_{\ell-\rho:\ell}^{\otimes 1}\Big].$$

This enables us to calculate the covariances of Laplace progressively Type-II censored order statistics. These formulas will be utilized later on (see Sect. 11.2.4) in the derivation of best linear unbiased estimators.

Example 7.4.2. Assuming the censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$, i.e., $m = 10$ and $n = 20$, the mean and the variance–covariance matrix of progressively Type-II censored order statistics from a Laplace$(0,1)$-distribution are computed using the preceding formulas as (up to three decimals)

$$E\mathbf{X}^{\mathscr{R}} = (-2.905, -1.850, -1.332, -0.990, -0.703, \\ -0.478, -0.291, -0.126, 0.075, 0.290)',$$

$$\mathsf{Cov}(\mathbf{X}^{\mathscr{R}}) = \begin{pmatrix} 1.596 & 0.569 & 0.322 & 0.215 & 0.150 & 0.111 & 0.085 & 0.069 & 0.057 & 0.052 \\ 0.569 & 0.595 & 0.339 & 0.226 & 0.158 & 0.117 & 0.090 & 0.073 & 0.061 & 0.055 \\ 0.322 & 0.339 & 0.344 & 0.230 & 0.161 & 0.119 & 0.092 & 0.074 & 0.062 & 0.056 \\ 0.215 & 0.226 & 0.230 & 0.232 & 0.162 & 0.120 & 0.093 & 0.075 & 0.062 & 0.056 \\ 0.150 & 0.158 & 0.161 & 0.162 & 0.169 & 0.125 & 0.097 & 0.079 & 0.066 & 0.059 \\ 0.111 & 0.117 & 0.119 & 0.120 & 0.125 & 0.130 & 0.101 & 0.082 & 0.068 & 0.062 \\ 0.085 & 0.090 & 0.092 & 0.093 & 0.097 & 0.101 & 0.105 & 0.085 & 0.072 & 0.065 \\ 0.069 & 0.073 & 0.074 & 0.075 & 0.079 & 0.082 & 0.085 & 0.092 & 0.077 & 0.071 \\ 0.057 & 0.061 & 0.062 & 0.062 & 0.066 & 0.068 & 0.072 & 0.077 & 0.100 & 0.092 \\ 0.052 & 0.055 & 0.056 & 0.056 & 0.059 & 0.062 & 0.065 & 0.071 & 0.092 & 0.128 \end{pmatrix}.$$

7.5 Bounds for Moments

7.5.1 Bounds Based on the Cauchy–Schwarz Inequality

Classical distribution-free upper bounds for means of order statistics have been derived by Hartley and David [432], Gumbel [421], and Ludwig [620] by means of the Cauchy–Schwarz inequality. Reviews and more details on this topic are given in Arnold et al. [58, Sect. 5.4], Arnold and Balakrishnan [51, Chap. 3], and Rychlik [764]. The upper bound (7.33) due to Balakrishnan et al. [129] extends these results to means of progressively Type-II censored order statistics.

Theorem 7.5.1. Let $X_{r:m:n}$ be a progressively Type-II censored order statistic from a distribution function F with $EX = 0$, $\mathsf{Var}(X) = 1$ for $X \sim F$. Then,

$$EX_{r:m:n} \leq \left\{ \prod_{i=1}^{r} \gamma_i^2 \sum_{i=1}^{r} \frac{a_{i,r}}{\prod_{j=1}^{r}(\gamma_j + \gamma_i - 1)} - 1 \right\}^{1/2} = B_{\mathsf{CS}}. \tag{7.33}$$

Proof. First, notice that $\int_0^1 F^{\leftarrow}(t)dt = EX = 0$ by assumption. Then, by introducing a constant $c \in \mathbb{R}$ and by Theorem 7.1.1, the Cauchy–Schwarz inequality yields the upper bound

$$EX_{r:m:n} = \int_0^1 F^{\leftarrow}(u)\Big(f^{U_{r:m:n}}(u) - c\Big)du \\ \leq \left\{ \int_0^1 \Big(f^{U_{r:m:n}}(u)\Big)^2 du - 2c + c^2 \right\}^{1/2}.$$

Obviously, the right-hand side attains its minimum for $c = 1$. Moreover, using (2.28) and the Lagrangian interpolation formula for $f \equiv 1$, we find

$$\int_0^1 \left(f^{U_{r:m:n}}(u)\right)^2 du = \prod_{i=1}^r \gamma_i^2 \sum_{i=1}^r \sum_{j=1}^r a_{i,r} a_{j,r} \int_0^1 (1-u)^{\gamma_i+\gamma_j-2} du$$

$$= \prod_{i=1}^r \gamma_i^2 \sum_{i=1}^r a_{i,r} \sum_{j=1}^r \frac{a_{j,r}}{\gamma_i+\gamma_j-1} = \prod_{i=1}^r \gamma_i^2 \sum_{i=1}^r \frac{a_{i,r}}{\prod_{j=1}^r(\gamma_i+\gamma_j-1)}.$$

This proves the desired bound. □

In the particular case $R_1 = \ldots = R_{r-1} = R$, the upper bound for $EX_{r:m:n}$ simplifies to (cf. Kamps [498, p. 184])

$$\left\{ \frac{\prod_{i=1}^r \gamma_i^2}{\prod_{i=0}^{2r-2} (2\gamma_r - 1 + i(R+1))} \binom{2(r-1)}{r-1} - 1 \right\}^{1/2}.$$

In view of the bound (7.33), it is interesting to ask whether it can be attained. Equality in (7.33) holds if and only if $|F^{\leftarrow}(u)| = d \cdot |f^{U_{r:m:n}}(u) - 1|$, $0 < u < 1$, where d is specified by the equation

$$1 = \int_0^1 \left(F^{\leftarrow}(u)\right)^2 du = d^2 \int_0^1 \left(f^{U_{r:m:n}}(u) - 1\right)^2 du = d^2 \cdot B_{\mathrm{CS}}^2.$$

Hence, $|F^{\leftarrow}(u)| = B_{\mathrm{CS}}^{-1} \cdot |f^{U_{r:m:n}}(u) - 1|$. Noticing that

$$f^{U_{r:m:n}}(0) = \begin{cases} 0, & r \geq 2 \\ \gamma_1, & r = 1 \end{cases} \quad \text{and that}$$

$$f^{U_{r:m:n}}(1) = \begin{cases} 0, & r \leq m-1 \text{ or } r = m, \gamma_m > 1 \\ \prod_{j=1}^{n-1} \frac{\gamma_j}{\gamma_j - 1}, & r = m, \gamma_m = 1 \end{cases},$$

we deduce from the monotonicity of $F^{\leftarrow}$ that the bound is attainable iff either $r = 1$ or $r = m$, $\gamma_m = 1$ (see also Theorem 2.7.5). If $r = 1$, then the density function of $U_{1:m:n}$ is given by $f^{U_{1:m:n}}(t) = n(1-t)^{n-1}$, which is a decreasing function. Hence, the distribution function attaining the upper bound $B_{\mathrm{CS}} = (n-1)/\sqrt{2n-1}$ is given by

$$F(x) = 1 - \left(\frac{1}{n}(1 - \frac{n-1}{\sqrt{2n-1}}x)\right)^{\frac{1}{n-1}}, \quad x \in \left(-\sqrt{2n-1}, \frac{\sqrt{2n-1}}{n-1}\right).$$

For order statistics, this result was given by Hartley and David [432]. For $r = m$ and $\gamma_m = 1$, we find from Lemma 7.5.3 that $\Phi_m(t) = F^{U_{m:m:n}}(t) - t$ is a convex function. Hence, $(f^{U_{m:m:n}})'(t) \geq 0$ such that $f^{U_{m:m:n}}$ is an increasing function on $(0, 1)$. This proves that the upper bound is attained. The support of the respective distribution function F is given by the interval $\left[-B_{\mathsf{CS}}, B_{\mathsf{CS}}\left(\prod_{j=1}^{m-1} \frac{\gamma_j}{\gamma_j - 1} - 1\right)\right]$. For $r = m = n \geq 2$ and order statistics, the explicit representation

$$F(x) = \Big(\frac{1}{n}\Big(1 + \frac{n-1}{\sqrt{2n-1}}x\Big)\Big)^{\frac{1}{n-1}}, \quad x \in \Big(-\frac{\sqrt{2n-1}}{n-1}, \sqrt{2n-1}\Big)$$

results (cf. Arnold et al. [58, p. 121]).

In the proof of Theorem 7.5.1, an expression for the integrated square of the density function of some progressively Type-II censored order statistic was found. This representation is a special case of formula (7.34). Let $U^{\mathscr{R}_1}_{r_1:m_1:n_1}$ and $U^{\mathscr{R}_2}_{r_2:m_2:n_2}$ be progressively Type-II censored order statistics based on the standard uniform distribution, and let $\gamma_i(\mathscr{R}_j)$ and $a^{\mathscr{R}_j}_{i,r_j}$, $j = 1, 2$, denote the corresponding parameters. Then,

$$\begin{aligned}
&\int_0^1 f^{U^{\mathscr{R}_1}_{r_1:m_1:n_1}}(u) f^{U^{\mathscr{R}_2}_{r_2:m_2:n_2}}(u) du \\
&\quad = \prod_{i=1}^{r_1} \gamma_i(\mathscr{R}_1) \prod_{i=1}^{r_2} \gamma_i(\mathscr{R}_2) \sum_{i=1}^{r_1} a^{\mathscr{R}_1}_{i,r_1} \Big(\prod_{j=1}^{r_2} \big(\gamma_j(\mathscr{R}_2) + \gamma_i(\mathscr{R}_1) - 1\big)\Big)^{-1}. \quad (7.34)
\end{aligned}$$

7.5.2 *Bound Based on the Method of Greatest Convex Minorant*

Another concept to derive upper bounds, called the method of the greatest convex minorant, is due to Moriguti [657]. We apply this procedure subsequently in order to obtain bounds for means of order statistics. For related results for usual order statistics, we refer to David and Nagaraja [327, pp. 65–68], Balakrishnan [82], and Huang [459]. A comprehensive treatment of the method as well as many applications can be found in the book by Rychlik [764].

Let F be the underlying cumulative distribution function, μ its expectation, σ^2 its variance,

$$\varphi_r(z) = f^{U_{r:m:n}}(z) - 1, \text{ and } \Phi_r(z) = \int_0^z \varphi_r(t)\, dt = F^{U_{r:m:n}}(z) - z,\ z \in [0, 1].$$

Hence, $EX_{r:m:n} - \mu = \int_0^1 x\varphi_r(F(x))\,dF(x)$ and Φ_r fulfills the assumptions of Theorem 2 of Moriguti [657]. This yields the upper bound

$$\frac{EX_{r:m:n} - \mu}{\sigma} \leq \left\{ \int_0^1 \underline{\varphi}_r(t)^2\,dt \right\}^{1/2}, \tag{7.35}$$

where $\underline{\varphi}_r$ is the right-hand derivative of the greatest convex minorant $\underline{\Phi}_r$ of Φ_r on the interval $(0, 1)$. Subsequently, we present some preliminary results that are used to compute $\underline{\Phi}_r$. First, we prove the following lemma.

Lemma 7.5.2 (Balakrishnan et al. [129]). Let $r \geq 2$. The polynomial h_r defined by

$$h_r(t) = -\sum_{i=1}^{r} a_{i,r} \frac{\gamma_i - 1}{\gamma_i}(1-t)^{\gamma_i} = \left(\prod_{j=1}^{r} \frac{1}{\gamma_j}\right)\left[\overline{F}^{U_{r:m:n}}(t) - (1-t) f^{U_{r:m:n}}(t)\right]$$

has a unique root in the interval $(0, 1)$ if either $r \leq m-1$ or $(r = m$ and $\gamma_m > 1)$. For $\gamma_m = 1$, $h_m(t)$ is positive for every $t \in (0, 1)$.

Proof. Since $\gamma_1 > \cdots > \gamma_m \geq 1$, $\gamma_r = 1$ is possible only for $r = m$. For $\gamma_m = 1$, the function h_m simplifies for $t \in (0, 1)$ to

$$\begin{aligned} h_m(t) &= -\sum_{i=1}^{m-1} a_{i,m-1} \frac{\gamma_i - 1}{(\gamma_m - \gamma_i)\gamma_i}(1-t)^{\gamma_i} + a_{m,m} \frac{\gamma_m - 1}{\gamma_m}(1-t)^{\gamma_m} \\ &= \sum_{i=1}^{m-1} \frac{a_{i,m-1}}{\gamma_i}(1-t)^{\gamma_i} = \prod_{j=1}^{m-2} \gamma_j^{-1}(1 - F^{U_{m-1:m:n}}(t)). \end{aligned}$$

Since $F^{U_{m-1:m:n}}$ is strictly increasing on $(0, 1)$ with $F^{U_{m-1:m:n}}(0) = 0$ and $F^{U_{m-1:m:n}}(1) = 1$, the function h_m is positive for $t \in (0, 1)$ and obviously decreasing.

Suppose now that either $r < m$ or $r = m, \gamma_m > 1$. Let $g_r(t) = (1-t)^{-\gamma_r} h_r(t)$, $t \in (0, 1)$. Then, $g_r(0) = -\sum_{i=1}^r a_{i,r} \frac{\gamma_i - 1}{\gamma_i} = \sum_{i=1}^r \frac{a_{i,r}}{\gamma_i} - \sum_{i=1}^r a_{i,r} = \prod_{j=1}^r \gamma_j^{-1} > 0$ and $g_r(1) = -a_{r,r} \frac{\gamma_r - 1}{\gamma_r} < 0$. Hence, g_r has at least one root t_0 in $(0, 1)$.

Now, it is proved by induction on r that g_r has only one root in $(0, 1)$. For $r = 2$, the assertion is obvious since

$$\begin{aligned} g_2(t) &= -a_{1,2} \frac{\gamma_1 - 1}{\gamma_1}(1-t)^{\gamma_1 - \gamma_2} - a_{2,2} \frac{\gamma_2 - 1}{\gamma_2}, \\ t_0 &= 1 - \left(\frac{\gamma_1(\gamma_2 - 1)}{\gamma_2(\gamma_1 - 1)}\right)^{1/(\gamma_1 - \gamma_2)}. \end{aligned} \tag{7.36}$$

Let $m \geq r \geq 3$ and suppose that g_{r-1} has only one root in $(0, 1)$. Since $g_r'(t) = -(1-t)^{R_{r-1}} g_{r-1}(t)$, we conclude that g_r has at most one local extreme in $(0, 1)$. Since g_r has at least one root in $(0, 1)$ and $g_r(0) > 0 > g_r(1)$, it has at most one root. This proves the assertion. □

Before focusing on the greatest convex minorant, we prove another auxiliary result. Both lemmas enable us to calculate the greatest convex minorant (cf. Theorem 7.5.5).

Lemma 7.5.3. Let $2 \leq r \leq m$.

(i) If $r = m$ and $\gamma_m = 1$, then Φ_m is convex on $[0, 1]$;
(ii) Let $r < m$ or $\gamma_m > 1$, then there exists a number η_r such that Φ_r is convex on the interval $[0, \eta_r]$ and concave on the interval $[\eta_r, 1]$.

Proof. Suppose $r = m$ and that $\gamma_m = 1$. Then, it is easy to show that $\Phi_m''(t) = (1-t)^{-1} f^{U_{m-1:m:n}}(t) \geq 0$. This proves the convexity of Φ_m on the interval $[0, 1]$.

According to Corollary 2.7.4, $F^{U_{r:m:n}}$ is unimodal (see also Theorem 2.7.5). This yields the desired result. □

Remark 7.5.4. The identity $\sum_{i=1}^{r} a_{i,r}\gamma_i = 0$ utilized in the preceding proof can be generalized by applying the same argument to $v(x) = x^p$, $x \in [0, 1]$, $1 \leq p \leq r-1$. This yields

$$\sum_{i=1}^{r} a_{i,r}\gamma_i^{p-1} = 0, \quad 1 \leq p \leq r-1.$$

Theorem 7.5.5 (Balakrishnan et al. [129]). If $\gamma_m = 1$, then $\Phi_m = \underline{\Phi}_m$. Otherwise, the greatest convex minorant $\underline{\Phi}_r$ of Φ_r $(r \geq 2)$ on the interval $(0, 1)$ is given by

$$\underline{\Phi}_r(z) = \begin{cases} \Phi_r(z), & z \in [0, \xi], \\ \Phi_r(\xi)\frac{1-z}{1-\xi}, & z \in [\xi, 1], \end{cases} \tag{7.37}$$

where ξ is the unique solution of the polynomial equation

$$h_r(t) = -\sum_{i=1}^{r} a_{i,r}\frac{\gamma_i - 1}{\gamma_i}(1-t)^{\gamma_i} = 0, \quad t \in (0, 1). \tag{7.38}$$

The preceding results yield the upper bound

$$EX_{r:m:n} \le \left\{ B_{\mathsf{CS}}^2 + \prod_{j=1}^{r} \gamma_j^2 \sum_{i=1}^{r} \sum_{j=1}^{r} a_{i,r} a_{j,r} \left(\frac{(\gamma_i - 1)(\gamma_j - 1)}{\gamma_i \gamma_j (\gamma_i + \gamma_j - 1)} \right) (1-\xi)^{\gamma_i + \gamma_j - 1} \right\}^{1/2}, \quad (7.39)$$

where B_{CS} is the upper bound in (7.33).

Proof. The first assertion follows from Lemma 7.5.3, since Φ_m is convex on the interval $(0, 1)$. Suppose $r \le m - 1$ or that $r = m$ and $\gamma_m > 1$. Lemma 7.5.3 yields that Φ_r is convex on an interval $[0, \eta_r]$ and concave on $[\eta_r, 1]$, $2 \le r \le m$. Following the ideas of Moriguti [657] (see also David and Nagaraja [327, p. 65/68]), we know that the greatest convex minorant $\underline{\Phi}_r$ is initially equal to Φ_r up to a point ξ and then continues as a linear function. Hence, $\underline{\Phi}_r$ has the representation (7.37). The point ξ is the solution of the equation $\Phi_r'(\xi) = -\Phi_r(\xi)(1-\xi)$ or, equivalently, $h_r(\xi) = 0$. Lemma 7.5.2 reveals that h_r has a unique root in $(0, 1)$. This proves the second assertion.

The bound results from the right-hand derivative of the greatest convex minorant $\underline{\Phi}_r$ of Φ_r on the interval $(0, 1)$, i.e.,

$$\underline{\varphi}_r(z) = \begin{cases} \varphi_r(z), & z \in [0, \xi], \\ -\frac{\Phi_r(\xi)}{1-\xi}, & z \in [\xi, 1], \end{cases}$$

the bound given in (7.35) and some lengthy calculations (see Balakrishnan et al. [129]). □

Remark 7.5.6.

(i) For $r = 1$, Φ_1 is a concave function (see Theorem 2.7.5). Thus, Moriguti's method does not lead to an improved bound;
(ii) For order statistics, (7.38) simplifies to equation (10) of Balakrishnan [82];
(iii) For $r > 2$, the solution ξ of (7.38) has to be computed numerically. For $r = 2$, the equation can be solved explicitly as given in (7.36).

7.5.3 Further Bounds

Further bounds can be established using different distance measures. For instance, Raqab [742] considered the so-called p-norm bounds (see also Cramer et al. [309]). Choosing

$$\sigma_p = (E|X-\mu|^p)^{1/p} = \left(\int_0^1 |F^{\leftarrow}(x) - \mu|^p dx \right)^{1/p}, \quad 1 \le p < \infty,$$

bounds similar to (7.39) can be obtained in this scale unit using the method of greatest convex minorant. Results for order statistics are provided by Arnold [50] and Rychlik [763].

Many bounds for moments have been established in terms of generalized order statistics. However, most of them are only valid for m-generalized order statistics and, thus, not applicable to progressively Type-II censored order statistics with arbitrary censoring scheme $\mathscr{R}$. Kałuszka and Okolewski [492] obtained upper bounds in terms of Tsallis' entropy defined as

$$T_p(X) = E\Big(\frac{X^p - X}{p-1}\Big),$$

where $p > 0$, $p \neq 1$ is denoted as entropy index (see Tsallis [856]). They found the following bounds.

Theorem 7.5.7. Let $X_{1:m:n}, \dots, X_{m:m:n}$ be progressively Type-II censored order statistics from a cumulative distribution function F with $F(0) = 0$ and mean $\mu \in \mathbb{R}$. Moreover, suppose $EX^{p\vee 1} < \infty$ and let $c > 0$. Then, for $1 \le r \le m$:

(i) For $p > 1$,

$$EX_{r:m:n} \le \frac{1}{p}\Big(c^{p-1}T_p(X) + \frac{c^{p-1}-p}{p-1}\mu + \frac{1}{c}\eta_r(p)\Big),$$

where $\eta_r(p) = \int_0^1 [1 + (p-1)\underline{\varphi}_r(t)]^{p/(p-1)}dt$ and $\underline{\varphi}_r$ is the right-hand derivative of the greatest convex minorant $\underline{\Phi}_r$ of the cumulative distribution function $\Phi_r = F^{U_{r:m:n}}$ on the interval $(0,1)$. The bound is attained iff

$$F^{\leftarrow}(t) = \frac{1}{c}[1 + (p-1)\underline{\varphi}_r(t)]^{1/(p-1)}, \quad t \in (0,1);$$

(ii) For $0 < p < 1$,

$$EX_{r:m:n} \ge -\frac{1}{p}\Big(c^{p-1}T_p(X) + \frac{c^{p-1}-p}{p-1}\mu + \frac{1}{c}\kappa_r(p)\Big),$$

where $\kappa_r(p) = -\int_0^1 [1 - (p-1)\overline{\varphi}_r(t)]^{p/(p-1)}dt$ and $\overline{\varphi}_r$ is the right-hand derivative of the smallest concave majorant $\overline{\Phi}_r$ of the cumulative distribution function $\Phi_r = F^{U_{r:m:n}}$ on the interval $(0,1)$.The bound is attained iff

$$F^{\leftarrow}(t) = \frac{1}{c}[1 - (p-1)\overline{\varphi}_r(t)]^{1/(p-1)}, \quad t \in (0,1).$$

Rychlik [765] has discussed bounds on moments of generalized L-statistics of generalized order statistics with a special focus on L-statistics of progressively Type-II censored order statistics, i.e.,

$$L(\mathbf{a}) = \sum_{j=1}^{m} a_j X_{j:m:n}.$$

He assumed that the cumulative distribution function F has a bounded support $[\alpha, \beta]$ and obtained the following result.

Theorem 7.5.8. Let $X_{1:m:n}, \ldots, X_{m:m:n}$ be progressively Type-II censored order statistics from a cumulative distribution function F with bounded support $[\alpha, \beta]$ and mean $\mu \in [\alpha, \beta]$. Then,

$$E \sum_{j=1}^{m} a_j \frac{X_{j:m:n} - \mu}{\beta - \alpha} \leq \sum_{j=1}^{m} a_j \frac{\beta - \mu}{\beta - \alpha} - \underline{F}_{\mathbf{a}}\Big(\frac{\beta - \mu}{\beta - \alpha}\Big),$$

where $\underline{F}_{\mathbf{a}}$ denotes the greatest convex minorant of $F_{\mathbf{a}} = \sum_{j=1}^{m} a_j F^{X_{j:m:n}}$.

Conditions for a discrete distribution with at most three supporting points attaining the bound can be found in Rychlik [765]. Moreover, he showed that the upper bound can be written as

$$\alpha^* \sum_{j=1}^{m} a_j - \underline{F}_{\mathbf{a}}(\alpha^*),$$

where $0 \leq \alpha^* \leq 1$ satisfies the equation $\underline{f}_{\mathbf{a}}(\alpha) = \sum_{j=1}^{m} a_j$ and $\underline{f}_{\mathbf{a}}$ is the right-hand derivative of $\underline{F}_{\mathbf{a}}$.

Finally, Rychlik [765] pointed out that the results simplify for single progressively Type-II censored order statistics as well as for differences of progressively Type-II censored order statistics. For instance, he showed that, for $r \geq 2$ and $\gamma_r > 1$, the upper bound on the expectation of $X_{r:m:n}$ can be written as

$$E \frac{X_{r:m:n} - \mu}{\beta - \alpha} \leq \begin{cases} \frac{\mu-\alpha}{\beta-\alpha}[f_{r:m:n}(\theta) - 1], & \alpha \leq \mu \leq \beta - \theta(\beta - \alpha) \\ \frac{\beta-\mu}{\beta-\alpha} F_{r:m:n}\big(\frac{\beta-\mu}{\beta-\alpha}\big), & \beta - \theta(\beta - \alpha) \leq \mu \leq \beta \end{cases},$$

where $0 \leq \theta \leq 1$ is the unique solution of the equation $(1 - t) f_{r:m:n}(t) = 1 - F_{r:m:n}(t)$. The results for differences are based on the shapes of density functions of uniform progressively Type-II censored order statistics (see Theorem 2.7.5).

Further approaches in the derivation of bounds deal with restricted families of distributions. Cramer et al. [313] considered nonnegative distributions and established the bounds

$$\mu \leq EX_{r:m:n} \leq \Big(\prod_{j=1}^{r-1} \frac{\gamma_j}{\gamma_j - 1}\Big)\mu$$

provided $\gamma_r = 1$. For $\gamma_r > 1$, the bound is given by

$$0 \le EX_{r:m:n} \le f^{U_{r:m:n}}(\xi)\mu,$$

where ξ is the solution of (7.38). The bounds are sharp in the sense that they are attained in the limit by an appropriately chosen sequence of distributions (see Cramer et al. [313]). The respective results for order statistics are given in Papadatos [702].

Marohn [639] obtained an upper bound for the survival function of a progressively Type-II censored order statistic. In particular, for $1 \le r \le m-1$ and $t \ge F^{\leftarrow}(r/n)$, he got the expression

$$P(X_{r:m:n} \ge t) \le \begin{cases} \exp\Big\{-\frac{n\gamma_{r+1}}{4(n-\gamma_{r+1})}[-\log \overline{F}(t) + \log(1-r/n)]^2\Big\}, \\ \qquad 0 \le -\log \overline{F}(t) + \log(1-r/n) \le \frac{n-\gamma_{r+1}}{n\gamma_{r+1}\gamma_r} \\ \exp\Big\{-\frac{\gamma_r}{4}[-\log \overline{F}(t) + \log(1-r/n)]\Big\}, \\ \qquad -\log \overline{F}(t) + \log(1-r/n) \ge \frac{n-\gamma_{r+1}}{n\gamma_{r+1}\gamma_r} \end{cases}.$$

7.6 First-Order Approximations to Moments

Explicit expressions for moments can only be established for some distributions (see Sect. 7.2). For the remaining ones, numerical computations or approximations have to be used. Subsequently, we illustrate an approximation by a first-order Taylor series approximation which can also be found in Balakrishnan and Aggarwala [86]. Further details on this approach are provided by David and Nagaraja [327, Sect. 4.6]. Consider uniform progressively Type-II censored order statistics $U_{1:m:n}, \ldots, U_{m:m:n}$. Then, we know from representation (2.16) that

$$U_{r:m:n} \stackrel{\mathsf{d}}{=} 1 - \prod_{j=1}^{r} U_j^{1/\gamma_j}, \quad 1 \le r \le m,$$

where $U_1, \ldots, U_m$ are IID standard uniform random variables. From Theorem 7.2.3, we find $(1 \le r \le s \le m)$

$$EU_{r:m:n} = \Pi_r = 1 - \prod_{j=1}^{r} \frac{\gamma_j}{1+\gamma_j} = 1 - b_r,$$

$$\mathsf{Var}(U_{r:m:n}) = \prod_{j=1}^{r} \frac{\gamma_j}{2+\gamma_j} - \prod_{j=1}^{r} \frac{\gamma_j^2}{(1+\gamma_j)^2} = (a_r - b_r)b_r,$$

$$\mathsf{Cov}(U_{r:m:n}, U_{s:m:n}) = \Big(\prod_{j=r+1}^{s} \frac{\gamma_j}{1+\gamma_j}\Big)\mathsf{Var}(X_{r:m:n}) = (a_r - b_r)b_s,$$

where $a_\nu = \frac{d_\nu}{e_\nu}$, $b_\nu = \frac{c_\nu}{d_\nu}$ and

$$c_\nu = \prod_{j=1}^{\nu}\gamma_j,\ d_\nu = \prod_{j=1}^{\nu}(\gamma_j+1),\ e_\nu = \prod_{j=1}^{\nu}(\gamma_j+2), \quad 1 \le \nu \le m$$

[cf. (7.12)]. Suppose $X_{1:m:n}, \ldots, X_{m:m:n}$ are progressively Type-II censored order statistics from an arbitrary cumulative distribution function F, where F is strictly increasing, absolutely continuous, and differentiable and its density function f is positive on the support of F. Moreover, we assume that the quantile function $F^{\leftarrow}$ is twice differentiable and that it can be expanded in a Taylor series of the second-order at the point $\Pi_r = EU_{r:m:n}$, i.e.,

$$F^{\leftarrow}(t) = F^{\leftarrow}(\Pi_r) + (F^{\leftarrow})'(\Pi_r)(t-\Pi_r) + \frac{1}{2}(F^{\leftarrow})'(\xi)(t-\Pi_r)^2$$

with $\xi \in (0,1)$. Considering the quantile representation given in Theorem 2.3.6

$$X_{r:m:n} \stackrel{\mathsf{d}}{=} F^{\leftarrow}(U_{r:m:n}), \quad 1 \le r \le m,$$

and ignoring the remainder in the Taylor expansion, we find the approximations

$$\begin{aligned}
EX_{r:m:n} &\approx F^{\leftarrow}(\Pi_r),\\
\mathsf{Var}(X_{r:m:n}) &\approx \frac{1}{f^2(F^{\leftarrow}(\Pi_r))}(a_r - b_r)b_r,\\
\mathsf{Cov}(X_{r:m:n}, X_{s:m:n}) &\approx \frac{1}{f(F^{\leftarrow}(\Pi_r))f(F^{\leftarrow}(\Pi_s))}(a_r - b_r)b_s,
\end{aligned}$$

where we have used the fact that $(F^{\leftarrow})'(t) = \frac{1}{f(F^{\leftarrow}(t))}$. This yields an approximation of the covariance matrix

$$\Sigma_{\mathbf{X}^{\mathscr{R}}} \approx \Delta \Sigma_{\mathbf{U}^{\mathscr{R}}} \Delta,$$

with a diagonal matrix $\Delta^{-1} = \mathsf{diag}(f(F^{\leftarrow}(\Pi_1)), \ldots, f(F^{\leftarrow}(\Pi_m)))$.

For order statistics, this kind of expansion was given by David and Johnson [324]. They also showed that the Taylor series terms beyond the first order are of order $\frac{1}{n}$. A detailed account on approximating moments of progressively Type-II censored order statistics has been provided by Balasooriya and Saw [162].

Chapter 8
Simulation of Progressively Censored Order Statistics

Simulation of order statistics has been addressed by many authors (see, for instance, Schucany [787], Lurie and Hartley [621], Lurie and Mason [622], Ramberg and Tadikamalla [737], Horn and Schlipf [450], and Devroye [337]), who discussed the efficient generation of partial or complete sets of order statistics. Reviews on that topic are provided by Devroye [337, Chap. V], Arnold et al. [58], and Tadikamalla and Balakrishnan [831]. One general problem in generating order statistics from a cumulative distribution function F is the sorting process which may be time consuming. Thus, several approaches using quantile functions and generation of uniform or exponential order statistics have been employed, and various efficient algorithms were developed to enable fast generation of order statistics from a given distribution.

In this regard, we present here some algorithms to generate progressively censored order statistics. These procedures are used to evaluate the performance of many statistical procedures presented in Parts II and III of this book.

8.1 Generation of Progressively Type-II Censored Order Statistics

The first paper dealing with the generation of progressively Type-II censored order statistics is the one by Balakrishnan and Sandhu [122] who introduced Algorithm 8.1.3 (see also Balakrishnan and Aggarwala [86, Chap. 3]). Maybe, they proposed the most popular method to generate progressively Type-II censored order statistics from an IID sequence of random variables. However, the obvious approach to generate progressively Type-II censored order statistics is to mimic the Generation Procedure 1.1.3. Starting point of the procedure is a sample of random variables $X_1, \ldots, X_n$. Notice that Algorithm 8.1.1 works for any distributional

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_8,

assumption imposed on the sample $X_1, \dots, X_n$. In particular, the approach works for nonidentically distributed random variables as well as for dependent random variables.

Let $\mathscr{R} \in \mathscr{C}_{m,n}^m$ be a censoring scheme and $X_1, \dots, X_n$ be a sample of random variables.

Algorithm 8.1.1.

① Compute the order statistics $X_{1:n}, \dots, X_{n:n}$ for the sample $X_1, \dots, X_n$;
② Let $\mathcal{N}_1 = \{1, \dots, n\}$;
③ For i from 1 to m do
④ Let $k_i = \min \mathcal{N}_i$ and put $X_{i:m:n}^{\mathscr{R}} = X_{k_i:n}$;
⑤ Choose randomly a without-replacement sample $\mathcal{R}_i \subseteq \mathcal{N}_i \setminus \{k_i\}$ with $|\mathcal{R}_i| = R_i$;
⑥ If $i < m$, put $\mathcal{N}_{i+1} = \mathcal{N}_i \setminus (\{k_i\} \cup \mathcal{R}_i)$ and go to ④, or else stop.

The preceding simulation algorithm starts with the generation of order statistics in step ①. This, for instance, may be handled by an efficient algorithm first (see the references at the beginning of this chapter). In any case, it is important that the complete sample of order statistics $X_{1:n}, \dots, X_{n:n}$ must be available. Since n in comparison to m is usually large, this method generally will not be very efficient timewise. Alternatively, one may use the quantile representation in Theorem 2.1.1 provided that the generation of uniform order statistics with the same dependence structure is simple. This approach will be especially of interest when the calculation of the quantile function $F^{\leftarrow}$ is very time consuming or if n is small.

For an IID sample, the situation becomes simpler. Then, progressively Type-II censored order statistics can be generated directly from uniform random variables without employing order statistics and sorting algorithms. The method works by using the quantile representation (2.14) in Theorem 2.3.6, i.e.,

$$X_{j:m:n} \stackrel{\mathrm{d}}{=} F^{\leftarrow}\Big(1 - \prod_{k=1}^{j} U_k^{1/\gamma_k}\Big) \stackrel{\mathrm{d}}{=} F^{\leftarrow}(U_{j:m:n}),$$

where $U_1, \dots, U_m$ are IID standard uniform random variables. Thus, the resulting method is as follows: first generate uniform random variables, multiply them according to (2.14), and, finally, apply the quantile function $F^{\leftarrow}$. This yields Algorithm 8.1.2.

Algorithm 8.1.2.

① Generate m IID uniform random variables $U_1, \dots, U_m$;
② Compute $B_k = U^{1/\gamma_k}$, $k = 1, \dots, m$;
③ Let $V_0 = 1$; calculate $V_k = B_k V_{k-1}$, $k = 1, \dots, m$;
④ Let $U_{r:m:n} = 1 - V_r$, $r = 1, \dots, m$;
⑤ Let $X_{r:m:n} = F^{\leftarrow}(U_{r:m:n})$, $r = 1, \dots, m$.

0.0452946913	0.1333139682	0.3164322400	0.3265153709	0.3725533981
0.3822451715	0.4346257333	0.4348511608	0.5176778031	0.5515798663

Table 8.1 Simulation of $m = 10$ uniform progressively Type-II censored order statistics with censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$

0.0463525634	0.1430784995	0.3804294909	0.3952901058	0.4660967082
0.4816636179	0.5702673482	0.5706661503	0.7291429300	0.8020246876

Table 8.2 Simulation of $m = 10$ standard exponential progressively Type-II censored order statistics with censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$

This approach relies on a simple computation of the quantile function $F^{\leftarrow}$. It is especially useful when $F^{\leftarrow}$ has an explicit representation. Further, it is clear that it is sufficient to generate r uniform random variables only when only r progressively Type-II censored order statistics are required.

Algorithm 8.1.2 has been introduced by Balakrishnan and Sandhu [122] in a slightly different way. They used the fact that, for a uniform random variable U, the random variable $U^{1/t}$, $t > 0$, has a Beta$(t, 1)$-distribution. Their algorithm is given in Algorithm 8.1.3.

Algorithm 8.1.3.

① Generate m independent beta-distributed random variables $B_1, \ldots, B_m$ with $B_j \sim \text{Beta}(\gamma_j, 1)$;
② Let $V_0 = 1$; calculate $V_k = B_k V_{k-1}$, $k = 1, \ldots, m$;
③ Let $U_{r:m:n} = 1 - V_r$, $r = 1, \ldots, m$;
④ Let $X_{r:m:n} = F^{\leftarrow}(U_{r:m:n})$, $r = 1, \ldots, m$.

Example 8.1.4 (Simulated uniform progressively Type-II censored order statistics). Using Algorithm 8.1.2, we have generated a sample of $m = 10$ uniform progressively Type-II censored order statistics with censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$ from a sample of size $n = 20$ (see Table 8.1). This particular censoring scheme was applied in the simulation of progressively Type-II censored order statistics from a Laplace distribution in Balakrishnan and Aggarwala [86, p. 133].

Using Step ⑤ of Algorithm 8.1.2, we apply the quantile transformation to generate a random sample of progressively Type-II censored order statistics from several other distributions.

Example 8.1.5 (Simulated nonuniform progressively Type-II censored order statistics). Using the simulated uniform progressively Type-II censored order statistics in Example 8.1.4, we apply the quantile transform to get the following simulated data:

(i) Standard exponential data: $F^{\leftarrow}(t) = -\log(1-t)$, $t \in (0, 1)$ (see Table 8.2).

−2.401418263	−1.321901089	−0.4574989714	−0.4261310730
−0.2942277203	−0.2685458855	−0.1401228209	−0.1396042850
0.03599574918	0.1088775075		

Table 8.3 Simulation of $m = 10$ standard Laplace progressively Type-II censored order statistics with censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$

12.99290868	18.39049456	22.71250514	22.86934464	23.52886140
23.65727057	24.29938590	24.30197858	25.17997875	25.54438754

Table 8.4 Simulation of $m = 10$ Laplace progressively Type-II censored order statistics with $\mu = 25$ and $\vartheta = 5$ and with censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$

(ii) Standard Laplace data: $F^{\leftarrow}(t) = \begin{cases} -\log(2t), & 0 < t < 1/2 \\ -\log(2(1-t)), & 1/2 \le t < 1 \end{cases}$ (see Table 8.3).

(iii) Laplace data with $\mu = 25$ and $\vartheta = 5$ (see Table 8.4).

Alternatively, one can simulate progressively Type-II censored data by using independent standard exponential random variables $Z_1, \ldots, Z_m$.

Algorithm 8.1.6.

① Generate m independent standard exponential random variables $Z_1, \ldots, Z_m$;
② Calculate $Y_k = \frac{1}{\gamma_k} Z_k$, $k = 1, \ldots, m$;
③ Let $Y_0 := 0$; calculate $Z_{k:m:n} = Y_k + Z_{k-1:m:n}$, $k = 1, \ldots, m$;
④ Let $X_{r:m:n} = F^{\leftarrow}(1 - \exp\{-Z_{r:m:n}\})$, $r = 1, \ldots, m$.

A disadvantage of Algorithms 8.1.2, 8.1.3, and 8.1.6 is that in order to simulate a single uniform/exponential progressively Type-II censored order statistic one has to compute all forerunners. In some exceptional cases, this can be simplified by simulating beta variates (as is possible for order statistics). This comment applies to progressively Type-II censored order statistics with equi-balanced censoring scheme $\mathscr{R} = (R^{*m})$. Moreover, it is possible for the first r progressively Type-II censored order statistics if the corresponding censoring numbers coincide (in fact, this means that the right truncated censoring scheme is equi-balanced). In this case, one has to generate a beta variable which can be done efficiently using well-known procedures (see Devroye [337, Chap. IX.4]).

8.1.1 Generation of General Progressively Type-II Censored Order Statistics

General uniform progressively Type-II censored order statistics $U^{\mathscr{R}_{\lhd r}}_{r+1:m:n}, \ldots, U^{\mathscr{R}_{\lhd r}}_{m:m:n}$ can be interpreted as a left censored sample of uniform progressively Type-II censored order statistics $U^{\mathscr{R}}_{r+1:m:n}, \ldots, U^{\mathscr{R}}_{m:m:n}$ with censoring scheme $\mathscr{R} = (0^{*r}, R_{r+1}, \ldots, R_m) \in \mathscr{C}^m_{m,n}$ [see (1.7)]. Hence, the first general uniform progressively Type-II censored order statistic satisfies the distributional identity $U^{\mathscr{R}_{\lhd r}}_{r+1:m:n} \stackrel{d}{=} U_{r+1:n}$, where $U_{r+1:n} \sim \mathsf{Beta}(r+1, n-r)$. This yields the following method to simulate a sample of general progressively Type-II censored order statistics.

Algorithm 8.1.7.

① Generate a Beta$(n-r, r+1)$-distributed random variable B_{r+1} and $m-r-1$ independent standard uniform random variables $U_{r+2}, \ldots, U_m$;
② Compute $B_k = U^{1/\gamma_k}$, $k = r+2, \ldots, m$;
③ Let $V_{r+1} = B_{r+1}$; calculate $V_k = B_k V_{k-1}$, $k = r+2, \ldots, m$;
④ Let $U^{\mathscr{R}_{\lhd r}}_{j:m:n} = 1 - V_j$, $j = r+1, \ldots, m$;
⑤ Let $X^{\mathscr{R}_{\lhd r}}_{j:m:n} = F^{\leftarrow}(U^{\mathscr{R}_{\lhd r}}_{j:m:n})$, $j = r+1, \ldots, m$.

8.1.2 Generation of Progressively Type-II Censored Order Statistics from a One-Step Censoring Plan

Simulation of progressively Type-II censored order statistics from a one-step-censoring plan can be carried out along the following lines by using (2.15).

Algorithm 8.1.8. Suppose that the one-step censoring plan $\mathscr{O}_k$ is given, $1 \le k \le m$.

① Generate uniform order statistics $U_{1:n}, \ldots, U_{k:n}$;
② Generate independent uniform order statistics $\widetilde{U}_{1:m-k}, \ldots, \widetilde{U}_{m-k:m-k}$;
③ Define $U_{r:m:n} = U_{r:n}$, $r = 1, \ldots, k$;
④ Define $U_{r:m:n} = 1 - (1 - U_{k:n}) \cdot \widetilde{U}_{m-r+1:m-k}$, $r = k+1, \ldots, m$;
⑤ Define $X_{r:m:n} = F^{\leftarrow}(U_{r:m:n})$, $r = 1, \ldots, m$.

Algorithm 8.1.8 shows that progressively Type-II censored order statistics from a one-step censoring plan $\mathscr{O}_k$ can be generated from two independent samples of order statistics drawn from a uniform population. In particular, this shows that, in this scenario, single progressively Type-II censored order statistics can be simulated from at most two independent beta variables.

8.1.3 Simulation of Progressively Hybrid Censored Data

Using the simulation procedures for progressively Type-II censored samples presented above, progressively hybrid censored data may be simulated. For instance, Type-I progressive hybrid censored data with threshold T may be obtained by simulating progressively Type-II censored order statistics and using the relation $X_j^{(\mathsf{I})} = \min(X_{j:m:n}, T)$, $1 \le j \le m$ [see (5.1)].

For Type-II progressive hybrid censored data, the sample

$$X_{1:m+R_m:n}, \dots, X_{m:m+R_m:n}, X_{m+1:m+R_m:n}, \dots, X_{m+R_m:m+R_m:n}$$

[see (5.17)] is important. However, one generates first a progressively Type-II censored sample $X_{1:m:n}, \dots, X_{m:m:n}$. If $X_{m:m:n}$ exceeds the threshold T, then we are done and $X_{1:m:n}, \dots, X_{m:m:n}$ yields the desired sample. If $T > X_{m:m:n}$, the order statistics of the remaining R_m observations $X^*_{j:R_m} = X_{m+j:m+R_m:n}$, $1 \le j \le m$, are added to the sample subject to the constraint $X_{m+j:m+R_m:n} \le T$, $1 \le j \le R_m$. Notice that this approach assumes that the random lifetimes exceeding $X_{m:m:n}$ are available. Therefore, this simulation can be done using Algorithm 8.1.1. Alternatively, one can also use a modified version of Algorithm 8.1.2.

Algorithm 8.1.9.

① Generate m independent standard uniform random variables $U_1, \dots, U_m$;
② Compute $B_k = U^{1/\gamma_k}$, $k = 1, \dots, m$;
③ Let $V_0 = 1$; calculate $V_k = B_k V_{k-1}$, $k = 1, \dots, m$;
④ Let $U_{r:m:n} = 1 - V_r$, $r = 1, \dots, m$;
⑤ If $U_{m:m:n} > F(T)$, then $X_r^{(\mathrm{II})} = F^{\leftarrow}(U_{r:m:n})$, $r = 1, \dots, m$, and stop the algorithm;
else $\ell = m + 1$ and go to ⑥;
⑥ While $V_{\ell-1} > 1 - F(T)$ and $\ell \le R_m + m$ do

❶ generate a uniform random variable U_ℓ,
❷ define $V_\ell = V_{\ell-1} U_\ell^{1/(R_m-\ell+m+1)}$,
❸ $\ell := \ell + 1$;

⑦ Let $X_j^{(\mathrm{II})} = F^{\leftarrow}(1 - V_j)$, $j = 1, \dots, \ell$.

8.2 Progressively Type-I Censored Data

Progressively Type-I censored order statistics can be simulated following the construction presented in Procedure 1.1.7. Suppose the initially planned censoring scheme is given by $\mathscr{R}^0 = (R_1^0, \dots, R_{k-1}^0)$, and let $T_1 < \dots < T_k$ be ordered real numbers, $T_0 = -\infty, T_{k+1} = \infty$.

Algorithm 8.2.1 (Simulation of progressively Type-I censored order statistics).

① Simulate a sample $X_1, \dots, X_n$;
② Calculate the order statistics $X_{1:n} \le \dots \le X_{n:n}$;
③ Define $\mathcal{P}_j = \{\alpha | T_{j-1} \le X_{\alpha:n}(\omega) < T_j\}$, $1 \le j \le k + 1$;
④ Let $\mathcal{N}_1 = \{1, \dots, n\}$; $\ell = 0$;

⑤ Increase ℓ by 1, and let $\mathcal{Q}_\ell = \mathcal{P}_\ell \cap \mathcal{N}_\ell = \{\nu_{\ell,1}, \ldots, \nu_{\ell,s_\ell}\}$ with $\nu_{\ell,1} < \cdots < \nu_{\ell,s_\ell}$;

⑥ If $|\mathcal{N}_\ell \cap \bigcup_{\alpha=\ell+1}^{k+1} \mathcal{P}_\alpha| > R_\ell^0$, then

choose randomly a without-replacement sample $\mathcal{R}_\ell \subseteq \mathcal{N}_\ell \cap \bigcup_{\alpha=\ell+1}^{k+1} \mathcal{P}_\alpha$ with $|\mathcal{R}_\ell| = R_\ell^0$.

else

let $R_\ell = |\mathcal{N}_\ell \cap \bigcup_{\alpha=\ell+1}^{k+1} \mathcal{P}_\alpha|$, $R_\alpha = 0$, $\alpha = \ell+1, \ldots, k$, put $\mathcal{Q}_\alpha = \emptyset$, $\alpha = \ell, \ldots, k$, and go to ⑦;

⑦ Put $\mathcal{N}_{\ell+1} = \mathcal{N}_\ell \setminus (\mathcal{R}_\ell \cup \mathcal{P}_\ell)$ and go to ④;

⑧ $\mathcal{M} = \bigcup_{\ell=1}^{k} \mathcal{Q}_\ell$; If $m = |\mathcal{M}| > 0$, then

$$(X_{j:m:n}^{\mathscr{R},T})_{j=1,\ldots,m} = (X_{j:n})_{j \in \mathcal{M}}.$$

Notice that, as in the case of Algorithm 8.1.1 for progressively Type-II censored order statistics, the original sample $X_1, \ldots, X_n$ may exhibit any distributional assumption as well as any dependence structure. In the IID case, one can alternatively start with uniform random variables and use a quantile transformation at the end of the simulation. In this case, the threshold T_j has to be replaced by $F(T_j)$, $j = 1, \ldots, k$.

8.3 Progressively Type-I Interval Censored Data

Extending an algorithm of Kemp and Kemp [514], Aggarwala [11] proposed a simulation algorithm for progressively Type-I interval censored data $D_1, \ldots, D_k$. Using the notation given in Fig. 1.8, Algorithm 8.3.1 can be used to generate these data. The construction is based on the following distributional results (see the conditional probability mass function in (4.9) and the derivations in Sect. 4.2):

$$D_1 \sim \mathsf{bin}(n, F(T_1)),$$

and, for $2 \le j \le k$,

$$D_j | D_{j-1}, \ldots, D_1, R_{j-1}, \ldots, R_1 \sim \mathsf{bin}\Big(n - \sum_{i=1}^{j-1}(D_j + R_j), p_j\Big),$$

where $p_j = \frac{F(T_j) - F(T_{j-1})}{1 - F(T_{j-1})}$.

Algorithm 8.3.1. Let $\mathscr{R}^0 = (R_1^0, \dots, R_{k-1}^0)$ be the initially planned censoring scheme.

① Let $j = 0$, $x = 0$, $r = 0$;
② next j;
③ Generate a binomial random variable $D_j \sim \text{bin}(n - x - r, p_j)$;
④ Determine the censoring number $R_j = \min\{R_j^0, n - x - r - D_j\}$;
⑤ Let $x = x + D_j$, $r = r + R_j$;
⑥ Go to ②.

Chapter 9
Information Measures

9.1 Fisher Information in Progressively Type-II Censored Samples

Fisher information in order statistics and, more generally, in ordered data has been considered by many authors (see, e.g., Park [704, 705], Zheng and Gastwirth [942], Nagaraja and Abo-Eleneen [668], Park and Zheng [709], Wang and He [887], Zheng and Park [944], and Burkschat and Cramer [232]). A review has recently been provided by Zheng et al. [945]. Results on progressively Type-II censored samples will be presented in the following sections. From the definition of Fisher information and the joint density function of a progressively Type-II censored sample $\mathbf{X}^{\mathscr{R}} \sim F_\theta^{\mathbf{X}^{\mathscr{R}}}$, the Fisher information about the parameter θ can be written as

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \sum_{i=1}^{m} E\left[-\frac{\partial^2}{\partial\theta^2}\log f_\theta(X_{i:m:n}) - R_i\frac{\partial^2}{\partial\theta^2}\log(1-F_\theta(X_{i:m:n}))\right]$$

provided that all derivatives as well as the expected values exist. This formula may be used for the computation of Fisher information, but it turns out to be not so efficient in general. Therefore, we present some alternative, more convenient representations.

9.1.1 Hazard Rate Representation of Fisher Information

Under some regularity conditions, the Fisher information about a single parameter θ in a progressively Type-II censored sample $\mathbf{X}^{\mathscr{R}}$ is given by

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = E\left[-\frac{\partial^2}{\partial\theta^2}\log f_\theta^{\mathscr{R}}(\mathbf{X}^{\mathscr{R}})\right],$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_9,

where $f_\theta^{\mathscr{R}}$ denotes the density function of $\mathbf{X}^{\mathscr{R}}$ (see, for example, Lehmann and Casella [582]). The expected Fisher information in a progressively Type-II censored data has been considered in Zheng and Park [943] who obtained an expression of the Fisher information in terms of the hazard rate function which is similar to a representation for uncensored data established by Efron and Johnstone [348]. More information as well as details on this representation and regularity conditions are provided by Burkschat and Cramer [232]. Throughout this chapter, let $f_{j:m:n;\theta}$, $1 \le j \le m$, denote the marginal densities of progressively Type-II censored order statistics based on f_θ (in the multiparameter case, we shall write $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ instead of θ).

Theorem 9.1.1 (Zheng and Park [943]). Under some regularity conditions, the Fisher information about a single parameter θ in $\mathbf{X}^{\mathscr{R}}$ is given by

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \int_{-\infty}^{\infty} \left\{ \frac{\partial}{\partial\theta} \log \lambda_\theta(x) \right\}^2 \sum_{j=1}^{m} f_{j:m:n;\theta}(x)\,dx, \tag{9.1}$$

where $\lambda_\theta = f_\theta/(1-F_\theta)$ denotes the hazard rate of the sample cumulative distribution function F_θ.

Proof. The proof is based on the Markov property of progressively Type-II censored order statistics (see Sect. 2.5.1). In particular, the factorization of the density function in (2.36) yields the decomposition

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \sum_{r=1}^{m} \mathscr{I}(X_{r:m:n}|X_{r-1:m:n};\theta),$$

where $\mathscr{I}(X_{r:m:n}|X_{r-1:m:n};\theta)$ is the expected Fisher information about θ in $X_{r:m:n}$, given $X_{r-1:m:n}$, $2 \le r \le m$, and $\mathscr{I}(X_{1:m:n}|X_{0:m:n};\theta) = \mathscr{I}(X_{1:m:n};\theta)$. Park [705] has established the identity

$$\mathscr{I}(X_{1:m:n};\theta) = \int_{-\infty}^{\infty} \left\{ \frac{\partial}{\partial\theta} \log \lambda_\theta(x) \right\}^2 f_{1:m:n;\theta}(x)\,dx.$$

The distribution of $X_{r:m:n}$, given $X_{r-1:m:n} = s$, is that of a minimum from a random sample of size γ_r with truncated density function $f_\theta/(1-F_\theta(s))$. Moreover, the corresponding hazard rate is given by λ_θ. Hence, we get by a second application of Park's [705] result, for $2 \le r \le m$, the identity

$$\begin{aligned}\mathscr{I}&(X_{r:m:n}|X_{r-1:m:n};\theta)\\ &= \int_{\mathbb{R}} \int_s^{\infty} \left\{ \frac{\partial}{\partial\theta} \log \lambda_\theta(t) \right\}^2 f_{r|r-1:m:n;\theta}(t|s)dt f_{r-1:m:n;\theta}(s)ds\end{aligned}$$

and by Fubini's theorem

$$= \int_{\mathbb{R}} \left\{ \frac{\partial}{\partial \theta} \log \lambda_\theta(t) \right\}^2 \int_{-\infty}^{t} f_{r|r-1:m:n;\theta}(t|s) f_{r-1:m:n;\theta}(s) ds\, dt$$
$$= \int_{\mathbb{R}} \left\{ \frac{\partial}{\partial \theta} \log \lambda_\theta(t) \right\}^2 f_{r:m:n;\theta}(t)\, dt.$$

This proves the desired representation. □

Fisher Information in Location or Scale Family

Suppose $F_\theta, \theta \in \Theta$, forms either a location or a scale family [see also (11.1)]. Moreover, we assume that the standard member F is absolutely continuous with density function function f and satisfies Regularity Conditions 9.1.2 (see, e.g., Escobar and Meeker [356], Balakrishnan et al. [140], and Burkschat and Cramer [232]).

Regularity Condition 9.1.2. Let F_θ, $\theta \in \Theta$, be either a location or a scale family of cumulative distribution functions with density functions f_θ. The density function of the standard member is denoted by f. It satisfies the conditions:

(i) $f(x) > 0$ for all $x \in \mathbb{R}$;
(ii) $\lim_{x\to\pm\infty} x^2 f'(x) = 0$, where f' is the derivative of f;
(iii) The second derivative f'' is continuous;
(iv) The expectation $\mathscr{I}(\theta) = -\int \frac{\partial^2 \log(f_\theta(x))}{\partial \theta^2} f_\theta(x)\, dx$ is finite, where f_θ denotes the density function from the location or scale family.

Then, the Fisher information $\mathscr{I}(\mathbf{X}^{\mathscr{R}}; \theta)$ can be written as in (9.1). Notice that the Fisher information in a right censored sample of progressively Type-II censored order statistics can be obtained in the same way by replacing m by $r \le m$. Therefore,

$$\mathscr{I}(X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{r:m:n}; \theta) = \int_{-\infty}^{\infty} \left\{ \frac{\partial}{\partial \theta} \log \lambda_\theta(x) \right\}^2 \sum_{j=1}^{r} f_{j:m:n;\theta}(x)\, dx.$$

In a location or scale family, the Fisher information about the parameter θ simplifies considerably in the sense that it could be directly computed from the Fisher information of the standard member F. The proof is straightforward from (9.1). A similar result is presented in Lehmann and Casella [582, p. 118/119].

Theorem 9.1.3. Let Regularity Condition 9.1.2 be satisfied. Then:

(i) If F_θ, $\theta \in \mathbb{R}$, forms a location family, then $\mathscr{I}(\mathbf{X}^{\mathscr{R}}; \theta) = \mathscr{I}(\mathbf{X}^{\mathscr{R}}; 0)$. In particular, the Fisher information in θ does not depend on the location parameter;

(ii) If F_θ, $\theta > 0$, forms a scale family, then $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \theta^{-2}\mathscr{I}(\mathbf{X}^{\mathscr{R}};1)$. Thus, the Fisher information in θ is proportional to θ^{-2}.

The result will be quite useful in deriving the Fisher information for particular distributions. For location or scale families with standard member F and corresponding hazard rate λ, expression (9.1) also has a simpler form. For a location family, we obtain

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};0) = \int_{-\infty}^{\infty} \left\{\frac{\lambda'(x)}{\lambda(x)}\right\}^2 \sum_{j=1}^{m} f_{j:m:n;0}(x)\,dx = \sum_{i=1}^{m} E\left(\frac{\lambda'(Y_{i:m:n})}{\lambda(Y_{i:m:n})}\right)^2, \quad (9.2)$$

where $Y_{1:m:n},\ldots,Y_{m:m:n}$ are progressively Type-II censored order statistics from F. For a scale family, the corresponding formula reads

$$\begin{aligned}\mathscr{I}(\mathbf{X}^{\mathscr{R}};1) &= \int_{-\infty}^{\infty} \left\{1 + x\frac{\lambda'(x)}{\lambda(x)}\right\}^2 \sum_{j=1}^{m} f_{j:m:n;1}(x)\,dx \quad (9.3)\\ &= \sum_{i=1}^{m} E\left(1 + Y_{i:m:n}\frac{\lambda'(Y_{i:m:n})}{\lambda(Y_{i:m:n})}\right)^2.\end{aligned}$$

Fisher Information in the Multiparameter Case

In the following, we assume that $F_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$, is a family of cumulative distribution functions having a multiple parameter $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_p)$. Fisher information in a multiparameter setting with progressively Type-II censored order statistics has been addressed first by Ng et al. [689] who consider Weibull and log-normal distributions, or, by applying log-transformations, location–scale models of both extreme value and normal distributions, respectively. Beside computing it directly from the definition of the Fisher information matrix, they employed the missing information principle to calculate the Fisher information in a progressively Type-II censored sample (see Sect. 9.1.2). Abo-Eleneen [6] obtained some computational results for both two-parameter extreme value and normal distributions using an expression of the Fisher information as a weighted sum of Fisher information in minima (see also Abo-Eleneen [7]).

Dahmen et al. [320] established a multiparameter version of the hazard rate representation of Fisher information matrix. It is assumed that Regularity Condition 9.1.4 imposed on the joint density function $f_{\boldsymbol{\theta}}^{\mathscr{R}}$ of $\mathbf{X}^{\mathscr{R}}$ holds. Conditions in terms of the baseline distribution $F_{\boldsymbol{\theta}}$ are presented in Burkschat and Cramer [232].

Regularity Condition 9.1.4.

(i) The parameter space $\Theta \subseteq \mathbb{R}^p$ is open;
(ii) The support $\widetilde{A} = \{x \in \mathbb{R} : f_{\boldsymbol{\theta}}(x) > 0\}$ of $F_{\boldsymbol{\theta}}$ does not depend on $\boldsymbol{\theta}$;
(iii) For $x_1,\ldots,x_m \in \tilde{A}$ with $x_1 \leq \ldots \leq x_m$, $\boldsymbol{\theta} \in \Theta$, and $1 \leq i \leq p$, the derivative $\frac{\partial}{\partial\theta_i} f_{\boldsymbol{\theta}}^{\mathscr{R}}(\mathbf{x}_m)$ exists and is finite;

(iv) The partial derivatives $\frac{\partial^2}{\partial\theta_i\,\partial\theta_j} f_{\boldsymbol{\theta}}^{\mathscr{R}}(\mathbf{x}_m), 1 \le i, j \le p$, exist and satisfy the condition

$$\int \frac{\partial^2}{\partial\theta_i\,\partial\theta_j} f_{\boldsymbol{\theta}}^{\mathscr{R}}(\mathbf{x}_m)\,d\mathbf{x}_m = 0.$$

Let $[\![\cdot]\!]$ be defined by the matrix $[\![a]\!] = a \cdot a'$, $a \in \mathbb{R}^p$, where a' denotes the transposed vector of a. Moreover, let $\frac{\partial}{\partial\boldsymbol{\theta}} \log\lambda_{\boldsymbol{\theta}} = (\frac{\partial}{\partial\theta_i}\log\lambda_{\boldsymbol{\theta}})_{1\le i\le p}$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$, where $\lambda_{\boldsymbol{\theta}} = f_{\boldsymbol{\theta}}/(1 - F_{\boldsymbol{\theta}})$ denotes the hazard rate of $F_{\boldsymbol{\theta}}$.

Theorem 9.1.5. Let Regularity Condition 9.1.4 be satisfied. Then, the Fisher information matrix of $\mathbf{X}^{\mathscr{R}}$ can be expressed in terms of the hazard rate $\lambda_{\boldsymbol{\theta}}$ of $F_{\boldsymbol{\theta}}$, i.e.,

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\boldsymbol{\theta}) = \int_{-\infty}^{\infty} \left[\!\!\left[\frac{\partial}{\partial\boldsymbol{\theta}}\log\lambda_{\boldsymbol{\theta}}(x)\right]\!\!\right] \sum_{s=1}^{m} f_{s:m:n;\boldsymbol{\theta}}(x)\,dx,$$

i.e., for $1 \le i, j \le p$ and $\boldsymbol{\theta} \in \Theta$, the components are given by

$$\begin{aligned}\Big(\mathscr{I}(\mathbf{X}^{\mathscr{R}};\boldsymbol{\theta})\Big)_{ij} &= \mathscr{I}_{ij}(\mathbf{X}^{\mathscr{R}};\boldsymbol{\theta}) \\ &= \int_{-\infty}^{\infty} \left\{\frac{\partial}{\partial\theta_i}\log\lambda_{\boldsymbol{\theta}}(x)\right\}\left\{\frac{\partial}{\partial\theta_j}\log\lambda_{\boldsymbol{\theta}}(x)\right\} \sum_{s=1}^{m} f_{s:m:n;\boldsymbol{\theta}}(x)\,dx.\end{aligned}$$

The proof can be found in Dahmen et al. [320]. Here, it should be mentioned that a representation in terms of Fisher information in minima as in (9.16) and Remark 9.1.10 also holds in the multiparameter case. In particular, we get an analogous expression in terms of the Fisher information matrices $\mathscr{I}(X_{1:\gamma_j};\boldsymbol{\theta})$ of minima.

In location–scale families, the necessary regularity conditions ensuring the hazard rate representation of the Fisher information matrix can be formulated by analogy with Regularity Condition 9.1.2. Notice that assumptions (i)–(iii) of Regularity Conditions 9.1.2 and 9.1.6 are identical.

Regularity Condition 9.1.6. Let $F_{\mu,\vartheta}$, $(\mu,\vartheta) \in \Theta$, be a location–scale family of cumulative distribution functions with density functions $f_{\mu,\vartheta}$. The density function of the standard member is denoted by f. f satisfies the following conditions:

(i) $f(x) > 0$ for all $x \in \mathbb{R}$;
(ii) $\lim_{x\to\pm\infty} x^2 f'(x) = 0$, where f' is the derivative of f;
(iii) The second derivative f'' is continuous;
(iv) The expectations $\mathscr{I}_{ij}(\mu,\vartheta) = -\int_{-\infty}^{\infty}\Big(\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\log f_{\mu,\vartheta}(x)\Big) f_{\mu,\vartheta}(x)\,dx$ exist and are finite, $i, j \in \{1, 2\}$, where $\theta_1 = \mu$, $\theta_2 = \vartheta$ and $f_{\mu,\vartheta} = F'_{\mu,\vartheta}$.

Given the preceding assumptions, Dahmen et al. [320] established an expression for the Fisher information matrix in location–scale families. Let $f_{s:m:n;0,1}$ be the density function of a progressively Type-II censored order statistic $X_{s:m:n}$ from the standard member F. Then, the following representations holds. It follows directly from the identity $f_{s:m:n;\mu,\vartheta}(x) = \frac{1}{\vartheta} \cdot f_{s:m:n;0,1}\big((x-\mu)/\vartheta\big)$, $x \in \mathbb{R}$, $1 \le s \le m$, and Theorem 9.1.5 [see also (9.2) and (9.3)].

Theorem 9.1.7. Let Regularity Condition 9.1.6 be satisfied, $g = \sum_{s=1}^{m} f_{s:m:n;0,1}$, and $\lambda = f/(1-F)$ denote the hazard rate of F. Then,

$$\mathscr{I}_{11}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) = \vartheta^2 \cdot \mathscr{I}_{11}(\mathbf{X}^{\mathscr{R}};0,1) = \vartheta^2 \int_{-\infty}^{\infty} \left\{\frac{\lambda'(x)}{\lambda(x)}\right\}^2 g(x)\,dx,$$

$$\mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) = \frac{1}{\vartheta^2} \cdot \mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}};0,1) = \frac{1}{\vartheta^2} \int_{-\infty}^{\infty} \left\{1 + x\frac{\lambda'(x)}{\lambda(x)}\right\}^2 g(x)\,dx, \quad (9.4)$$

$$\mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) = \mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}};0,1) = -\int_{-\infty}^{\infty} \frac{\lambda'(x)}{\lambda(x)} \left\{1 + x\frac{\lambda'(x)}{\lambda(x)}\right\} g(x)\,dx.$$

The expressions presented in Theorem 9.1.7 can also be directly applied in both location and scale models [see (9.2) and (9.3)] since $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\mu) = \mathscr{I}_{11}(\mathbf{X}^{\mathscr{R}};\mu,1)$ and $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\vartheta) = \mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}};0,\vartheta)$, respectively.

9.1.2 Fisher Information via Missing Information Principle

As an alternative to the above approaches, Ng et al. [688] calculated the expected Fisher information via missing information principle (see Louis [618], Tanner [838]), i.e., they used the relation

Observed information = Complete information − Missing information.

They interpreted the observed progressively Type-II censored order statistics $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ as observed information. The random vector $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_m)$ is defined as the vector of progressively censored random variables, where $\mathbf{W}_j = (W_{j1}, \ldots, W_{jR_j})$ denotes those random variables corresponding to units withdrawn in the jth step of the progressive censoring procedure. It can be thought of as missing data. Merging $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{W}$ yields the complete data set $\mathbf{X}$. Ng et al. [688] established the conditional density function of $\mathbf{W}_j$, given $\mathbf{X}^{\mathscr{R}}$.

Theorem 9.1.8. Given $\mathbf{X}_j^{\mathscr{R}} = \mathbf{x}_j = (x_1, \ldots, x_j)$, the conditional density function of W_{jk}, $k \in \{1, \ldots, R_j\}$, is given by

$$f^{W_{jk}|\mathbf{X}_j^{\mathscr{R}}}(w|\mathbf{x}_j) = f^{W_{jk}|X_{j:m:n}}(w|x_j) = \frac{f(w)}{1-F(x_j)}, \quad w > x_j, \quad (9.5)$$

and W_{jk} and $W_{j\ell}$, $k \neq \ell$, are conditionally independent given $X_{j:m:n} = x_j$.

Proof. In order to calculate the desired density function, we apply an expression similar to (10.6) obtained in the derivation of the joint density function of progressively Type-II censored order statistics. Suppose the first m random variables in the original sample $X_1, \ldots, X_m$ correspond to the progressively Type-II censored order statistics. The remaining ones are partitioned as described above. Let $\mathbf{w}_j \in \mathbb{R}^{R_j}$ with $\min \mathbf{w}_j > x_j$. Evaluating the probability

$$P\Big(X_r \le \min\{X_{r+1}, \ldots, X_m, \mathbf{W}_r, \ldots, \mathbf{W}_m\}, X_r \le x_r, \mathbf{W}_r \le \mathbf{w}_r, 1 \le r \le m\Big),$$

we get, by arguments similar to those in the derivation of (10.6), the expression

$$\int_{-\infty}^{x_1}\int_{t_1}^{x_2}\cdots\int_{t_{m-1}}^{x_m} P\left(t_r \le \min\{\mathbf{W}_r, \ldots, \mathbf{W}_m\}, \mathbf{W}_r \le \mathbf{w}_r, 1 \le r \le m\right) \times \prod_{j=1}^{m} f(t_j)\, dt_m \ldots dt_1. \quad (9.6)$$

The integrand in this integral can be written as

$$\begin{aligned} &P\left(t_r \le \min\{\mathbf{W}_r, \ldots, \mathbf{W}_m\}, \mathbf{W}_r \le \mathbf{w}_r,\ 1 \le r \le m\right) \\ &= \prod_{r=1}^{m} P\left(t_r \le \min\{\mathbf{W}_r\}, \mathbf{W}_r \le \mathbf{w}_r\right) \\ &= \prod_{r=1}^{m}\prod_{k=1}^{R_j} P\left(t_r \le W_{rk} \le w_{rk}\right) \\ &= \prod_{r=1}^{m}\prod_{k=1}^{R_j} [F(w_{rk}) - F(t_r)], \quad t_1 \le \cdots \le t_m. \end{aligned}$$

Differentiating (9.6) w.r.t. $x_1, \ldots, x_m$ and w_{rk}, $1 \le r \le m$, $1 \le k \le R_r$, and taking into account that $X_1, \ldots, X_m$ are a particular assignment of the progressively Type-II censored order statistics $\mathbf{X}^{\mathscr{R}}$, we arrive at the joint density function of $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{W}$ as

$$f^{\mathbf{X}^{\mathscr{R}},\mathbf{W}}(\mathbf{x}, \mathbf{w}) = \prod_{r=1}^{m}\left[\gamma_r f(x_r) \prod_{k=1}^{R_r} f(w_{rk})\right], \quad \min \mathbf{w}_r > x_r, 1 \le r \le m.$$

From Corollary 2.1.3, we have the density function of $\mathbf{X}^{\mathscr{R}}$ as

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}) = \prod_{r=1}^{m}\left[\gamma_r f(x_r)(1 - F(x_r))^{R_r}\right].$$

Therefore, for $x_1 < \cdots < x_r$, the conditional density function of $\mathbf{W}$, given $\mathbf{X}^{\mathscr{R}}$, is given by

$$f^{\mathbf{W}|\mathbf{X}^{\mathscr{R}}}(\mathbf{w}|\mathbf{x}) = \prod_{r=1}^{m}\prod_{k=1}^{R_r}\frac{f(w_{rk})}{1-F(x_r)} = \prod_{r=1}^{m}\prod_{k=1}^{R_r} f^{W_{rk}|X_{r:m:n}}(w_{rk}|x_r),$$
$$\min \mathbf{w}_r > x_r, 1 \le r \le m. \quad (9.7)$$

Notice that the γ_r's cancel out. Hence, we deduce from the factorization theorem that, given $\mathbf{X}^{\mathscr{R}} = x$, W_{rk}, $1 \le k \le R_r$, $1 \le r \le m$, are conditionally independent with density function as in (9.5). □

Representing the observed information, complete information, and missing information by $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta)$, $\mathscr{I}((\mathbf{X}^{\mathscr{R}},\mathbf{W});\theta)$, and $\mathscr{I}(\mathbf{W}|\mathbf{X}^{\mathscr{R}};\theta)$, respectively, we relate these quantities to each other via the missing information principle. The complete information is given by

$$\mathscr{I}((\mathbf{X}^{\mathscr{R}},\mathbf{W});\theta) = -E\Big[\frac{\partial^2}{\partial\theta^2}\log f^{\mathbf{X}^{\mathscr{R}},\mathbf{W}}(\mathbf{X}^{\mathscr{R}},\mathbf{W})\Big] = \mathscr{I}(\mathbf{X};\theta) = n\mathscr{I}(X_1;\theta).$$

Defining the Fisher information about θ in a random variable W progressively censored at the time of the jth failure, given $X_{j:m:n}$, by

$$\mathscr{I}^{(j)}(W|X_{j:m:n};\theta) = -E\Big[\frac{\partial^2}{\partial\theta^2}\log f_\theta^{W|X_{j:m:n}}(W|X_{j:m:n})\Big], \quad (9.8)$$

we get the expected Fisher information of $\mathbf{W}$, given $\mathbf{X}^{\mathscr{R}}$, as

$$\mathscr{I}(\mathbf{W}|\mathbf{X}^{\mathscr{R}};\theta) = \sum_{j=1}^{m} R_j\mathscr{I}^{(j)}(\mathbf{W}_j|X_{j:m:n};\theta).$$

Hence, the expected Fisher information can be written as

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \mathscr{I}(\mathbf{X};\theta) - \sum_{j=1}^{m} R_j\mathscr{I}^{(j)}(\mathbf{W}_j|X_{j:m:n};\theta). \quad (9.9)$$

This relation will be used while computing maximum likelihood estimates of the parameters via the EM-algorithm.

9.1.3 Fisher Information for Particular Distributions

Invariance of Fisher Information Under Progressive Censoring

Suppose F_θ is a cumulative distribution function from a one-parameter exponential family as given in (2.12). Then,

$$-\frac{\partial^2}{\partial\eta^2}\log(1-F_\eta(x)) = 0, \quad -\frac{\partial^2}{\partial\eta^2}\log f_\eta(x) = \eta^{-2}.$$

Reparametrization transforms the Fisher information to

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \mathscr{I}(\mathbf{X}^{\mathscr{R}};\eta)[\eta'(\theta)]^2 \tag{9.10}$$

(see Lehmann and Casella [582, p. 115]), so that

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \frac{m}{\eta^2(\theta)}[\eta'(\theta)]^2. \tag{9.11}$$

Hence, for all F_θ of the form (2.12), the Fisher information in the progressively censored sample does not depend on the censoring scheme $\mathscr{R}$.

Remark 9.1.9. The result (9.11) can also be deduced using the following argument. Let F be a cumulative distribution function as in (2.12) with $\eta(\theta) = 1/\theta$. Then, the joint density function is

$$f_\theta^{\mathscr{R}}(\mathbf{x}) = c_{m-1}\theta^{-m}\exp\Big\{-m\frac{\widehat{\theta}}{\theta}\Big\}$$

with $\widehat{\theta} = \frac{1}{m}\sum_{j=1}^m \gamma_j(d(X_{j:m:n}) - d(X_{j-1:m:n}))$, $d(X_{0:m:n}) = 0$. Obviously, $\widehat{\theta}$ forms a sufficient statistic with $E\widehat{\theta} = \theta$ and $2m\widehat{\theta}/\theta \sim \chi^2(2m)$ (see Remark 2.3.4 and Theorem 12.1.1). Then, $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \frac{1}{\mathsf{Var}_\theta(\widehat{\theta})} = \frac{m}{\theta^2}$ (cf. Lehmann and Casella [582, p. 116, Theorem 5.4]), which is independent of the censoring scheme.

Weibull Distribution (Shape) and Extreme Value Distribution (Scale)

Using the transformation property of Fisher information in (9.10), it becomes clear that the Fisher information in the shape parameter of a Weibull distribution and that in the scale parameter of an extreme value distribution are related. Suppose $\mathbf{X}^{\mathscr{R}}$ forms a progressively Type-II censored sample from a Weibull$(1,\beta)$-distribution. Then, $Y_{j:m:n} = \log X_{j:m:n}$, $1 \le j \le m$, is a progressively Type-II censored sample from an extreme value distribution with scale parameter β. Then, $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\beta) = \mathscr{I}(\mathbf{Y}^{\mathscr{R}};1/\beta)$. Using (9.10) and $\beta = 1/\vartheta$, we find

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\beta) = \mathscr{I}(\mathbf{Y}^{\mathscr{R}};1/\beta) = \mathscr{I}(\mathbf{Y}^{\mathscr{R}};\vartheta)\frac{1}{\beta^4}.$$

Notice that we have used the parametrization $F_\vartheta(t) = 1 - \mathrm{e}^{-\mathrm{e}^{t/\vartheta}}$, $t \in \mathbb{R}$, of the extreme value distribution which is different than that used in Balakrishnan et al. [140] and Dahmen et al. [320]. Therefore, it is sufficient to calculate the Fisher information for the extreme value distribution. Using (9.4) and $\lambda(x) = \mathrm{e}^x$, we get from (9.3)

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\vartheta) = \vartheta^2 \sum_{r=1}^{m} E(1 + Y_{r:m:n})^2.$$

This expression shows that, for fixed sample size n, the Fisher information is increasing w.r.t. the number of observed failures. Moreover, the expectation can be directly evaluated. As given in Dahmen et al. [320], we get the expression

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\vartheta) = \frac{m\pi^2\vartheta^2}{6} + \vartheta^2 \sum_{s=1}^{m} c_{s-1} \sum_{i=1}^{s} \frac{a_{i,s}}{\gamma_i}(\gamma - 1 + \log\gamma_i)^2 \tag{9.12}$$

(see also Cramer and Ensenbach [293]). Different expressions are given in Balakrishnan et al. [140] and Ng et al. [689] (see also Burkschat and Cramer [232, in terms of generalized order statistics]). Expression (9.12) is useful to establish a representation for the Fisher information included in the first k order statistics $X_{1:n}, \ldots, X_{k:n}$. After some rearrangements, we obtain the expression ($1 \le k \le n-1$)

$$\begin{aligned}
&\mathscr{I}(X_{1:n}, \ldots, X_{k:n};\vartheta)\\
&= \frac{k\pi^2\vartheta^2}{6} + n\vartheta^2 \sum_{i=1}^{k} \left(\sum_{s=i}^{k} (-1)^{i-s} \binom{n-1}{s-1}\binom{s-1}{i-1} \right) \frac{(\gamma - 1 + \log\gamma_i)^2}{\gamma_i}\\
&= \frac{k\pi^2\vartheta^2}{6} + nk\binom{n-1}{k}\vartheta^2 \sum_{i=1}^{k} (-1)^{k+i} \binom{k-1}{i-1} \frac{(\gamma - 1 + \log\gamma_i)^2}{(n-i)\gamma_i}.
\end{aligned}$$

Since $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\mu) = \mathscr{I}_{11}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) = m\vartheta^{-2}$, $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\vartheta) = \mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta)$, and

$$\begin{aligned}
\mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) = \mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}};0,1) &= -\sum_{s=1}^{m} \{1 + EZ_{s:m:n}\}\\
&= -m + \sum_{s=1}^{m} c_{s-1} \sum_{i=1}^{s} \frac{a_{i,s}}{\gamma_i}(\gamma + \log\gamma_i)\\
&\overset{(2.26)}{=} (\gamma - 1)m + \sum_{s=1}^{m} c_{s-1} \sum_{i=1}^{s} \frac{a_{i,s}}{\gamma_i} \log\gamma_i,
\end{aligned}$$

we also get the Fisher information matrix in the location–scale model. Escobar and Meeker [355] suggested an algorithm to compute the Fisher information in the Weibull case which can also be applied in the progressive censoring setting.

Laplace Distribution (Location)

Since the scale Laplace distribution forms an exponential family, the Fisher information is constant. For the location setting, the situation is different. In particular, the location Laplace distribution provides an example where the standard regularity conditions for Fisher information are not satisfied (see Burkschat and Cramer [232]). However, one can overcome these technical difficulties and obtain a closed form expression even in this case. Burkschat and Cramer [232] established the following result which also holds in the more general framework of generalized order statistics:

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};0) = \sum_{r=1}^{m} c_{r-1} \sum_{j=1}^{r} a_{j,r}\, \kappa(\gamma_j - 2),$$

where

$$\kappa(d) = \int_{1/2}^{1} x^{d-1}\, dx = \begin{cases} \frac{1}{d}\left(1 - \left(\frac{1}{2}\right)^d\right), & d \neq 0, \\ \log(2), & d = 0. \end{cases}$$

Logistic Distribution

In this section, we consider a location family of logistic distributions with location parameter $\mu \in \mathbb{R}$ given by

$$f_\mu(x) = \frac{\mathrm{e}^{-(x-\mu)}}{(1 + \mathrm{e}^{-(x-\mu)})^2}, \quad x \in \mathbb{R}.$$

Since the Fisher information does not depend on the location parameter, it is sufficient to calculate it when $\mu = 0$. Burkschat and Cramer [232] established the following explicit representation in this case:

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};0) = \sum_{r=1}^{n} \prod_{j=1}^{r} \frac{\gamma_j}{\gamma_j + 2}. \tag{9.13}$$

Normal Distribution

For normal distribution, explicit formulas are not available. Computational results for selected censoring schemes are provided by Balakrishnan et al. [140] and Abo-Eleneen [6, 7]. For regularity conditions, we refer to Burkschat and Cramer [232].

Lomax Distribution

Dahmen et al. [320] have established explicit expressions for the Fisher information in a two-parameter family of Lomax distributions with cumulative distribution function

$$F_{q,\vartheta}(x) = 1 - (1 + \vartheta x)^{-q}, \quad x > 0, \tag{9.14}$$

with parameters $q, \vartheta > 0$. Notice that this family does not form a location–scale family of distributions. In this case, they found that

$$\begin{aligned} \mathscr{I}_{11}(\mathbf{X}^{\mathscr{R}}; q, \vartheta) &= \frac{m}{q^2}, \\ \mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}}; q, \vartheta) &= \frac{1}{\vartheta^2} \sum_{s=1}^{m} \prod_{j=1}^{s} \left(1 - \frac{2}{q\gamma_j + 2}\right), \\ \mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}}; q, \vartheta) &= \frac{1}{q\vartheta} \sum_{s=1}^{m} \prod_{j=1}^{s} \left(1 - \frac{1}{q\gamma_j + 1}\right). \end{aligned} \tag{9.15}$$

Notice that, for $q = 1$, $\mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}}; 1, \vartheta)$ equals the expression of the Fisher information for the logistic distribution as given in (9.13). This is due to a transformation result as mentioned in Sect. 9.1.3.

9.1.4 Recurrence Relations for Fisher Information

Using the hazard representation of Fisher information in progressively Type-II censored order statistics, Abo-Eleneen [6] established a representation of Fisher information in terms of Fisher information of minima. In particular, using the density function of the cumulative distribution function given in (2.27), we arrive at

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}}; \theta) = \sum_{r=1}^{m} c_{r-1} \sum_{j=1}^{r} \frac{1}{\gamma_j} a_{j,r} \mathscr{I}(X_{1:\gamma_j}; \theta). \tag{9.16}$$

This representation is quite useful for computational purposes, particularly, when numerical integration is necessary like in the case of the normal distribution. In order to compute the Fisher information in a sample $\mathbf{X}^{\mathscr{R}}$, we have to compute only the Fisher information $\mathscr{I}(X_{1:\gamma_j};\theta)$ in the minima $X_{1:\gamma_j}$, $1 \le j \le m$. Therefore, if you aim at computing it for any censoring scheme, we wish to compute the Fisher information $\mathscr{I}(X_{1:k};\theta)$, $1 \le k \le n$, and use (9.16). For the location parameter of a normal distribution, we have to evaluate the integral

$$\mathscr{I}(X_{1:k};\mu) = 1 + \frac{k^2(k-1)}{2}\int_{-\infty}^{\infty} \varphi^3(x)(1-\Phi(x))^{k-3}dx,$$

where φ and Φ denote the density function and cumulative distribution function of the standard normal distribution, respectively (see Park [704] and Nagaraja and Abo-Eleneen [668]).

Remark 9.1.10. Exchanging summation in (9.16), an alternative formula can be obtained as follows. We get

$$\begin{aligned}\mathscr{I}(\mathbf{X}^{\mathscr{R}};\boldsymbol{\theta}) &= \sum_{j=1}^{m} \mathscr{I}(X_{1:\gamma_j};\boldsymbol{\theta}) \sum_{r=j}^{m} \frac{c_{r-1}}{\gamma_j} a_{j,r} \\ &= \sum_{j=1}^{m} \mathscr{I}(X_{1:\gamma_j};\boldsymbol{\theta}) \sum_{r=j}^{m} \prod_{i=1, i\neq j}^{r} \frac{\gamma_i}{\gamma_i - \gamma_j}.\end{aligned}$$

Recurrence relations for the Fisher information of order statistics have been established by Park [704]. An extension to progressively Type-II censored order statistics has been tried by Abo-Eleneen [7], but the results seem to be in error since they are based on a wrong version of (2.46). However, using representation (9.16) or its matrix version, a simple recurrence relation can be established. For instance, we get for $1 \le r \le m$,

$$\begin{aligned}\mathscr{I}(X_{1:m:n},\ldots,X_{r:m:n};\boldsymbol{\theta}) &= \sum_{k=1}^{r} c_{k-1} \sum_{j=1}^{k} \frac{1}{\gamma_j} a_{j,k} \mathscr{I}(X_{1:\gamma_j};\boldsymbol{\theta}) \\ &= \mathscr{I}(X_{1:m:n},\ldots,X_{r-1:m:n};\boldsymbol{\theta}) + c_{r-1} \sum_{j=1}^{r} \frac{1}{\gamma_j} a_{j,r} \mathscr{I}(X_{1:\gamma_j};\boldsymbol{\theta}) \\ &= \mathscr{I}_{r-1}(\gamma_1,\ldots,\gamma_{r-1}) + s_r(\gamma_1,\ldots,\gamma_r), \qquad (9.17)\end{aligned}$$

where $\mathscr{I}_0(\gamma_1) = \mathscr{I}_0(n) = \mathscr{I}(X_{1:n})$. Since $\mathscr{I}_{m-1}$ depends only on the first $m-1$ γ_j's, this relation provides a recurrence relation in terms of $\gamma_1,\ldots,\gamma_m$ which can be easily implemented. It can also be used to compute the Fisher information for different censoring schemes. Let m and n be fixed. Then, $\mathscr{I}_{m-1}$ is a function of n, m, and the right truncated censoring scheme $(R_1, \ldots, R_{m-2})$. Hence, we can compute the Fisher information for any censoring scheme

$$(R_1, \ldots, R_{m-2}, t, \gamma_{m-1} - t - 1), \quad t = 0, \ldots, \gamma_{m-1} - 1,$$

by evaluating $s_m(\gamma_1, \ldots, \gamma_{m-1}, \gamma_{m-1} - t + 1)$ in (9.17). This observation suggests the following way to compute the Fisher information (matrix) for any censoring scheme $(R_1, \ldots, R_m)$.

Algorithm 9.1.11. Let $m, n \in \mathbb{N}$ be given with $2 \le m \le n$. Suppose the Fisher information $\mathscr{I}(X_{1:m:n}, \ldots, X_{r:m:n}; \boldsymbol{\theta})$ has been computed for any censoring scheme of dimension $1 \le r < m$. Then, the Fisher information of progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{r+1:m:n}$ about the parameter $\boldsymbol{\theta}$ for any censoring scheme can be obtained as follows:

① for any $\gamma_2 > \cdots > \gamma_r \ge m - r + 1$, compute $s_{r+1}(\gamma_1, \ldots, \gamma_r, \ell)$, $\ell = m - r, \ldots, \gamma_r - 1$;
② use (9.17) to compute $\mathscr{I}(X_{1:m:n}, \ldots, X_{r+1:m:n}; \boldsymbol{\theta})$.

Remark 9.1.12. Since $\mathscr{I}(X_{1:r:n}, \ldots, X_{r:r:n}; \boldsymbol{\theta}) = \mathscr{I}(X_{1:m:n}, \ldots, X_{r:m:n}; \boldsymbol{\theta})$ with an appropriately truncated censoring scheme, Algorithm 9.1.11 can be iteratively applied to compute the desired Fisher informations for the next dimension.

9.2 Fisher Information in Progressive Hybrid Censoring

Results on the Fisher information $\mathscr{I}_{m\wedge T:m:n}(\theta)$ about a single parameter θ in progressive hybrid censoring have been established in Park et al. [711] when the population cumulative distribution function F_θ, $\theta \in \Theta \subseteq \mathbb{R}$, has a continuous density function with support contained in the positive real line. For Type-I progressively hybrid censored data, they noticed that

$$\mathscr{I}_{m\wedge T:m:n}(\theta) = \sum_{i=1}^{m} \mathscr{I}(\min\{Y_{1:\gamma_i}, T\}; \theta)$$

using the structure of the joint density function. This leads to the following result [cf. (9.1)] which extends expressions of Wang and He [887] and Park et al. [710] for Type-I hybrid censored data.

Theorem 9.2.1. The Fisher information about θ in a Type-I progressively hybrid censored sample is given by

$$\mathscr{I}_{m\wedge T:m:n}(\theta) = \int_0^T \left\{ \frac{\partial}{\partial\theta} \log \lambda_\theta(x) \right\}^2 \sum_{i=1}^{m} f_{i:m:n;\theta}(x)\, dx.$$

Notice that this expression is very similar to the Fisher information in a progressively Type-II censored sample as given in (9.1). In fact, the integration area is restricted to the interval $[0, T]$.

For an $\mathsf{Exp}(\vartheta)$-distribution, this expression simplifies to

$$\mathscr{I}_{m\wedge T:m:n}(\vartheta) = \frac{1}{\vartheta^2}\sum_{i=1}^{m} F_{i:m:n}\Big(\frac{T}{\vartheta}\Big), \tag{9.18}$$

where $F_{i:m:n}$ denotes the cumulative distribution function of the ith exponential progressively Type-II censored order statistic.

For Type-II progressively hybrid censored data, Park et al. [711] established the following result for the Fisher information $\mathscr{I}_{m\vee T:m:n}(\theta)$ in the Type-II progressively hybrid censored data:

$$\mathscr{I}_{m\vee T:m:n}(\theta) = \mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) + \mathscr{I}(T;\theta) - \mathscr{I}_{(n-R_m)\wedge T:n}(\theta),$$

where $\mathscr{I}(T;\theta)$ and $\mathscr{I}_{(n-R_m)\wedge T:n}(\theta)$ denote the Fisher information in Type-I censored data and Type-I hybrid censored data, respectively. Using that $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = m/\vartheta^2$ (see Sect. 9.1.3), $\mathscr{I}(T;\theta) = nF(T/\vartheta)/\vartheta^2$ (see Park et al. [710]), and carrying out some simple rearrangements, this expression is similar to (9.18) for an $\mathsf{Exp}(\vartheta)$-distribution, i.e.,

$$\mathscr{I}_{m\vee T:m:n}(\vartheta) = \frac{1}{\vartheta^2}\left[m + \sum_{i=n-R_m+1}^{n} F_{i:n}\Big(\frac{T}{\vartheta}\Big)\right]. \tag{9.19}$$

In fact, this expression shows that the gain of information using the Type-II hybrid censoring procedure is given by $\frac{1}{\vartheta^2}\sum_{i=n-R_m+1}^{n} F_{i:n}\Big(\frac{T}{\vartheta}\Big)$ which is increasing in T and bounded by the maximum gain R_m/ϑ^2.

9.3 Tukey's Linear Sensitivity Measure

Fisher information is a widely used concept in statistics, but it cannot be applied in any situations. Therefore, alternative measures like Tukey's linear sensitivity measure have been proposed (see Nagaraja [666], Tukey [863]). A multiparameter situation has been discussed in Chandrasekar and Balakrishnan [243]. Let $\mathbf{Y} = (Y_1,\ldots,Y_r)'$ be a random vector with cumulative distribution function $F_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_k)'$ denotes the vector of parameters. Its mean is given by $\boldsymbol{\alpha} = (\alpha_1,\ldots,\alpha_r)' = E\mathbf{Y}$ and its variance–covariance matrix by $\Sigma = \mathsf{Cov}(\mathbf{Y})$. The matrix $D = (d_{ij})_{i,j}$ of partial derivatives of the mean is defined by

$$d_{ij} = \frac{\partial\alpha_i}{\partial\theta_j}, \quad i = 1,\ldots,r,\ j = 1,\ldots,k.$$

Then, Chandrasekar and Balakrishnan [243] defined a multiparameter version of Tukey's linear sensitivity measure by

$$\mathscr{I}_S(\mathbf{Y};\boldsymbol{\theta}) = \sup_{A\in\mathfrak{A}} D'A'(A\Sigma A')^{-1}AD,$$

where $\mathfrak{A}$ denotes the set of matrices such that $A\Sigma A'$ is non-singular. They established some properties of this measure. For instance, it is additive provided that the random vectors are uncorrelated and it exhibits some monotonicity properties. In location–scale models, it is the inverse of the variance–covariance matrix of the BLUE of $\boldsymbol{\theta} = (\mu, \vartheta)$.

For progressively Type-II censored order statistics $\mathbf{X}^{\mathscr{R}}$ from a two-parameter exponential distribution, this leads to the expression [see (11.7)]

$$\mathscr{I}_S(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) = \left[\frac{\vartheta^2}{n^2(m-1)}\begin{pmatrix} m & -n \\ -n & n^2\end{pmatrix}\right]^{-1} = \frac{1}{\vartheta^2}\begin{pmatrix} n^2 & n \\ n & m\end{pmatrix}.$$

For generalized Pareto distributions, we deduce from (11.12) the expression

$$\begin{aligned}\mathscr{I}_S(\mathbf{X}^{\mathscr{R}};\mu,\vartheta) &= \left[\vartheta^2\frac{q^2}{\Psi e_1 n - e_1^2}\begin{pmatrix} \Psi & -\operatorname{sgn}(q)(\Psi+\frac{e_1}{q}) \\ -\operatorname{sgn}(q)(\Psi+\frac{e_1}{q}) & \Psi+\frac{e_1^2}{q^2}\end{pmatrix}\right]^{-1} \\ &= \frac{1}{\vartheta^2}\begin{pmatrix} \Psi+\frac{e_1^2}{q^2} & \operatorname{sgn}(q)(\Psi+\frac{e_1}{q}) \\ \operatorname{sgn}(q)(\Psi+\frac{e_1}{q}) & \Psi\end{pmatrix}.\end{aligned}$$

9.4 Entropy

The joint differential entropy (Shannon entropy) of $P^{\mathbf{X}^{\mathscr{R}}}$ is defined via the expectation

$$\mathscr{H}^{\mathscr{R}}_{1,\dots,m:m:n} = -E\log f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{X}^{\mathscr{R}}) \tag{9.20}$$

(see Cover and Thomas [282], Park [706]). This quantity will also be discussed in the area of goodness-of-fit tests (see Sect. 19.2). The joint differential entropy (9.20) can be seen as a measure of uncertainty in the progressively Type-II censored sample $\mathbf{X}^{\mathscr{R}}$ which measures uniformity of the density $f^{\mathbf{X}^{\mathscr{R}}}$. On the other hand, the negative entropy can be interpreted as a measure of concentration and, thus, measures the information in $\mathbf{X}^{\mathscr{R}}$ (see Soofi [820]). A connection to Fisher information is the isoperimetric inequality for entropies (see Dembo and Cover [335]): The Fisher information $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta)$ can be bounded from below by

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) \ge 2\pi\mathrm{e}\, m\exp\Big\{-\frac{2}{m}\mathscr{H}^{\mathscr{R}}_{1,\dots,m:m:n}\Big\}.$$

This bound may be useful if the Fisher information is hard to compute. Further information on entropy and information measures can be found in, e.g., Kullback [554] and Cover and Thomas [282]. Other entropy measures for progressively Type-II censored order statistics have been discussed recently by Haj Ahmad and Awad [427, 428].

Balakrishnan et al. [138] established the following representation of the entropy of a progressively censored sample. Notice that the Fisher information can also be expressed in terms of the hazard rate function [see (9.1)].

Theorem 9.4.1. Let F be an absolutely continuous cumulative distribution function with density function f. Then, the entropy in a progressively Type-II censored sample $\mathbf{X}^{\mathscr{R}}$ is given by

$$\mathscr{H}^{\mathscr{R}}_{1,\dots,m:m:n} = -\log c(\mathscr{R}) + m - \sum_{j=1}^{m} E \log \lambda(X_{j:m:n}), \tag{9.21}$$

where $\lambda = f/(1-F)$ and $c(\mathscr{R})$ denote the population hazard rate function and the normalizing constant of the density function, respectively.

Proof. From the representation (2.4) of the joint density function, we get

$$\begin{aligned}\mathscr{H}^{\mathscr{R}}_{1,\dots,m:m:n} &= -\log c(\mathscr{R}) - \sum_{j=1}^{m} E \log \Big(f(X_{j:m:n})[1-F(X_{j:m:n})]^{R_j} \Big) \\ &= -\log c(\mathscr{R}) - \sum_{j=1}^{m} E \log \lambda(X_{j:m:n}) - \sum_{j=1}^{m} (R_j+1) E \log[1-F(X_{j:m:n})].\end{aligned}$$

From Corollary 2.3.7, we get by the continuity of F

$$-E \log[1-F(X_{j:m:n})] = E\left(\sum_{i=1}^{j} \frac{1}{\gamma_i} Z_i \right) = \sum_{i=1}^{j} \frac{1}{\gamma_i},$$

where $Z_1, \dots, Z_j$ are independent Exp(1)-distributed random variables. Furthermore, by interchanging the summation and using the fact that $\gamma_i = \sum_{j=i}^{m} (R_j+1)$, we get $\sum_{j=1}^{m} (R_j+1) \sum_{i=1}^{j} \frac{1}{\gamma_i} = m$. This proves the desired expression. □

Expression (9.21) can be used to calculate the entropy in particular distributions. In particular, it yields directly for standard exponential distribution

$$\mathscr{H}^{\mathscr{R}}_{1,\dots,m:m:n} = -\log c(\mathscr{R}) + m.$$

For a Pareto(α)-distribution with parameter $\alpha > 0$, the following expression holds (see Cramer and Bagh [291]):

$$\mathscr{H}^{\mathscr{R}}_{1,\dots,m:m:n} = -\log c(\mathscr{R}) + m - m \log \alpha + \frac{1}{\alpha} \sum_{j=1}^{m} \frac{m-j+1}{\gamma_j}. \tag{9.22}$$

For a reflected power distribution RPower(α), the expression is almost the same as that in (9.22) except for a minus sign in front of the sum:

$$\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n} = -\log c(\mathscr{R}) + m - m\log\alpha - \frac{1}{\alpha}\sum_{j=1}^{m}\frac{m-j+1}{\gamma_j}. \tag{9.23}$$

For $\alpha = 1$, the entropy in the uniform distribution results.

For a Weibull(ϑ, β)-distribution, the hazard rate function $\lambda(x) = \frac{\beta}{\vartheta}x^{\beta-1}$, $x > 0$, yields the expression

$$\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n} = -\log c(\mathscr{R}) + m - m\log(\beta/\vartheta) - \frac{\beta-1}{\beta}\sum_{j=1}^{m} E\log X^{\beta}_{j:m:n}.$$

As shown in Cramer and Bagh [291], this expression can be written as

$$\begin{aligned}\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n} = -\log c(\mathscr{R}) + m(1+\gamma) - m\log\beta - \frac{m(-\log\vartheta + \gamma)}{\beta}\\ - \frac{\beta-1}{\beta}\sum_{j=1}^{m} c_{j-1}\sum_{i=1}^{j} a_{i,j}\frac{\log\gamma_i}{\gamma_i},\end{aligned}$$

where γ denotes Euler's constant.

Finally, we state an integral representation of the entropy $\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n}$ in terms of the quantile function $F^{\leftarrow}$. A similar version has been established by Balakrishnan et al. [138] (see also Ahmadi [16]). Similar representations for the entropy in a single observation and in a Type-II censored sample can be found in Vasicek [872] and Park [706].

Lemma 9.4.2. Let $\mathbf{X}^{\mathscr{R}}$ be a progressively Type-II censored sample with population cumulative distribution function F and differentiable quantile function $F^{\leftarrow}$. Moreover, let $\gamma_m > 1$. Then,

$$\begin{aligned}\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n} = -\log c(\mathscr{R}) + m + nE\left(\int_0^T \log\left(\frac{d}{dp}F^{\leftarrow}(p)\right)dp\right)\\ -\sum_{j=1}^{m}\frac{m-j+1}{\gamma_j},\end{aligned} \tag{9.24}$$

where T is a random variable with density function

$$f^T(t) = \frac{1}{n}\sum_{r=1}^{m} c_{r-1}\sum_{i=1}^{r} a_{i,r}(\gamma_i - 1)(1-t)^{\gamma_i - 2}, \quad t \in (0,1), \tag{9.25}$$

and $ET = \frac{m}{n}$.

Proof. From (9.21), we can write

$$E\log\lambda(X_{j:m:n}) = E\log f(X_{j:m:n}) - E\log(1-F(X_{j:m:n}))$$
$$= E\log f(X_{j:m:n}) - E\log(1-U_{j:m:n}),$$

where $U_{j:m:n}$ denotes a uniform progressively Type-II censored order statistic. Then, noticing that $-\log(1-U_{j:m:n}))$ can be interpreted as a progressively Type-II censored order statistic from a standard exponential distribution, we get from Theorem 7.2.1 the identity

$$-E\log(1-U_{j:m:n})) = \sum_{i=1}^{j}\frac{1}{\gamma_i}.$$

The first expectation can be written as

$$-E\log f(X_{j:m:n}) = E\log\frac{1}{f(F^{\leftarrow}(U_{j:m:n}))} = E\log\Big(\frac{d}{dp}F^{\leftarrow}(p)\Big)\Big|_{p=U_{j:m:n}}.$$

Then, writing $(1-p)^{\gamma_i-1} = (\gamma_i-1)\int_p^1(1-u)^{\gamma_i-2}du$, $p\in(0,1)$, and $\nu(p) = \log\big(\frac{d}{dp}F^{\leftarrow}(p)\big)$ and using (2.24), we arrive at

$$\begin{aligned}-E\log f(X_{j:m:n}) &= E\nu(U_{j:m:n})\\ &= c_{j-1}\sum_{i=1}^{j}a_{i,j}(\gamma_i-1)\int_0^1\int_p^1(1-u)^{\gamma_i-2}du\,\nu(p)dp\\ &= c_{j-1}\sum_{i=1}^{j}a_{i,j}(\gamma_i-1)\int_0^1(1-u)^{\gamma_i-2}\int_0^u\nu(p)dp\,du.\end{aligned}$$

Combining these results and using (9.21) and (9.25), we arrive at the desired expression. □

Remark 9.4.3.

(i) Balakrishnan et al. [138] obtained the result in Lemma 9.4.2 without imposing the condition $\gamma_m \geq 2$. But, this assumption must not be dropped as illustrated in Remark 9.4.4. Further, they claimed that f^T is a density function and that $ET = \frac{m}{n}$. However, although it has been proved that $\int f^T(t)dt = 1$, it has not been shown that $f^T(t)\geq 0$, $t\in(0,1)$.

(ii) Using Corollary 2.4.7, it can be shown that the cumulative distribution function of T with density function (9.25) is given by

$$F^T(t) = 1 - \frac{1}{n}\sum_{j=1}^{m} f^{U_{j:m:n}}(t)$$
$$= 1 - \frac{1}{1-t}\sum_{j=1}^{m} \frac{R_j+1}{n}\,\overline{F}^{U_{j:m:n}}(t), \quad t \in \mathbb{R}, \tag{9.26}$$

where $\overline{F}^{U_{j:m:n}}$ denotes the survival function of $U_{j:m:n}$, $1 \le j \le m$.

Remark 9.4.4. If the condition $\gamma_m \ge 2$ in Lemma 9.4.2 does not hold, the representation (9.24) with the density function f^T in (9.25) is not true. First, notice that $\gamma_i > 1$ for $i < m$ since $\gamma_1 > \cdots > \gamma_m \ge 1$. Given $\gamma_m = 1$, we can proceed as in the proof of Lemma 9.4.2 except for the term with $j = m$. In this case, we get

$$-E\log f(X_{m:m:n}) = Ev(U_{m:m:n})$$
$$= c_{m-1}\sum_{i=1}^{m-1} a_{i,m}(\gamma_i - 1)\int_0^1\int_p^1 (1-u)^{\gamma_i-2}du\, v(p)dp$$
$$+ c_{m-1}a_{m,m}\int_0^1 v(p)dp.$$

With $c_{m-1}a_{m,m} = \prod_{i=1}^{m-1}\frac{\gamma_i}{\gamma_i-1} = \delta$ and f^T as in (9.25), we get

$$-\sum_{j=1}^{m} E\log\lambda(X_{j:m:n}) = (n-\delta)E\Big(\int_0^{T^*}\log\Big(\frac{d}{dp}F^{\leftarrow}(p)\Big)dp\Big) - \sum_{j=1}^{m}\frac{m-j+1}{\gamma_j}$$
$$+ \delta\int_0^p \log\Big(\frac{d}{dp}F^{\leftarrow}(p)\Big)dp,$$

where T^* is a random variable with density function $f^{T^*} = \frac{n}{n-\delta}f^T$. Notice that $\mathscr{H}_{1:1:1} = \int_0^p \log\big(\frac{d}{dp}F^{\leftarrow}(p)\big)dp$ as shown in Vasicek [872] and Park [706]. Therefore, we can write (9.24) as

$$\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n} = -\log c(\mathscr{R}) + m + (n-\delta)E\Big(\int_0^{T^*}\log\Big(\frac{d}{dp}F^{\leftarrow}(p)\Big)dp\Big)$$
$$-\sum_{j=1}^{m}\frac{m-j+1}{\gamma_j} + \delta\mathscr{H}_{1:1:1}(X).$$

Remark 9.4.5. The entropy in a progressively Type-II censored sample has also been discussed in Abo-Eleneen [8]. He addressed recurrence relations and presented results for selected censoring schemes for normal and logistic distributions. Ahmadi [16] discussed characterizations of distributions via the Müntz-Szász theorem (see also Kamps [498, p. 103]) and stochastic ordering w.r.t. the entropy order. In particular, he extended a result of Ebrahimi et al. [347] for order statistics to the case of progressive censoring.

9.5 Kullback–Leibler Information

The Kullback–Leibler information serves as a measure to compare two distributions. For two progressively Type-II censored samples $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{Y}^{\mathscr{R}}$ with the same censoring scheme, but possibly different population density functions f and g, it is defined as (see Kullback [554])

$$\mathscr{I}_{\mathscr{R}}(f\|g) = \mathscr{I}(f^{\mathbf{X}^{\mathscr{R}}}\|g^{\mathbf{Y}^{\mathscr{R}}}) = \int_S f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}) \log \frac{f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x})}{g^{\mathbf{Y}^{\mathscr{R}}}(\mathbf{x})} d\mathbf{x}, \tag{9.27}$$

where $f^{\mathbf{X}^{\mathscr{R}}}$ and $g^{\mathbf{Y}^{\mathscr{R}}}$ denote the joint density functions of the progressively Type-II censored samples. Balakrishnan et al. [138] and Rad et al. [736] have used Kullback–Leibler information to construct goodness-of-fit tests (see Sect. 19.2). Using properties of the logarithm, it follows that

$$\mathscr{I}_{\mathscr{R}}(f\|g) = -\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n} - \int_S f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}) \log g^{\mathbf{Y}^{\mathscr{R}}}(\mathbf{x}) d\mathbf{x}, \tag{9.28}$$

where $\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n}$ is the entropy of $P^{\mathbf{X}^{\mathscr{R}}}$ as in (9.20).

The Kullback–Leibler information in favor of an IID sample $Y_1, \ldots, Y_m$ from a population with population density function f against a progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$ from the same population has been discussed in Cramer and Bagh [291]. Here, a slightly more general distance is considered for two samples of progressively Type-II censored order statistics with the same sample size m but possibly different censoring schemes $\mathscr{R}$ and $\mathscr{S}$ and original sample sizes $\gamma_1(\mathscr{R})$ and $\gamma_1(\mathscr{S})$. According to (9.27), the Kullback–Leibler distance of $P^{\mathbf{X}^{\mathscr{R}}}$ and $P^{\mathbf{X}^{\mathscr{S}}}$ is defined by

$$\mathscr{I}_0(\mathscr{R}\|\mathscr{S}) = \mathscr{I}_0(f^{\mathbf{X}^{\mathscr{R}}}\|f^{\mathbf{X}^{\mathscr{S}}}) = \int_S f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}) \log \frac{f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x})}{f^{\mathbf{X}^{\mathscr{S}}}(\mathbf{x})} d\mathbf{x}, \tag{9.29}$$

where S denotes the support of $f^{\mathbf{X}^{\mathscr{R}}}$ and $f^{\mathbf{X}^{\mathscr{S}}}$. The Kullback–Leibler information given in (9.29) has the following representations.

Theorem 9.5.1. Let $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{X}^{\mathscr{S}}$ be progressively Type-II censored order statistics with population density function f and censoring schemes $\mathscr{R}$ and $\mathscr{S}$, respectively. Then, the Kullback–Leibler information is given by

$$\mathscr{I}_0(\mathscr{R}\|\mathscr{S}) = \log c(\mathscr{R}) - \log c(\mathscr{S}) - m + \sum_{j=1}^{m} \frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})}.$$

Proof. Using the quantile representation given in (2.14), this yields

$$\begin{aligned}\mathscr{I}_0(\mathscr{R}\|\mathscr{S}) &= \log c(\mathscr{R}) - \log c(\mathscr{S}) + \sum_{i=1}^{m}(R_i - S_i) E \log\left[1 - F(X_{i:m:n})\right] \\ &= \log c(\mathscr{R}) - \log c(\mathscr{S}) + \sum_{i=1}^{m}(R_i - S_i) \sum_{j=1}^{i} \frac{1}{\gamma_j(\mathscr{R})} E \log U_j.\end{aligned}$$

Since $-\log U_j$ has a standard exponential distribution and, thus, $E \log U_j = -1$, we get

$$\begin{aligned}\mathscr{I}_0(\mathscr{R}\|\mathscr{S}) &= \log c(\mathscr{R}) - \log c(\mathscr{S}) - \sum_{i=1}^{m}(R_i - S_i) \sum_{j=1}^{i} \frac{1}{\gamma_j(\mathscr{R})} \\ &= \log c(\mathscr{R}) - \log c(\mathscr{S}) - m + \sum_{j=1}^{m} \frac{1}{\gamma_j(\mathscr{R})} \sum_{i=j}^{m}(S_i + 1) \\ &= \log c(\mathscr{R}) - \log c(\mathscr{S}) - m + \sum_{j=1}^{m} \frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})}.\end{aligned}$$

This yields the desired expression. □

Given the censoring scheme $\mathscr{R}$, it is immediate from the properties of the Kullback–Leibler information that the closest distribution $P^{\mathbf{X}^{\mathscr{S}}}$ is specified by the censoring scheme $\mathscr{R} = \mathscr{S}$. On the other hand, one might ask which distribution is most dissimilar to $P^{\mathbf{X}^{\mathscr{S}}}$. From Theorem 9.5.1, we can write

$$\mathscr{I}_0(\mathscr{R}\|\mathscr{S}) = \sum_{j=1}^{m}\Big[-\log \frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})} - 1 + \frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})}\Big] = \sum_{j=1}^{m} h\Big(\frac{\gamma_j(\mathscr{S})}{\gamma_j(\mathscr{R})}\Big)$$

with $h(x) = x - \log x - 1 \geq 0$, $x > 0$. Since h is a strictly convex function, the maximum is attained at the boundary. Notice that the original sample size $\gamma_1(\mathscr{R})$ of the progressively censored sample $\mathbf{X}^{\mathscr{R}}$ has to be bounded because otherwise a maximum does not exist. Moreover, this shows that right censoring will be most dissimilar when $\gamma_1(\mathscr{R})$ exceeds a certain level. Therefore, it makes sense to bound the sample size.

Suppose that $\gamma_1(\mathscr{R}) = \gamma_1(\mathscr{S})$. Then, taking the ordering $1 \leq \gamma_m(\mathscr{R}) < \cdots < \gamma_1(\mathscr{R})$ into account, we get, e.g., that $\mathscr{O}_1$ and $\mathscr{O}_m$ are most dissimilar.

Denoting by $\mathscr{H}_{1,\ldots,m:m:n}^{\mathscr{R};\mathsf{uniform}}$ the entropy of uniform progressively Type-II censored order statistics with censoring scheme $\mathscr{R}$ and let $\mathscr{S} = (0^{*m})$, we get from (9.23)

$$\mathscr{I}_0(\mathscr{R}\|\mathscr{S}) = -\mathscr{H}_{1,\ldots,m:m:n}^{\mathscr{R};\mathsf{uniform}}$$

as given in (9.23) with $\alpha = 1$.

Remark 9.5.2. The Kullback–Leibler information between two progressively Type-II censored samples with identical censoring scheme but different population distributions has been discussed in Park [707]. Here,

$$\mathscr{I}_0(g^{\mathbf{X}^{\mathscr{R}}}\|f^{\mathbf{X}^{\mathscr{R}}}) = \int_S g^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}) \log \frac{g^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x})}{f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x})} d\mathbf{x}. \tag{9.30}$$

Park [707] showed that the Kullback–Leibler information in (9.30) can be expressed in terms of the hazard rates λ_F and λ_G, i.e., as

$$\mathscr{I}_0(g^{\mathbf{X}^{\mathscr{R}}}\|f^{\mathbf{X}^{\mathscr{R}}}) = \int_{-\infty}^{\infty} h\Big(\frac{\lambda_F(x)}{\lambda_G(x)}\Big) \sum_{j=1}^{m} g_{j:m:n}^{\mathscr{R}}(x)\, dx,$$

where $h(t) = t - \log t - 1$, $t > 0$. A similar representation holds for the marginal distributions of the first r progressively Type-II censored order statistics. Notice that this result parallels expression (9.1) obtained for the Fisher information.

The $\mathscr{I}_\alpha$-information is defined by

$$\mathscr{I}_\alpha(\mathscr{R}\|\mathscr{S}) = \frac{1}{\alpha(1-\alpha)}\Big(1 - \int_S \Big[f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x})\Big]^{\alpha}\Big[f^{\mathbf{X}^{\mathscr{S}}}(\mathbf{x})\Big]^{1-\alpha} d\mathbf{x}\Big),$$

where $0 < \alpha < 1$ is a given parameter and S denotes the support of $f^{\mathbf{X}^{\mathscr{R}}}$. Properties as well as generalizations of $\mathscr{I}_\alpha(\mathscr{R}\|\mathscr{S})$ can be found in Vajda [864].

Theorem 9.5.3. For $\mathscr{R}, \mathscr{S} \in \mathscr{C}_{m,n}^m$, $\mathscr{I}_\alpha(\mathscr{R}\|\mathscr{S})$ is given by the expression

$$\mathscr{I}_\alpha(\mathscr{R}\|\mathscr{S}) = \frac{1}{\alpha(1-\alpha)}\Big(1 - \frac{c(\mathscr{R})^\alpha c(\mathscr{S})^{1-\alpha}}{d_\alpha(\mathscr{R})}\Big)$$

with $d_\alpha(\mathscr{R}) = \prod_{j=1}^m [(1-\alpha)\gamma_j(\mathscr{S}) + \alpha\gamma_j(\mathscr{R})]$.

Proof. From (2.4), we find for $x_1 < \cdots < x_m$

$$\begin{aligned}
&\Big(f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x})\Big)^{\alpha}\Big(f^{\mathbf{X}^{\mathscr{S}}}(\mathbf{x})\Big)^{1-\alpha}\\
&= c(\mathscr{R})^\alpha c(\mathscr{S})^{1-\alpha}\Big(\prod_{i=1}^m f(x_i)\Big)^{\alpha}\Big(\prod_{i=1}^m [1-F(x_i)]^{R_i}\Big)^{\alpha}\\
&\qquad \times \Big(\prod_{i=1}^m f(x_i)\Big)^{1-\alpha}\Big(\prod_{i=1}^m [1-F(x_i)]^{S_i}\Big)^{1-\alpha}
\end{aligned}$$

$$= c(\mathscr{R})^{\alpha} c(\mathscr{S})^{1-\alpha} \prod_{i=1}^{m} \left[f(x_i)(1 - F(x_i))^{\alpha R_i + (1-\alpha) S_i} \right]$$

$$= \frac{c(\mathscr{R})^{\alpha} c(\mathscr{S})^{1-\alpha}}{d_\alpha(\mathscr{R})} \cdot d_\alpha(\mathscr{R}) \prod_{i=1}^{m} \left[f(x_i)(1 - F(x_i))^{\alpha R_i + (1-\alpha) S_i} \right].$$

The term $d_\alpha(\mathscr{R}) \prod_{i=1}^{m} \left[f(x_i)(1 - F(x_i))^{\alpha R_i + (1-\alpha) S_i} \right]$ is the joint density function of generalized order statistics based on parameters $(1-\alpha)\gamma_j(\mathscr{S}) + \alpha\gamma_j(\mathscr{R})$, $j = 1, \ldots, m$, and the cumulative distribution function F (see Kamps [498, 499], and Cramer and Kamps [301]). This proves the result. □

9.6 Pitman Closeness

Pitman closeness is a measure to compare the closeness of two estimators to the estimated parameter which was introduced in 1937 by Pitman [722]. A detailed discussion has been provided in the monograph by Keating et al. [512]. The notion of Pitman closeness has been introduced by Balakrishnan et al. [142] to determine the closest order statistic in a sample of size n to a given population quantile (for the median, we refer to Balakrishnan et al. [143]). Given a quantile ξ_p of the baseline distribution, they were interested in identifying an order statistic $X_{\ell:n}$ such that

$$P(|X_{\ell:n} - \xi_p| < |X_{i:n} - \xi_p|) \geq \frac{1}{2} \quad \text{for all } i \in \{1, \ldots, n\} \setminus \{\ell\}.$$

Volterman et al. [880] adapted this idea to two independent progressively Type-II censored samples $X_{1:m:n}^{\mathscr{R}}, \ldots, X_{m:m:n}^{\mathscr{R}}$ and $Y_{1:s:k}^{\mathscr{S}}, \ldots, Y_{s:s:k}^{\mathscr{S}}$ from the same cumulative distribution function F. For a quantile ξ_p, they introduced the Pitman closeness probabilities

$$\pi_{i,j}^{\mathscr{R},\mathscr{S}}(\xi_p) = P(|X_{i:m:n}^{\mathscr{R}} - \xi_p| < |Y_{j:s:k}^{\mathscr{S}} - \xi_p|), \quad 1 \leq i \leq m, 1 \leq j \leq s.$$

They found that

$$\pi_{i,j}^{\mathscr{R},\mathscr{S}}(\xi_p) = F^{U_{i:m:n}^{\mathscr{R}}}(F(2\xi_p - \xi_0)) - \prod_{\ell=1}^{i} \gamma_\ell(\mathscr{R}) \prod_{\ell=1}^{j} \gamma_\ell(\mathscr{S}) \sum_{\ell_1=1}^{i} \sum_{\ell_2=1}^{j} \frac{a_{\ell_1,i}^{\mathscr{R}} a_{\ell_2,j}^{\mathscr{S}}}{\gamma_{\ell_2}(\mathscr{S})}$$
$$\times \left\{ \frac{1 - 2(1-p)^{\gamma_{\ell_1}(\mathscr{R}) + \gamma_{\ell_2}(\mathscr{S})}}{\gamma_{\ell_1}(\mathscr{R}) + \gamma_{\ell_2}(\mathscr{S})} - \int_{F(2\xi_p - \xi_1)}^{p} w(u)du + \int_{p}^{F(2\xi_p - \xi_0)} w(u)du \right\}, \tag{9.31}$$

where $w(u) = (1-u)^{\gamma_{\ell_1}(\mathscr{R})-1}\big[1 - F(2\xi_p - \xi_u)\big]^{\gamma_{\ell_2}(\mathscr{S})}$. In the case of a symmetric population distribution, the above expression simplifies considerably for the population median $\xi_{0.5}$ and can be written in terms of incomplete beta functions as

$$\pi_{i,j}^{\mathscr{R},\mathscr{S}}(\xi_{0.5}) = 1 - \prod_{\ell=1}^{i}\gamma_\ell(\mathscr{R})\prod_{\ell=1}^{j}\gamma_\ell(\mathscr{S})\sum_{\ell_1=1}^{i}\sum_{\ell_2=1}^{j}\frac{a_{\ell_1,i}^{\mathscr{R}}a_{\ell_2,j}^{\mathscr{S}}}{\gamma_{\ell_2}(\mathscr{S})}$$
$$\times\left\{\frac{1-2^{1-\gamma_{\ell_1}(\mathscr{R})-\gamma_{\ell_2}(\mathscr{S})}}{\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})} + \mathsf{B}_1(\gamma_{\ell_2}(\mathscr{S})+1, \gamma_{\ell_1}(\mathscr{R}))\right.$$
$$\left. - 2\mathsf{B}_{1/2}(\gamma_{\ell_2}(\mathscr{S})+1, \gamma_{\ell_1}(\mathscr{R}))\right\},$$

where $\mathsf{B}_t(\cdot,\cdot)$ denotes the incomplete beta functions with $t \in [0,1]$.

Remark 9.6.1. As a corollary, Volterman et al. [880] obtained an expression for exceedance probabilities

$$P(X_{i:m:n}^{\mathscr{R}} < Y_{j:s:k}^{\mathscr{S}}) = \prod_{\ell=1}^{i}\gamma_\ell(\mathscr{R})\prod_{\ell=1}^{j}\gamma_\ell(\mathscr{S})$$
$$\times\sum_{\ell_1=1}^{i}\sum_{\ell_2=1}^{j}\frac{a_{\ell_1,i}^{\mathscr{R}}a_{\ell_2,j}^{\mathscr{S}}}{\gamma_{\ell_2}(\mathscr{S})[\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})]}. \tag{9.32}$$

Example 9.6.2. Volterman et al. [880] have evaluated the expression in (9.31) for exponential and uniform distributions. For a standard exponential distribution, they obtained the expression

$$\pi_{i,j}^{\mathscr{R},\mathscr{S}}(\xi_p) = F^{U_{i:m:n}^{\mathscr{R}}}(2p-p^2) - \prod_{\ell=1}^{i}\gamma_\ell(\mathscr{R})\prod_{\ell=1}^{j}\gamma_\ell(\mathscr{S})\sum_{\ell_1=1}^{i}\sum_{\ell_2=1}^{j}\frac{a_{\ell_1,i}^{\mathscr{R}}a_{\ell_2,j}^{\mathscr{S}}}{\gamma_{\ell_2}(\mathscr{S})}$$
$$\times\left\{\frac{1-2(1-p)^{\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})}}{\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})} - \kappa(p;\gamma_{\ell_1}(\mathscr{R}),\gamma_{\ell_2}(\mathscr{S}))\right\},$$

where $\kappa(p;\gamma_{\ell_1}(\mathscr{R}),\gamma_{\ell_2}(\mathscr{S})) = \begin{cases} 0, & \gamma_{\ell_1}(\mathscr{R}) = \gamma_{\ell_2}(\mathscr{S}) \\ \frac{\left[(1-p)^{\gamma_{\ell_1}(\mathscr{R})}-(1-p)^{\gamma_{\ell_2}(\mathscr{S})}\right]^2}{\gamma_{\ell_1}(\mathscr{R})-\gamma_{\ell_2}(\mathscr{S})}, & \gamma_{\ell_1}(\mathscr{R}) \neq \gamma_{\ell_2}(\mathscr{S}) \end{cases}$.

For a standard uniform distribution, the probability is given by

$$\pi_{i,j}^{\mathscr{R},\mathscr{S}}(\xi_p) = F^{U_{i:m:n}^{\mathscr{R}}}(\min\{1,2p\}) - \prod_{\ell=1}^{i}\gamma_\ell(\mathscr{R})\prod_{\ell=1}^{j}\gamma_\ell(\mathscr{S})\sum_{\ell_1=1}^{i}\sum_{\ell_2=1}^{j}\frac{a_{\ell_1,i}^{\mathscr{R}}a_{\ell_2,j}^{\mathscr{S}}}{\gamma_{\ell_2}(\mathscr{S})}$$
$$\times\left\{\frac{1-2(1-p)^{\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})}}{\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})} + p^{\star\,\gamma_{\ell_1}(\mathscr{R})+\gamma_{\ell_2}(\mathscr{S})}\right.$$

$$\times \Big[\mathsf{B}_{[1-1/p^\star]_+}(\gamma_{\ell_2}(\mathscr{S})+1,\gamma_{\ell_1}(\mathscr{R})) + \mathsf{B}_{1-[1-1/p^\star]_+}(\gamma_{\ell_2}(\mathscr{S})+1,\gamma_{\ell_1}(\mathscr{R})) - 2\mathsf{B}_{1/2}(\gamma_{\ell_2}(\mathscr{S})+1,\gamma_{\ell_1}(\mathscr{R}))\Big]\Big\},$$

where $p^\star = 2(1-p)$.

The above idea has been further discussed by Balakrishnan et al. [146]. They introduced the notion of simultaneous-closeness probability following a proposal of Blyth [209]. Replacing the order statistics by progressively Type-II censored order statistics, the same idea has been introduced into the framework of progressively Type-II censored order statistics by Volterman et al. [879].

Definition 9.6.3. The simultaneous-closeness probability (SCP) of $X_{i:m:n}$, $i \in \{1,\ldots,m\}$, among the order statistics $X_{1:m:n},\ldots,X_{m:m:n}$ in the estimation of a population parameter θ is defined as

$$\pi_{i:m:n}(\theta) = P\Big(|X_{i:m:n}-\theta| < \min_{j\neq i}|X_{j:m:n}-\theta|\Big).$$

They showed that this probability can we written using the pairwise Voronoi region associated with $X_{i:m:n}$ defined by

$$\mathcal{A}_{i:m:n} = \{X_{i-1:m:n}+X_{i:m:n} \le 2\theta\},\quad i\in\{2,\ldots,m\},$$

and the simultaneous Voronoi region associated with $X_{i:m:n}$ defined by

$$\mathcal{B}_{i:m:n} = \{|X_{i:m:n}-\theta| \le \min_{j\neq i}|X_{j:m:n}-\theta|\},\quad i\in\{1,\ldots,m\}.$$

Then,

$$\pi_{i:m:n}(\theta) = P(\mathcal{B}_{i:m:n}) = P(\mathcal{A}_{i:m:n}) - P(\mathcal{A}_{i+1:m:n}),\quad i\in\{1,\ldots,m\},$$

where $\mathcal{A}_{m+1:m:n} = \emptyset$. Using this connection, Volterman et al. [879] established the following expression for the SCP.

Theorem 9.6.4. Suppose the population cumulative distribution function F has a bounded support (α,ω). Then, for any $i\in\{2,\ldots,m-1\}$ and any quantile ξ_p, $p\in(0,1)$, the SCP of $X_{i:m:n}$ to ξ_p is given by

$$\begin{aligned}\pi_{i:m:n}(\xi_p) &= F_{i-1:m:n}(\xi_p) - F_{i:m:n}(\xi_p)\\ &+ \prod_{j=1}^{i}\gamma_j\sum_{j=1}^{i}a_{j,i}\int_0^p(1-u)^{\gamma_j-\gamma_{i+1}-1}\big[1-F\big(\min\{\omega,2\xi_p-F^{\leftarrow}(u)\}\big)\big]^{\gamma_{i+1}}du\\ &- \prod_{j=1}^{i-1}\gamma_j\sum_{j=1}^{i-1}a_{j,i-1}\int_0^p(1-u)^{\gamma_j-\gamma_i-1}\big[1-F\big(\min\{\omega,2\xi_p-F^{\leftarrow}(u)\}\big)\big]^{\gamma_i}du.\end{aligned}$$

Furthermore, $\pi_{1:m:n}(\xi_p) = 1-P(\mathcal{A}_{2:m:n})$ and $\pi_{m:m:n}(\xi_p) = P(\mathcal{A}_{m:m:n})$.

If the distribution has an infinite right endpoint of support, i.e., $\omega = \infty$, then the minimum in the argument of the cumulative distribution function can be replaced by $2\xi_p - F^{\leftarrow}(u)$. Volterman et al. [879] discussed also some particular baseline distributions like exponential, uniform, and normal distributions. For exponential distribution, they found the explicit expression

$$P(\mathcal{A}_{i+1:m:n}) = F_{i:m:n}(\xi_p) - \Big(\prod_{j=1}^{i}\gamma_j\Big)(1-p)^{2\gamma_{i+1}}$$
$$\times \sum_{j=1}^{i} a_{j,i} \begin{cases} \frac{1}{\gamma_j - 2\gamma_{i+1}}\big[1-(1-p)^{\gamma_j-2\gamma_{i+1}}\big], & \gamma_j \neq 2\gamma_{i+1} \\ -\log(1-p), & \gamma_j = 2\gamma_{i+1} \end{cases}.$$

For the uniform distribution, a simple expression is available in terms of the incomplete beta function $\mathsf{B}_t(\alpha,\beta)$, i.e.,

$$P(\mathcal{A}_{i+1:m:n}) = F_{i:m:n}(p) - \prod_{j=1}^{i}\gamma_j \sum_{j=1}^{i} a_{j,i}[2(1-p)]^{\gamma_j}$$
$$\times \Big(\mathsf{B}_{1/2}(\gamma_{i+1}+1, \gamma_j - \gamma_{i+1}) - \mathsf{B}_{[1-1/(2(1-p))]_+}(\gamma_{i+1}+1, \gamma_j-\gamma_{i+1})\Big).$$

Volterman et al. [879] provided extensive computations for selected censoring schemes, i.e., for

$$\mathscr{R}_1 = (20, 0^{*9}), \qquad \mathscr{R}_2 = (0^{*9}, 20), \qquad \mathscr{R}_3 = (2^{*10}),$$
$$\mathscr{R}_4 = (5^{*2}, 0^{*6}, 5^{*2}), \quad \mathscr{R}_5 = (0^{*4}, 20, 0^{*5}), \quad \mathscr{R}_6 = (0^{*4}, 10^{*2}, 0^{*4}).$$

They found that all probabilities are quite close. For late progressive censoring (like for $\mathscr{R}_2$ and $\mathscr{R}_4$), the largest progressively Type-II censored order statistic $X_{m:m:n}$ is Pitman closest to the upper quantiles. However, if $R_m > 0$ and, thus, right censoring is present in the data, $X_{m:m:n}$ may not be very close to the desired quantile. In order to overcome this difficulty, one may extend the life test by observing more items (like in adaptive censoring or in Type-II progressive hybrid censoring).

The notion of simultaneous Pitman closeness has been applied by Volterman et al. [880] to measure the distance of predictors of a future sample of progressively Type-II censored order statistics.

Definition 9.6.5. Given a progressively Type-II censored sample $Y^{\mathscr{S}}_{1:s:k}, \ldots, Y^{\mathscr{S}}_{s:s:k}$, the simultaneous-closeness probability of $Y^{\mathscr{S}}_{j:s:k}$, $j \in \{1,\ldots,s\}$, to a progressively Type-II censored order statistic $X^{\mathscr{R}}_{i:m:n}$, $i \in \{1,\ldots,m\}$ from a future sample is defined as

$$\pi^{\mathscr{S}}_j(i,\mathscr{R}) = P\Big(|Y^{\mathscr{S}}_{j:s:k} - X^{\mathscr{R}}_{i:m:n}| < \min_{\ell\neq j}|Y^{\mathscr{S}}_{\ell:s:k} - X^{\mathscr{R}}_{i:m:n}|\Big).$$

They found the following expression for this SCP.

Theorem 9.6.6. Given a progressively Type-II censored sample $Y_{1:s:k}^{\mathscr{S}}, \ldots, Y_{s:s:k}^{\mathscr{S}}$, the simultaneous-closeness probability of $Y_{j:s:k}^{\mathscr{S}}$, $j \in \{1, \ldots, s\}$, to a progressively Type-II censored order statistic $X_{i:m:n}^{\mathscr{R}}$, $i \in \{1, \ldots, m\}$ from a future sample is given by

$$\pi_j^{\mathscr{S}}(i, \mathscr{R}) = \begin{cases} 1 - \pi_{1,2}^{\mathscr{S}}(i, \mathscr{R}), & j = 1 \\ \pi_{j,j-1}^{\mathscr{S}}(i, \mathscr{R}) - \pi_{j+1,j}^{\mathscr{S}}(i, \mathscr{R}), & j \in \{2, \ldots, s-1\} \\ \pi_{s,s-1}^{\mathscr{S}}(i, \mathscr{R}), & j = s \end{cases},$$

where, for $j = 1, \ldots, s-1$,

$$\begin{aligned} \pi_{j+1,j}^{\mathscr{S}}(i, \mathscr{R}) = P(Y_{j:s:k}^{\mathscr{S}} < X_{i:m:n}^{\mathscr{R}}) - \prod_{\ell=1}^{i} \gamma_\ell(\mathscr{R}) \prod_{\ell=1}^{j} \gamma_\ell(\mathscr{S}) \sum_{\ell_1=1}^{i} \sum_{\ell_2=1}^{j} \frac{a_{\ell_1,i}^{\mathscr{R}} a_{\ell_2,j}^{\mathscr{S}}}{\gamma_{\ell_2}(\mathscr{S})} \\ \times \left\{ \int_0^1 \int_0^v (1-u)^{\gamma_{\ell_2}(\mathscr{S}) - \gamma_{j+1}(\mathscr{S}) - 1} (1-v)^{\gamma_{\ell_1}(\mathscr{R}) - 1} \right. \\ \left. \times \overline{F}^{\gamma_{j+1}(\mathscr{S})}\big(\min\{\xi_1, 2\xi_v - \xi_u\}\big) du\, dv \right\}, \end{aligned}$$

and the exceedance probability is given as in (9.32).

Chapter 10
Progressive Type-II Censoring Under Nonstandard Conditions

Order statistics under nonstandard conditions have been studied under various assumptions. Surveys on these results can be found in Arnold and Balakrishnan [51], Rychlik [764], David and Nagaraja [327, Chap. 5], and Balakrishnan [83]. For progressively Type-II censored order statistics, first results in this regard were established in Balakrishnan and Cramer [93] and Cramer et al. [314] for independent but not necessarily identically distributed (INID) random variables $X_1, \ldots, X_n$ from absolutely continuous cumulative distribution functions $F_1, \ldots, F_n$ with density functions $f_1, \ldots, f_n$. Fischer et al. [371] obtained a general mixture representation of the distributions which generalizes the results of Thomas and Wilson [843] for IID random variables (see also Guilbaud [418, 419]). Results for dependent samples are due to Rezapour et al. [752, 753].

10.1 Mixture Representation for Progressively Type-II Censored Order Statistics with Arbitrary Distribution

The following theorem presents a mixture representation for the joint distribution of the progressively Type-II censored order statistics in the general setup. It relates the cumulative distribution function of a progressively Type-II censored sample to the cumulative distribution functions of m-dimensional marginals of order statistics $X_{1:n}, \ldots, X_{n:n}$ from the original sample $X_1, \ldots, X_n$.

Theorem 10.1.1. (Fischer et al. [371]) Let $x_1, x_2, \ldots, x_m \in \mathbb{R}$. Then, we have in Model 1.2.5,

$$P(\mathbf{X}^{\mathscr{R}} \le \mathbf{x}_m) = \frac{1}{\prod_{i=1}^{m} \binom{\gamma_i - 1}{R_i}} \sum_{\mathfrak{Z}^{\mathscr{R}}} P(X_{k_1:n} \le x_1, X_{k_2:n} \le x_2, \ldots, X_{k_m:n} \le x_m), \quad (10.1)$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_10,

where the summation is over all elements $(k_1, \mathcal{R}_1, \ldots, k_m, \mathcal{R}_m)$ of the set

$$\mathfrak{Z}^{\mathscr{R}} = \Bigg\{ (l_1, \mathcal{S}_1, l_2, \mathcal{S}_2, \ldots, l_m, \mathcal{S}_m) \,:\, l_1 = 1, \mathcal{S}_1 \subseteq \{2, \ldots, n\}, |\mathcal{S}_1| = R_1,$$
$$l_i = \min\Big(\{1, \ldots, n\} \setminus \Big[\bigcup_{j=1}^{i-1} (\mathcal{S}_j \cup \{l_j\}) \Big]\Big),$$
$$\mathcal{S}_i \subseteq \{1, \ldots, n\} \setminus \Big[\bigcup_{j=1}^{i-1} \mathcal{S}_j \cup \bigcup_{j=1}^{i} \{l_j\} \Big], |\mathcal{S}_i| = R_i \;\forall\; i = 2, \ldots, m \Bigg\}. \quad (10.2)$$

Proof. Let Z be a random variable with values in $\mathfrak{Z}^{\mathscr{R}}$ that models the choice of the sets $\mathcal{R}_1, \ldots, \mathcal{R}_m$ in Construction 1.1.3. Notice that this determines the indices $k_1, \ldots, k_m$ of the failure times $X_{k_i:n}$ which we include in the notation. According to our assumptions, Z is a random variable with values in $\mathfrak{Z}^{\mathscr{R}}$, which has a discrete uniform distribution and is independent of the original sample $X_1, \ldots, X_n$. Consequently, we may assume without loss of any generality that Z is defined on the same probability space as $X_1, X_2, \ldots, X_n$, for otherwise we can consider an appropriate product space. Since γ_i denotes the number of units in the experiment before the ith failure, i.e., γ_i is the cardinality of the set $\mathcal{N}_i$ in the algorithm before the ith run, a simple combinatorial argument readily yields

$$|\mathfrak{Z}^{\mathscr{R}}| = \prod_{i=1}^{m} \binom{\gamma_i - 1}{R_i}$$

(see also Stigler [825]). Hence, it follows that

$$\begin{aligned}
&P(\mathbf{X}^{\mathscr{R}} \le \mathbf{x}_m)\\
&= \sum_{\mathfrak{Z}^{\mathscr{R}}} P\Big(\mathbf{X}^{\mathscr{R}} \le \mathbf{x}_m \Big| Z = (k_1, \mathcal{R}_1, \ldots, k_m, \mathcal{R}_m)\Big)\\
&\qquad\qquad \times P\Big(Z = (k_1, \mathcal{R}_1, \ldots, k_m, \mathcal{R}_m)\Big)\\
&= \sum_{\mathfrak{Z}^{\mathscr{R}}} \Bigg[\prod_{i=1}^{m} \binom{\gamma_i - 1}{R_i}\Bigg]^{-1} P\Big(\mathbf{X}^{\mathscr{R}} \le \mathbf{x}_m \Big| Z = (k_1, \mathcal{R}_1, \ldots, k_m, \mathcal{R}_m)\Big)\\
&= \frac{1}{\prod_{i=1}^{m} \binom{\gamma_i - 1}{R_i}} \sum_{\mathfrak{Z}^{\mathscr{R}}} P(X_{k_1:n} \le x_1, X_{k_2:n} \le x_2, \ldots, X_{k_m:n} \le x_m),
\end{aligned}$$

where the summation is over $\mathfrak{Z}^{\mathscr{R}}$ defined in (10.2). □

It is worth mentioning that Theorem 10.1.1 does not impose any assumptions on the distribution of the original sample $X_1, \ldots, X_n$. Thus, these assumptions are implicitly included in the distribution of the marginal order statistics $(X_{k_1:n}, \ldots, X_{k_m:n})$, $\{k_1, \ldots, k_m\} \subseteq \{1, \ldots, n\}$.

In the IID case, the mixture representation simplifies since the probabilities in (10.1) depend only on $(k_1, \ldots, k_m)$ but not on $\big(\mathcal{R}_1, \ldots, \mathcal{R}_m\big)$. Using the probability mass function of $(K_1, \ldots, K_m)$ given in (10.9), we find the mixture representation (see Thomas and Wilson [843], Guilbaud [419], and Cramer and Lenz [303])

$$P(\mathbf{X}^{\mathscr{R}} \leq \mathbf{x}_m) = \sum_{1=k_1<k_2<\cdots<k_m\leq n-R_m} \omega(k_2, \ldots, k_m) P(X_{k_1:n} \leq x_1, \ldots, X_{k_m:n} \leq x_m),$$

where $\omega(k_2, \ldots, k_m) = \prod_{j=2}^m \frac{\binom{n-k_j}{n-\gamma_j+1-k_j}}{\binom{n-k_{j-1}}{n-\gamma_j-k_{j-1}}}$ and $X_{1:n}, X_{k_2:n}, \ldots, X_{k_m:n}$, $1 < k_2 < \cdots < k_m \leq n - R_m$, are order statistics from an IID sample $X_1, \ldots, X_n$. Guilbaud [418, 419] proposed algorithms to compute the marginal distributions which are applied in the construction of nonparametric confidence intervals, prediction intervals, and tolerance intervals (see Sects. 17.1.5, 17.4.1, and 17.5).

10.2 Joint Density Function of Progressively Type-II Censored Order Statistics

In Model 1.2.2, let $F_1, \ldots, F_n$ be absolutely continuous cumulative distribution functions with density functions $f_1, \ldots, f_n$, respectively. Then, $P^{\mathbf{X}^{\mathscr{R}}}$ has a density function $f^{\mathbf{X}^{\mathscr{R}}}$ as given in the following theorem.

Theorem 10.2.1 (Balakrishnan and Cramer [93]). For $n \in \mathbb{N}$, let $\mathfrak{S}_n$ be the set of all permutations π of $(1, \ldots, n)$. For brevity, let $R_{\bullet r} = \sum_{j=1}^r R_j$, $1 \leq r \leq m$, with $R_{\bullet 0} = 0$ and $R_{\bullet m} = n - m$. Then, the joint density function of $X_{1:m:n}, \ldots, X_{m:m:n}$ is given by

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \frac{1}{(n-1)!}\left(\prod_{j=2}^m \gamma_j\right) \sum_{\pi\in\mathfrak{S}_n} \prod_{j=1}^m f_{\pi(j)}(t_j) \times \left\{ \prod_{r=m+R_{\bullet j-1}+1}^{m+R_{\bullet j}} \overline{F}_{\pi(r)}(t_j) \right\}, \quad t_1 \leq \cdots \leq t_m, \quad (10.3)$$

where $\pi(i)$ is the ith component of the permutation vector $\pi \in \mathfrak{S}_n$, $1 \leq i \leq n$.

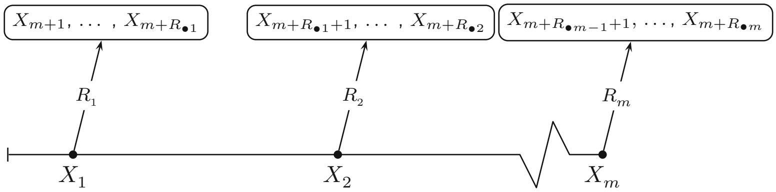

Fig. 10.1 Generation process of progressively Type-II censored order statistics with certain outcome and corresponding lifetimes

Proof. To prove this result, we consider the joint cumulative distribution function

$$P\left(X_{1:m:n}^{\mathscr{R}} \le x_1, \ldots, X_{m:m:n}^{\mathscr{R}} \le x_m\right). \tag{10.4}$$

Without loss of generality, we assume that $x_1 \le \cdots \le x_m$ holds in (10.4).

First, taking into account the construction process 1.1.3 of progressively Type-II censored order statistics, we assume a certain outcome of this procedure. Namely, we assign the failures to the lifetimes $X_1, \ldots, X_m$, i.e., $X_{1:m:n}^{\mathscr{R}} = X_1, \ldots, X_{m:m:n}^{\mathscr{R}} = X_m$, and the lifetimes of the progressively censored units as in Fig. 10.1. This means that, before the rth failure, the units with indices $r, \ldots, m$ and $m + R_{\bullet r-1} + 1, \ldots, n$ are still in the experiment. Moreover, the previous setup fixes that the rth failure is assigned to the unit number r, $1 \le r \le m$. In order to simplify the notation, we introduce the random vectors $\mathbf{W}_r = (X_{m+R_{\bullet r-1}+1}, \ldots, X_{m+R_{\bullet r}})$, $1 \le r \le m$. The components of $\mathbf{W}_r$ represent the lifetimes of those units which are progressively censored immediately after the rth failure. Using this notation, we can write

$$\begin{aligned}\min\{X_{r+1}, \ldots, X_m, X_{m+R_{\bullet r-1}+1}, \ldots, X_n\}& \\ &= \min\{X_{r+1}, \ldots, X_m, \mathbf{W}_r, \ldots, \mathbf{W}_m\}.\end{aligned}$$

Then,

$$P\left(X_r \le \min\{X_{r+1}, \ldots, X_m, \mathbf{W}_r, \ldots, \mathbf{W}_m\}, X_r \le x_r, 1 \le r \le m\right) \tag{10.5}$$

denotes the probability that the random variables X_r represent the failure times (which are supposed to be less than x_r) and that the units corresponding to the components of $\mathbf{W}_r$ are progressively censored (after the rth failure), $1 \le r \le m$.

The probability in (10.5) can be calculated as follows:

$$\begin{aligned}&P\left(X_r \le \min\{X_{r+1}, \ldots, X_m, \mathbf{W}_r, \ldots, \mathbf{W}_m\}, X_r \le x_r, 1 \le r \le m\right)\\ &\quad= \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} P\left(t_r \le \min\{t_{r+1}, \ldots, t_m, \mathbf{W}_r, \ldots, \mathbf{W}_m\}, 1 \le r \le m\right)\\ &\qquad\qquad \times \prod_{j=1}^{m} f_j(t_j)\, dt_m \ldots dt_1\end{aligned}$$

$$= \int_{-\infty}^{x_1} \int_{t_1}^{x_2} \cdots \int_{t_{m-1}}^{x_m} P\left(t_r \le \min\left\{\mathbf{W}_r, \ldots, \mathbf{W}_m\right\}, 1 \le r \le m\right) \times \prod_{j=1}^{m} f_j(t_j)\, dt_m \ldots dt_1. \quad (10.6)$$

Now, let us consider the probability term in the integrand in (10.6). Using the definition of the minimum and taking into account the ordering $t_1 \le \cdots \le t_m$, we obtain the expression

$$\begin{aligned} &P\left(t_r \le \min\left\{\mathbf{W}_r, \ldots, \mathbf{W}_m\right\}, 1 \le r \le m\right) \\ &\quad = P\big(t_r \le \min\{\mathbf{W}_r\}, 1 \le r \le m\big) \\ &\quad = P\left(t_1 \le \min\{\mathbf{W}_1\}\right) P\left(t_2 \le \min\{\mathbf{W}_2\}\right) \cdots P\left(t_m \le \min\{\mathbf{W}_m\}\right) \\ &\quad = \prod_{j=1}^{m} \prod_{r=m+R_{\bullet j-1}+1}^{m+R_{\bullet j}} \overline{F}_r(t_j). \end{aligned}$$

Hence, we obtain

$$\begin{aligned} &P\left(X_r \le \min\left\{X_{r+1}, \ldots, X_m, \mathbf{W}_r, \ldots, \mathbf{W}_m\right\}, X_r \le x_r, 1 \le r \le m\right) \\ &\quad = \int_{-\infty}^{x_1} \int_{t_1}^{x_2} \cdots \int_{t_{m-1}}^{x_m} \prod_{j=1}^{m} f_j(t_j) \left\{ \prod_{r=m+R_{\bullet j-1}+1}^{m+R_{\bullet j}} \overline{F}_r(t_j) \right\} dt_m \ldots dt_1. \end{aligned}$$

Differentiation of (10.6) w.r.t. $x_1, \ldots, x_m$ yields the function

$$h_I(\mathbf{x}_m) = \prod_{j=1}^{m} f_j(x_j) \left\{ \prod_{r=m+R_{\bullet j-1}+1}^{m+R_{\bullet j}} \overline{F}_r(x_j) \right\}, \quad x_1 \le \cdots \le x_m,$$

where $I = (1, \ldots, n)$. Choosing a permutation $\pi(I) = (\pi(1), \ldots, \pi(n))$ of $(1, \ldots, n)$, this leads to the expression

$$h_{\pi(I)}(\mathbf{x}_m) = \prod_{j=1}^{m} f_{\pi(j)}(x_j) \left\{ \prod_{r=m+R_{\bullet j-1}+1}^{m+R_{\bullet j}} \overline{F}_{\pi(r)}(x_j) \right\}, \quad x_1 \le \cdots \le x_m.$$

In the next step, we have to take into account that those units that are censored at the rth failure are censored at random. The procedure works as follows. First, we specify a number i_1 out of $n = \gamma_1$ units and assign this number to the first failure. Then, from the remaining numbers, we choose randomly R_1 values $i_2, \ldots, i_{R_1+1}$ out of $\gamma_1 - 1$ (with ordering!) and remove the associated units from the experiment. The corresponding probability is $\frac{(\gamma_1 - R_1 - 1)!}{(\gamma_1 - 1)!} = \frac{\gamma_2!}{(\gamma_1 - 1)!}$. Then, we choose a new failure time $X_{i_{R_1+2}}$ out of γ_2 possible random variables and so on. Continuing this process, we obtain a permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$, i.e, $(i_1, \ldots, i_n) = \pi(I)$.

The probability to choose a specific permutation π leading to the previous outcome is then given by

$$K = \frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j .$$

Denoting by the event A_π the assignment of the permutation π to the unit indices, we obtain

$$\begin{aligned} P\left(\mathbf{X}^{\mathscr{R}} \leq \mathbf{x}_m\right) &= \sum_{\pi \in \mathfrak{S}_n} P\left(\mathbf{X}^{\mathscr{R}} \leq \mathbf{x}_m \mid A_\pi\right) \, P(A_\pi) \\ &= K \sum_{\pi \in \mathfrak{S}_n} P\left(X_{\pi(r)} \leq \min\left\{X_{\pi(r+1)}, \ldots, X_{\pi(m)}, \mathbf{W}_r^\pi, \ldots, \mathbf{W}_m^\pi\right\},\right. \\ &\qquad\qquad \left. X_{\pi(r)} \leq x_r, 1 \leq r \leq m\right), \end{aligned}$$

where $\mathbf{W}_j^\pi = \left(X_{\pi(m+R_{\bullet j-1}+1)}, \ldots, X_{\pi(m+R_{\bullet j})}\right)$, $1 \leq j \leq m$. By construction (see Assumption 1.2.6), the event A_π is independent of the random variables $X_1, \ldots, X_n$ so that the condition can be omitted after specifying the associated outcome of the progressive censoring procedure. Any term in the above sum is of the form in (10.5). In order to apply the preceding results, the specific permutation of the indices only needs to be taken into account. Therefore, differentiation of the preceding expression yields

$$\frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j \sum_{\pi \in \mathfrak{S}_n} h_{\pi(I)}(\mathbf{x}_m),$$

which is the joint density function of $X_{1:m:n}, \ldots, X_{m:m:n}$ presented in Eq. (10.3). □

Assuming $F_1 = \cdots = F_n$, (10.3) yield directly the joint density function of progressively Type-II censored order statistics in the IID case, i.e., $f^{\mathbf{X}^{\mathscr{R}}}$ has the representation (2.4). Hence, the proof of Theorem 10.2.1 also provides an alternate proof for the joint density function in the IID case.

Modeling of Outliers

The case when $F_1 = \cdots = F_{n-1} = F$ and $F_n = G$ is of special interest in the modeling of a single outlier (see Barnett and Lewis [171]). With f and g as the corresponding densities, the joint density function of $X_{1:m:n}, \ldots, X_{m:m:n}$ in (10.3) simplifies considerably. For $t_1 \leq \cdots \leq t_m$, Balakrishnan and Cramer [93] showed that

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \left(\prod_{j=2}^{m} \gamma_j\right) \sum_{\nu=1}^{m} \left\{g(t_\nu)\overline{F}(t_\nu) + R_\nu f(t_\nu)\overline{G}(t_\nu)\right\} \left\{\overline{F}(t_\nu)\right\}^{R_\nu - 1} \times \prod_{\substack{j=1 \\ j\neq\nu}}^{m} f(t_j) \left\{\overline{F}(t_j)\right\}^{R_j}. \quad (10.7)$$

For $m = n$ and $\mathscr{R} = (0^{*n})$, Eq. (10.7) reduces to the well-known formula for density functions of order statistics from a single-outlier model presented, for example, in Kale and Sinha [491] and Joshi [485] (see also Arnold and Balakrishnan [51]).

Connection to Permanents

Vaughan and Venables [874] established a connection of distributions of order statistics from INID random variables to permanents. They pointed out that the joint density function of order statistics $X_{1:n}, \ldots, X_{n:n}$ from a sample $X_1, \ldots, X_n$ of independent random variables with cumulative distribution functions F_i and density functions f_i, $i = 1, \ldots, n$, can be written as the permanent of the matrix

$$\begin{pmatrix} f_1(t_1) & \cdots & f_n(t_1) \\ f_1(t_2) & \cdots & f_n(t_2) \\ \vdots & \ddots & \vdots \\ f_1(t_n) & \cdots & f_n(t_n) \end{pmatrix}, \qquad t_1 \leq \cdots \leq t_n.$$

Similar expressions hold for all marginal density and cumulative distribution functions. In particular, the one-dimensional marginal cumulative distribution function of the rth order statistic can be expressed as a weighted sum of permanents of matrices

$$B_i = \begin{pmatrix} (F_1(t), \ldots, F_n(t)) \otimes \mathbb{1}_i \\ (\overline{F}_1(t), \ldots, \overline{F}_n(t)) \otimes \mathbb{1}_{n-i} \end{pmatrix}, \quad t \in \mathbb{R},$$

where the notation $\otimes$ denotes the Kronecker product of matrices/vectors. Notice that the permanent $\mathsf{per}(A)$ of a square matrix $A = (a_{ij})_{i,j} \in \mathbb{R}^{n\times n}$ is defined by

$$\mathsf{per}(A) = \sum_{\pi\in\mathfrak{S}_n} \prod_{i=1}^{n} a_{i\pi(i)}.$$

As pointed out by Balakrishnan and Cramer [93], Theorem 10.2.1 yields a representation of the joint density function in terms of a permanent

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j \operatorname{per}\begin{pmatrix} (f_1(t_1),\ldots,f_n(t_1)) \otimes \mathbb{1}_1 \\ (\overline{F}_1(t_1),\ldots,\overline{F}_n(t_1)) \otimes \mathbb{1}_{R_1} \\ \vdots \\ (f_1(t_m),\ldots,f_n(t_m)) \otimes \mathbb{1}_1 \\ (\overline{F}_1(t_m),\ldots,\overline{F}_n(t_m)) \otimes \mathbb{1}_{R_m} \end{pmatrix}.$$

Cramer et al. [314] obtained an alternative expression to (10.3) based on an expansion of permanents due to Ryser [766, p. 27] (see also Bebiano [183] and Minc [650, p. 124]). Notice that Ryser's method is known to be the most efficient way of computing a permanent. The procedure gets by with $(n-1)(2^n-1)$ multiplications which is the best known value for the evaluation of permanents (see, e.g., Minc [650, p. 126], Kräuter [549], Liang et al. [590]).

Theorem 10.2.2. Let $F_S = \frac{1}{|S|}\sum_{\alpha\in S} F_\alpha$. The joint density function of progressively Type-II censored order statistics $X_{1:m:n},\ldots,X_{m:m:n}$, with censoring scheme $\mathscr{R} = (R_1,\ldots,R_m)$ from n INID random variables $X_1,\ldots,X_n$, is given by

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \frac{1}{n!}\sum_{k=0}^{n-1}(-1)^k(n-k)^n \sum_{|S|=n-k}\Big(\prod_{j=1}^{m}\gamma_j\Big)\left(\prod_{i=1}^{m} f_S(t_i)(1-F_S(t_i))^{R_i}\right),$$

where $t_1 \le \cdots \le t_m$.

Thus, the result of Guilbaud [417] extends to the case of progressive censoring from INID variables, viz., that the progressively Type-II censored order statistics from INID variables can be seen as a generalized mixture of progressively Type-II censored order statistics from a population with cumulative distribution function F_S and censoring scheme $\mathscr{R}$. In particular, this representation illustrates that an explicit expression for the joint cumulative distribution function can be easily obtained via integration. In fact, the "mixture" probabilities correspond to those for INID order statistics. The representation presented in Theorem 10.2.2 has also been noted by Guilbaud [420] in his discussion of the paper by Balakrishnan [84].

Applications to Stochastic Orderings

Let $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$ be two samples of INID random variables with cumulative distribution functions $F_1,\ldots,F_n$ and $G_1,\ldots,G_n$, respectively. Then, Fischer et al. [371] found the following basic result which follows from a direct application of the mixture representation.

Theorem 10.2.3. Let $\mathbf{X}_{1,n:n} = (X_{1:n}, \dots, X_{n:n})'$ and $\mathbf{Y}_{1,n:n} = (Y_{1:n}, \dots, Y_{n:n})'$. Then, for any censoring scheme $\mathscr{R}$, stochastic ordering of the samples of order statistics implies stochastic ordering of the two samples of progressively Type-II censored order statistics, i.e.,

$$\mathbf{X}_{1,n:n} \geq_{\mathsf{st}} \mathbf{Y}_{1,n:n} \quad \Longrightarrow \quad \mathbf{X}^{\mathscr{R}} \geq_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}.$$

A direct application of this result yields an extension of a result for order statistics shown by Hu [453, 454]. They considered cumulative distribution functions of the type

$$\overline{F}_i(t) = \mathsf{e}^{-H(\lambda_i t)}, \qquad \overline{G}_i(t) = \mathsf{e}^{-H(\mu_i t)}, \quad t \geq 0, 1 \leq i \leq n,$$

where $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n > 0$ and H is a function defined by

$$H(t) = \int_0^t h(s)ds, \quad t \geq 0,$$

with a nonnegative integrable function h such that $\lim_{t\to\infty} H(t) = \infty$. Moreover, suppose h is decreasing as well as that the function h^* defined by $h^*(x) = xh(x)$ is increasing. Then, $\boldsymbol{\lambda}_n \succcurlyeq_{\mathsf{w}} \boldsymbol{\mu}_n$ implies $\mathbf{X}^{\mathscr{R}} \geq_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}$. The result can be directly applied to Weibull and gamma families as shown in Fischer et al. [371]. For order statistics, we refer to Hu [453, 454], Khaledi and Kochar [523], and Lihong and Xinsheng [596]). Moreover, an application to proportional hazards distributions is possible. It extends results of Proschan and Sethuraman [731] (for the univariate stochastic order, see also Boland et al. [212] and Pledger and Proschan [723]).

A family of cumulative distribution functions $\{F_1, \dots, F_n\}$ is said to have proportional hazard functions if for all $t \in \mathbb{R}$,

$$\overline{F}_i(t) = \mathsf{e}^{-\lambda_i H(t)}, \quad i = 1, 2, \dots, n,$$

where $\lambda_1, \dots, \lambda_n > 0$ and $H : \mathbb{R} \longrightarrow \mathbb{R}$ is a nondecreasing function with $H(t) = 0$ for all $t < 0$ and $\lim_{t\to\infty} H(t) = \infty$. The parameters $\lambda_1, \dots, \lambda_n$ are called constants of proportionality for the common hazard function H.

Theorem 10.2.4. Suppose $\mathbf{X}^{\mathscr{R}}$ and $\mathbf{Y}^{\mathscr{R}}$ are INID progressively Type-II censored order statistics based on $F_1, \dots, F_n$ and $G_1, \dots, G_n$, respectively, where

(1) $F_1, \dots, F_n$ have proportional hazard functions with constants of proportionality $\lambda_1, \dots, \lambda_n > 0$ for the common hazard function H,
(2) $G_1, \dots, G_n$ have proportional hazard functions with constants of proportionality $\mu_1, \dots, \mu_n > 0$ for the common hazard function H, and
(3) $\boldsymbol{\lambda}_n \succcurlyeq_{\mathsf{w}} \boldsymbol{\mu}_n$.

Then, $\mathbf{X}^{\mathscr{R}} \geq_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}$.

The preceding result can be directly applied to extend a result of Kochar and Korwar [540] on stochastic comparisons of generalized spacings from INID exponential distributions (see also Kochar and Rojo [541]). Using the notation introduced in Sect. 2.5.2, we get for the vector of $(r, 1)$-spacings $S^{\star\mathscr{R}}_{r,1} = X^{\mathscr{R}}_{r:m:n} - X^{\mathscr{R}}_{1:m:n}$, $r = 2, \ldots, m$, (cf. (2.38)) the following result.

Corollary 10.2.5. Let $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ be independent exponential random variables such that X_i has hazard rate λ_i and Y_i has hazard rate μ_i, $i = 1, \ldots, n$. Then, $\boldsymbol{\lambda}_n \succcurlyeq_{\mathsf{w}} \boldsymbol{\mu}_n$ implies

$$(S^{\star\mathscr{R}}_{X;r,1})_{r=2,\ldots,m} \geq_{\mathsf{st}} (S^{\star\mathscr{R}}_{Y;r,1})_{r=2,\ldots,m}.$$

If $Y_1, \ldots, Y_n$ is an IID sample with hazard rate $\overline{\lambda} = \frac{1}{n}\sum_{j=1}^{n} \lambda_j$, the preceding results show that $\mathbf{X}^{\mathscr{R}} \geq_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}$ since $\boldsymbol{\lambda}_n \succcurlyeq_{\mathsf{w}} \overline{\lambda}\mathbb{1}_n$. Moreover, Fischer et al. [371] pointed out that $\overline{\lambda}$ can be replaced by any value $\mu \in [\overline{\lambda}, \infty)$. For order statistics, such comparisons have been addressed by Ball [165], Barbour et al. [166], Li and Shaked [586], and Ma [624]. Under the assumption that $Y_1, \ldots, Y_n$ is an IID sample, Ma [624] showed that the stochastic order of the order statistics is determined by the stochastic ordering of the minimum and maximum of the samples. Thus, Theorem 10.2.3 can be written as

$$X_{1:n} \geq_{\mathsf{st}} Y_{1:n} \Longrightarrow \mathbf{X}^{\mathscr{R}} \geq_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}, \quad X_{n:n} \leq_{\mathsf{st}} Y_{n:n} \Longrightarrow \mathbf{X}^{\mathscr{R}} \leq_{\mathsf{st}} \mathbf{Y}^{\mathscr{R}}.$$

Mao and Hu [637] discussed likelihood ratio ordering of INID progressively Type-II censored order statistics and obtained the following result.

Theorem 10.2.6. Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from an INID sample $X_1, \ldots, X_n$ and $\mathscr{R} = (R_1, \ldots, R_m)$ be a censoring scheme with decreasing censoring numbers $R_1 \geq R_2 \geq \cdots \geq R_m$. Then, if a permutation $\pi \in \mathfrak{S}_n$ exists with $X_{\pi(1)} \leq_{\mathsf{lr}} \ldots \leq_{\mathsf{lr}} X_{\pi(n)}$, the progressively Type-II censored order statistics are likelihood ratio ordered, i.e.,

$$X_{i:m:n} \leq_{\mathsf{lr}} X_{j:m:n}, \quad 1 \leq i < j \leq m.$$

Finally, Fischer et al. [371] established an extension of a result of Sen [790] which we present for completion.

Theorem 10.2.7. Let $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ be INID progressively Type-II censored order statistics based on $F_1, \ldots, F_n$ and a censoring scheme $\mathscr{R} \in \mathscr{C}^m_{m,n}$ and $Y^{\mathscr{R}}_{1:m:n}, \ldots, Y^{\mathscr{R}}_{m:m:n}$ be progressively Type-II censored order statistics from the cumulative distribution function

$$F^* = \frac{1}{n}\sum_{k=1}^{n} F_k.$$

Further, let

$$\alpha_i = i - 1, \quad \beta_i = n - \gamma_i + 1 = i + \sum_{j=1}^{i-1} R_j, \quad i \in \{2, \dots, m\},$$

and $\xi_p = \inf\{u \in \mathbb{R} : F^*(u) \geq p\}$ be the pth quantile of F^*, $p \in (0, 1)$. Then, for $i \in \{2, \dots, m-1\}$ and $x \leq \xi_{\frac{\alpha_i}{n}} < \xi_{\frac{\beta_i}{n}} \leq y$,

$$P\big(x < X^{\mathscr{R}}_{i:m:n} \leq y\big) \;\geq\; P\big(x < Y^{\mathscr{R}}_{i:m:n} \leq y\big),$$

where equality holds if $F_k(x) = F^*(x)$ and $F_k(y) = F^*(y)$ for all $k \in \{1, \dots, n\}$.

If $R_m > 0$, then

$$P\big(x < X^{\mathscr{R}}_{m:m:n} \leq y\big) \;\geq\; P\big(x < Y^{\mathscr{R}}_{m:m:n} \leq y\big) \quad \forall\ x \leq \xi_{\frac{\alpha_m}{n}} < \xi_{\frac{\beta_m}{n}} \leq y,$$

with equality holding if $F_k(x) = F^*(x)$ and $F_k(y) = F^*(y)$ for all $k \in \{1, \dots, n\}$.

Furthermore, for $x \in \mathbb{R}$,

$$P\big(X^{\mathscr{R}}_{1:m:n} \leq x\big) \;\geq\; P\big(Y^{\mathscr{R}}_{1:m:n} \leq x\big)$$

with equality holding if $F_k(x) = F^*(x)$ for all $k \in \{1, \dots, n\}$.

10.3 Dependence Structure of INID Progressively Type-II Censored Order Statistics

The dependence structure of INID progressively Type-II censored order statistics has been studied in Cramer and Lenz [303] by considering the generation process of progressively Type-II censored order statistics presented in Fischer et al. [371] (see Sect. 1.1.1). Using the notation introduced in Assumption 1.2.6, Cramer and Lenz [303] obtained the following results. The probabilities in (10.8) can also be found in Thomas and Wilson [843].

Proposition 10.3.1. (i) $(K_j, \bigcup_{i=1}^{j} \mathcal{R}_{i,K_i})_{1 \leq j \leq m}$ is a Markov chain with transition probabilities

$$P(K_{j+1} = k_{j+1}, \bigcup_{i=1}^{j+1} \mathcal{R}_{i,K_i} = \bigcup_{i=1}^{j+1} \mathcal{S}_i \;\Big|\; K_l = k_l, \bigcup_{i=1}^{l} \mathcal{R}_{i,K_i} = \bigcup_{i=1}^{l} \mathcal{S}_i, 1 \leq l \leq j)$$

$$= \begin{cases} P(\mathcal{R}_{j+1,k_{j+1}} = \mathcal{S}_{j+1}) = \frac{1}{\binom{\gamma_{j+1}-1}{R_{j+1}}} & \text{if } G(k_j, \bigcup_{i=1}^{j} \mathcal{S}_i) = k_{j+1} \\ 0 & \text{if } G(k_j, \bigcup_{i=1}^{j} \mathcal{S}_i) \neq k_{j+1} \end{cases},$$

where

$$G(k_j, \bigcup_{i=1}^{j} \mathcal{S}_i) = \min\left(\{k_j + 1, \dots, n\} \setminus \bigcup_{i=1}^{j} \mathcal{S}_i\right);$$

(ii) $(K_j)_{1 \le j \le m}$ forms a Markov chain with transition probabilities

$$P(K_j = k | K_{j-1} = \nu) = \frac{\binom{n-k}{n-\gamma_j+1-k}}{\binom{n-\nu}{n-\gamma_j-\nu}}, \quad 2 \le j \le m, \tag{10.8}$$

and $P(K_1 = 1) = 1$;

(iii) In general, $(K_j, \mathcal{R}_{j,K_j})_{1 \le j \le m}$ is not a Markov chain.

In particular, the joint probability mass function of $K_1, \dots, K_m$ can be taken from (10.8) as

$$P(K_j = k_j, j = 1, \dots, m) = \prod_{j=2}^{m} \frac{\binom{n-k_j}{n-\gamma_j+1-k_j}}{\binom{n-k_{j-1}}{n-\gamma_j-k_{j-1}}},$$
$$1 = k_1 < \dots < k_m = n - R_m. \tag{10.9}$$

Dependence properties of (usual) INID order statistics $X_{1:n}, \dots, X_{n:n}$ have been discussed extensively in the literature. The following selection can be found in Boland et al. [211].

Theorem 10.3.2. Let $X_{1:n}, \dots, X_{n:n}$ be order statistics from INID random variables $X_1, \dots, X_n$. Then:

(i) $(X_{1:n}, \dots, X_{n:n})$ is associated;

(ii) For $i, j \in \{1, \dots, n\}$ with $i < j$, $(X_{i:n}, X_{j:n})$ are right tail increasing in sequence;

(iii) Let $X_1, \dots, X_n$ have differentiable density functions (w.r.t. the Lebesgue measure) and proportional hazard functions on an interval $[a, b]$ such that $[a, b]$ is the support of X_j for each $j \in \{1, \dots, n\}$. Then, $(X_{i:n}, X_{j:n})$ is conditionally increasing in sequence for all $j \in \{2, \dots, n\}$;

(iv) In general, $(X_{1:n}, \dots, X_{n:n})$ does not have the right corner set increasing property and, consequently, does not have the MTP_2 property (even if $X_1, \dots, X_n$ have differentiable (Lebesgue) densities with a common interval support).

The last statement implies directly that $\mathbf{X}^{\mathscr{R}}$ does not have the MTP_2 property in general as is true in the IID model (see Theorem 2.5.10). In order to prove association of INID progressively Type-II censored order statistics, Cramer and Lenz [303] proved the following ordering result which implies association of the random indices $(K_1, \ldots, K_m)$.

Proposition 10.3.3. $(K_1, \ldots, K_m)$ is conditionally increasing in sequence (CIS), i.e., for $j \in \mathbb{N}$ with $j < m$, $(k_1, \ldots, k_j), (k_1^*, \ldots, k_j^*) \in \{1, \ldots, n\}^j$ with $k_j \le k_j^*$ and $P(K_1 = k_1, \ldots, K_j = k_j) > 0$, $P(K_1 = k_1^*, \ldots, K_j = k_j^*) > 0$,

$$[K_{j+1} \mid K_1 = k_1, \ldots, K_j = k_j] \le_{\mathrm{st}} [K_{j+1} \mid K_1 = k_1^*, \ldots, K_j = k_j^*].$$

Moreover, $(K_1, \ldots, K_m)$ is associated.

Using that subsamples of associated samples are associated (cf. Müller and Stoyan [659]) and a result on mixtures and association of random variables due to Jogdeo [480], the following result holds.

Theorem 10.3.4. Let $(X_1, \ldots, X_n)$ be associated. Then,

$$(X_{1:m:n}^{\mathscr{R}}, \ldots, X_{m:m:n}^{\mathscr{R}}, K_1, \ldots, K_m)$$

is associated, too. In particular, this yields association of $\mathbf{X}^{\mathscr{R}}$.

Notice that independence of $(X_1, \ldots, X_n)$ is not necessary to establish association of $\mathbf{X}^{\mathscr{R}}$. Therefore, the above result holds for samples exhibiting some dependence between the random variables.

Finally, Cramer and Lenz [303] showed that $X_{j:m:n}$ is right tail increasing in $X_{1:m:n}$, i.e., $P(X_{j:m:n}^{\mathscr{R}} > t | X_{1:m:n} > s)$ is increasing in s for any t. This notion is denoted by $\mathrm{RTI}(X_{j:m:n} | X_{1:m:n})$. Further results in this direction have been established in Mao and Hu [637]. In particular, they extended the above results and proved $\mathrm{RTI}(X_{j:m:n} | X_{i:m:n})$ for $1 \le i < j \le m$. A similar result for the left tail decreasing property has also been shown. These findings extended results of Boland et al. [211] for INID order statistics to progressive censoring.

10.4 Dependence and Copulas

Whilst the preceding approaches relax the assumption of identical distribution, further efforts have been made to introduce dependence in progressive Type-II censoring (see Theorem 10.3.4). Rezapour et al. [752, 753] assumed an Archimedean copula C_ψ to model the dependence in the sample $X_1, \ldots, X_n$, i.e.,

$$C_\psi(u_1, \ldots, u_n) = \psi\left(\sum_{i=1}^{n} \psi^{-1}(u_j)\right), \tag{10.10}$$

where $\psi : [0,\infty) \longrightarrow [0,1]$ is a completely monotone generator function such that $\psi(0) = 1$ and $\lim_{x\to\infty}\psi(x) = 0$ (see McNeil and Nešlehová [643]). With $G(t) = \exp\{-\psi^{-1}(t)\}$, $t \in [0,1]$, and M_ψ be a cumulative distribution function with Laplace transform ψ, (10.10) can be written as

$$C_\psi(u_1,\dots,u_n) = \int_0^\infty \prod_{i=1}^n G^\alpha(u_i)\, dM_\psi(\alpha).$$

Using this expression, the sample is supposed to have the survival function

$$P(\mathbf{X} > \mathbf{x}) = \int_0^\infty \prod_{i=1}^n G^\alpha(\overline{F}_i(x_i))\, dM_\psi(\alpha), \quad \mathbf{x} = (x_1,\dots,x_n).$$

Let F_i have a density function f_i, $i = 1,\dots,n$, and g be the first derivative of G. Then, Rezapour et al. [753] showed that the density function $f^{\mathbf{X}^{\mathscr{R}}}$ can be expressed as a mixture of INID progressively Type-II censored order statistics

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \int_0^\infty f^{\mathbf{X}_\alpha^{\mathscr{R}}}(\mathbf{t}_m) dM_\psi(\alpha), \quad \mathbf{t}_m \in \mathbb{R}^m, \tag{10.11}$$

where $f^{\mathbf{X}_\alpha^{\mathscr{R}}}$ is the joint density function of INID progressively Type-II censored order statistics $\mathbf{X}_\alpha^{\mathscr{R}} = (X_{1:m:n;\alpha},\dots,X_{m:m:n;\alpha})$ with baseline distributions $G_i(\cdot,\alpha)$, defined by

$$G_i(t,\alpha) = 1 - G^\alpha(\overline{F}_i(t)) = \exp(-\alpha\psi^{-1}(\overline{F}_i(t))), \quad i = 1,\dots,n, \alpha \geq 0$$

(see (10.3)). Hence, the density function $f^{\mathbf{X}^{\mathscr{R}}}$ in (10.11) can be seen as a M_ψ-mixture of the distributions of INID progressively Type-II censored order statistics. Moreover, the joint cumulative distribution function has the representation

$$P(\mathbf{X}^{\mathscr{R}} \leq \mathbf{x}_m) = \int_0^\infty P(\mathbf{X}_\alpha^{\mathscr{R}} \leq \mathbf{x}_m) dM_\psi(\alpha), \quad \mathbf{x}_m \in \mathbb{R}^m. \tag{10.12}$$

Moreover, it follows from (10.12) that any marginal distribution is also a mixture of the corresponding INID progressively Type-II censored order statistics. As pointed out by Rezapour et al. [753], the integral in (10.11) can be explicitly calculated as given below.

Corollary 10.4.1. Let $K = \frac{1}{(n-1)!}\prod_{j=2}^m \gamma_j$. Then,

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = K \sum_{\pi\in\mathfrak{S}_n} \prod_{j=1}^m h_{\pi(j)}(t_j)$$
$$\times \psi^{(m)}\Big(\sum_{j=1}^m \Big(\psi^{-1}(\overline{F}_{\pi(j)}(t_j)) + \sum_{r=m+R_{\bullet j-1}+1}^{m+R_{\bullet j}} \psi^{-1}(\overline{F}_{\pi(r)}(t_j))\Big)\Big),$$

where $t_1 \leq \cdots \leq t_m$, $\psi^{(m)}$ denotes the mth derivative of ψ, and $h_i(t) = \frac{f_i(t)}{\psi'(\psi^{-1} \circ \overline{F}_i(t))}$.

In the case of identically distributed random variables $X_1, \dots, X_n$, the density function in (10.11) simplifies to

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \int_0^\infty \Big(\prod_{j=1}^m \gamma_j\Big) \prod_{j=1}^m g(t_j, \alpha) \overline{G}^{R_j}(t_j, \alpha) dM_\psi(\alpha),$$

where $\overline{G}(t, \alpha) = \exp\Big(-\alpha\psi^{-1}\Big(\overline{F}(t)\Big)\Big)$ and $g(t, \alpha) = \frac{\partial}{\partial t} G(t, \alpha)$. Thus, for Archimedean copulas, the distribution of progressively Type-II censored order statistics is a mixture of IID progressively Type-II censored order statistics with population distributions $G(\cdot, \alpha)$, $\alpha \geq 0$. As before, we can get rid of the integral (see Rezapour et al. [753]). Furthermore, it is interesting to see that the marginal distributions have simple expressions in this setting. For instance, the density function and cumulative distribution function are given by (cf. (2.25))

$$f^{X_{r:m:n}}(t) = c_{r-1} \sum_{i=1}^r a_{i,r} \frac{\psi'\Big(\gamma_i \psi^{-1}\big(\bar{F}(t)\big)\Big)}{\psi'\Big(\psi^{-1}\Big(\bar{F}(t)\Big)\Big)} f(t), \quad t \in \mathbb{R},$$

$$F^{X_{r:m:n}}(t) = 1 - c_{r-1} \sum_{i=1}^r \frac{a_{i,r}}{\gamma_i} \psi\Big(\gamma_i \psi^{-1}\Big(\bar{F}(t)\Big)\Big), \quad t \in \mathbb{R}.$$

Similar representations can be established for bivariate marginals (see Rezapour et al. [753]).

Moreover, it is worth mentioning that the triangle and quadruple rules given in (2.46) and (2.47) for the IID case also hold in this model. Additionally, the quantile representation given in Theorem 2.1.1 is also true in the present situation.

Remark 10.4.2. If we choose $\psi(u) = \mathrm{e}^{-u}$, $u \geq 0$, the inverse function is given by $\psi^{-1}(t) = -\log t$, $t \in (0, 1]$. Then, the copula in (10.10) is the independence copula yielding the IID case. Hence, $G(t) = t$, $t \in (0, 1]$, and M_ψ is the cumulative distribution function of a one-point distribution in $\alpha = 1$. In this situation, (10.11) reduces to (10.3).

10.5 Progressive Type-II Censoring for Multivariate Observations

Considering p-dimensional random vectors $\mathbf{X}$ and $\mathbf{Y}$ and a continuous function $\Pi : \mathbb{R}^p \longrightarrow \mathbb{R}$ such that

$$\lambda^p\big(\{\mathbf{x} \mid \Pi(\mathbf{x}) = z\}\big) = 0 \quad \text{for all } z \in \mathbb{R}, \tag{10.13}$$

conditionally ordered multivariate variables can be defined by

$$\mathbf{X} \preccurlyeq_\Pi \mathbf{Y} \iff \Pi(\mathbf{X}) \leq \Pi(\mathbf{Y})$$

(see, e.g., [60, 76, 79], [750, p. 66], for further details and interpretations). In particular, the function Π induces a partition of $\mathbb{R}^p$. Assuming an IID sample $\mathbf{X}_i \sim f^{\mathbf{X}}$, $1 \leq i \leq n$, this concept has been applied to define conditionally Π-ordered statistics $\mathbf{X}_{\Pi;1:n} \preccurlyeq_\Pi \cdots \preccurlyeq_\Pi \mathbf{X}_{\Pi;n:n}$ by Bairamov [76] using the order statistics $\Pi_{1:n}(\mathbf{X}), \ldots, \Pi_{n:n}(\mathbf{X})$ of the sample $\Pi(\mathbf{X}_1), \ldots, \Pi(\mathbf{X}_n)$. Notice that condition (10.13) implies that the distribution of $\Pi(\mathbf{X})$ is (absolutely) continuous.

Since $\Pi(\mathbf{X}_1), \ldots, \Pi(\mathbf{X}_n)$ are IID random variables with cumulative distribution function $F_\Pi(t) = P(\Pi(\mathbf{X}) \leq t)$, the density function of $\mathbf{X}_{\Pi;j:n}$ is given by

$$f^{\mathbf{X}_{\Pi;j:n}}(\mathbf{x}) = n\binom{n-1}{j-1} F_\Pi^{j-1}(\Pi(\mathbf{x}))\big(1 - F_\Pi(\Pi(\mathbf{x}))\big)^{n-j} f^{\mathbf{X}}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^p$$

(see [60, 76]), where $F_\Pi(\Pi(\mathbf{x}))$, $x \in \mathbb{R}^p$, is called the structural function. Arnold et al. [60] pointed out that the ordering can be viewed as a concomitant based ordering.

In a similar manner, Bairamov [76] defined progressively censored Π-ordered statistics $\mathbf{X}_{\Pi;1:m:n} \preccurlyeq_\Pi \cdots \preccurlyeq_\Pi \mathbf{X}_{\Pi;m:m:n}$ by applying the progressive censoring procedure to the sample $\Pi(\mathbf{X}_1), \ldots, \Pi(\mathbf{X}_n)$. This results in progressively Type-II censored order statistics $\Pi_{1:m:n}(\mathbf{X}), \ldots, \Pi_{m:m:n}(\mathbf{X})$ and induces the desired progressively censored multivariate sample. As a result, similar expressions can be established for density functions, moments, etc. as in the univariate case.

Part II
Inference

Chapter 11
Linear Estimation in Progressive Type-II Censoring

Linear inference for parametric distributions with ordered data attracted great attention since the class of *L-statistics* $\sum_{i=1}^{m} c_i X_i$ has many applications. Moreover, L-statistics have a simple and nice structure. In particular, least-squares estimation has been widely discussed (see, for instance, Arnold et al. [58, 59] and David and Nagaraja [327, Sect. 8.4]) since it leads to explicit expressions of the estimators provided that the single and product moments can be computed easily (see Chap. 7). For progressive Type-II censoring, an extensive treatment of this topic is presented in Balakrishnan and Aggarwala [86, Chap. 6] reflecting the state of research in 2000. Since the publication of this monograph, many results have been established in this direction. In the following, we summarize these results and present some applications of L-statistics in progressive Type-II censoring.

11.1 Preliminaries

In this chapter, we consider linear estimation of parameters. In particular, we assume that the progressively Type-II censored order statistics are based on location–scale families of distributions

$$\mathscr{F}_l = \{F(\cdot - \mu) | \mu \in \mathbb{R}\}, \tag{11.1a}$$

$$\mathscr{F}_s = \{F(\cdot/\vartheta) | \vartheta > 0\}, \tag{11.1b}$$

$$\mathscr{F}_{ls} = \left\{F\left(\frac{\cdot - \mu}{\vartheta}\right) \middle| \mu \in \mathbb{R}, \vartheta > 0\right\}, \tag{11.1c}$$

respectively, where F is a given continuous distribution. In these cases, the random variables form a linear model. We illustrate this approach for the location–scale model $\mathscr{F}_{ls}$ as the corresponding results for $\mathscr{F}_l$ and $\mathscr{F}_s$ are quite similar. Let $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})'$ be the random vector of progressively Type-II censored order

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_11,

statistics based on $F \in \mathscr{F}_{ls}$ with location parameter μ and scale parameter ϑ. Then, we define $\mathbf{Y}^{\mathscr{R}} = (Y_{1:m:n}, \ldots, Y_{m:m:n})'$ by

$$Y_{j:m:n} = \frac{X_{j:m:n} - \mu}{\vartheta}, \quad j = 1, \ldots, m.$$

Introducing $\mathbf{W}^{\mathscr{R}} = \vartheta(\mathbf{Y}^{\mathscr{R}} - E\mathbf{Y}^{\mathscr{R}})$ and $\mathbf{b} = E\mathbf{Y}^{\mathscr{R}}$, we arrive at the linear model

$$\mathbf{X}^{\mathscr{R}} = \mu \cdot \mathbb{1} + \vartheta \mathbf{Y}^{\mathscr{R}} = \mu \cdot \mathbb{1} + \vartheta E\mathbf{Y}^{\mathscr{R}} + \mathbf{W}^{\mathscr{R}} = [\mathbb{1}, \mathbf{b}]\begin{pmatrix}\mu \\ \vartheta\end{pmatrix} + \mathbf{W}^{\mathscr{R}} = B\boldsymbol{\theta} + \mathbf{W}^{\mathscr{R}},$$

where $E\mathbf{W}^{\mathscr{R}} = 0$, $\mathsf{Cov}(\mathbf{W}^{\mathscr{R}}) = \vartheta^2\Sigma$, $\Sigma = \mathsf{Cov}(\mathbf{Y}^{\mathscr{R}})$, $B = [\mathbb{1}, \mathbf{b}]$ is the known design matrix, and $\boldsymbol{\theta} = (\mu, \vartheta)'$ is the (unknown) parameter vector. Notice that the distribution of $\mathbf{Y}^{\mathscr{R}}$ is parameter-free, and it depends only on the standard member F.

Remark 11.1.1. The above representations are based on the assumption that the complete progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$ has been observed. However, by deleting the respective lines (and rows) in both the mean vector $E\mathbf{Y}^{\mathscr{R}}$ and the covariance matrix Σ, the subsequent results can be applied also to a multiply censored sample $X_{j_1:m:n}, \ldots, X_{j_r:m:n}$ with $1 \le j_1 < \cdots < j_r \le m$.

This observation is of particular interest in the context of general progressive censoring which can be seen as a left censored progressively censored sample (see p. 10). In this case, we have to delete the first r lines and rows, respectively.

11.1.1 Least-Squares Estimation

Least-squares estimators for $\boldsymbol{\theta}$ in the linear model $\mathbf{X}^{\mathscr{R}} = B\boldsymbol{\theta} + \mathbf{W}^{\mathscr{R}}$ are obtained by minimizing the squared (Mahalanobis) distance

$$Q(\boldsymbol{\theta}) = (\mathbf{X}^{\mathscr{R}} - B\boldsymbol{\theta})'\Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - B\boldsymbol{\theta}) = \left\|\mathbf{X}^{\mathscr{R}} - B\boldsymbol{\theta}\right\|_{\Sigma}^2$$

w.r.t. the parameter $\boldsymbol{\theta}$ for a given sample $\mathbf{X}^{\mathscr{R}}$, where Σ is assumed to be regular. It is well known (see, e.g., Christensen [264]) that the resulting least-squares estimator of $\boldsymbol{\theta}$ is given by

$$\widehat{\boldsymbol{\theta}} = (B'\Sigma^{-1}B)^{-1}B'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}. \tag{11.2}$$

Moreover, the variance–covariance matrix of $\widehat{\boldsymbol{\theta}}$ is given by

$$\mathsf{Cov}(\widehat{\boldsymbol{\theta}}) = \vartheta^2(B'\Sigma^{-1}B)^{-1}.$$

This shows that the least-squares estimator can be derived when the covariance matrix $\Sigma = \mathsf{Cov}(\mathbf{Y}^{\mathscr{R}})$ can be calculated. Explicit representations result if the inverse matrix of Σ has a nice structure. It has to be mentioned that absolute continuity of the distribution $P^{\mathbf{Y}^{\mathscr{R}}}$ (or of the cumulative distribution function F) ensures a non-singular covariance matrix Σ.

Given representation (11.2), the Gauß–Markov theorem yields the best linear unbiased estimators (BLUE) $\widehat{\mu}$ and $\widehat{\vartheta}$.

Similar representations result in location or scale families where we assume $\vartheta = 1$ and $\mu = 0$, respectively. In particular, we have the linear models

$$\mathbf{X}^{\mathscr{R}} - \mathbf{b} = \mu \cdot \mathbb{1} + \mathbf{W}^{\mathscr{R}},$$
$$\mathbf{X}^{\mathscr{R}} = \vartheta E\mathbf{Y}^{\mathscr{R}} + \mathbf{W}^{\mathscr{R}},$$

where $\mathbf{b} = E\mathbf{Y}^{\mathscr{R}}$. The representations of the least-squares estimators follow from (11.2) with $B = \mathbb{1}$ and $B = E\mathbf{Y}^{\mathscr{R}} = \mathbf{b}$, respectively. In particular, we find the following expressions for the best linear unbiased estimators.

Location model $\mathscr{F}_l$

The BLUE of μ is given by

$$\widehat{\mu}_{\mathsf{LU}} = \frac{1}{\mathbb{1}'\Sigma^{-1}\mathbb{1}}\mathbb{1}'\Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - \mathbf{b})$$

with variance $\mathsf{Var}(\widehat{\mu}_{\mathsf{LU}}) = \frac{\vartheta^2}{\mathbb{1}'\Sigma^{-1}\mathbb{1}}$.

Scale model $\mathscr{F}_s$

The BLUE of ϑ is given by

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{\mathbf{b}'\Sigma^{-1}\mathbf{b}}\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}$$

with variance $\mathsf{Var}(\widehat{\vartheta}_{\mathsf{LU}}) = \frac{\vartheta^2}{\mathbf{b}'\Sigma^{-1}\mathbf{b}}$.

Location–scale model $\mathscr{F}_{ls}$

Suppose $m \geq 2$. Using the fact that $n = \sum_{j=1}^{m}(R_j + 1)$, the BLUEs $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$ are given by

$$\widehat{\mu}_{\mathsf{LU}} = \frac{1}{\Delta}\big((\mathbf{b}'\Sigma^{-1}\mathbf{b})(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}) - (\mathbb{1}'\Sigma^{-1}\mathbf{b})(\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}})\big), \tag{11.3a}$$

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{\Delta}\big((\mathbb{1}'\Sigma^{-1}\mathbb{1})(\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}) - (\mathbb{1}'\Sigma^{-1}\mathbf{b})(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}})\big), \tag{11.3b}$$

where $\Delta = \mathbb{1}'\Sigma^{-1}\mathbb{1}\mathbf{b}'\Sigma^{-1}\mathbf{b} - (\mathbb{1}'\Sigma^{-1}\mathbf{b})^2$. The variances and covariance of these estimators are given by

$$\mathsf{Var}(\widehat{\mu}_{\mathsf{LU}}) = \frac{\mathbf{b}'\Sigma^{-1}\mathbf{b}}{\Delta}\vartheta^2, \quad \mathsf{Var}(\widehat{\vartheta}_{\mathsf{LU}}) = \frac{\mathbb{1}'\Sigma^{-1}\mathbb{1}}{\Delta}\vartheta^2,$$
$$\mathsf{Cov}(\widehat{\mu}_{\mathsf{LU}}, \widehat{\vartheta}_{\mathsf{LU}}) = -\frac{\mathbf{b}'\Sigma^{-1}\mathbb{1}}{\Delta}\vartheta^2.$$

Remark 11.1.2. (i) It has to be noted that

$$\mathbb{1}'\Sigma^{-1}\mathbb{1}\mathbf{b}'\Sigma^{-1}\mathbf{b} - (\mathbb{1}'\Sigma^{-1}\mathbf{b})^2 > 0$$

provided that $m \geq 2$. This can be seen from the Cauchy–Schwarz inequality which proves that equality holds if $\mathbf{b} = a\mathbb{1}$ for some $a \in \mathbb{R}$. However, this implies $EX_{1:m:n} = \cdots = EX_{m:m:n}$. Since, by definition, $X_{1:m:n} \leq \cdots \leq X_{m:m:n}$, this readily yields $X_{1:m:n} = \cdots = X_{m:m:n}$ almost everywhere. This contradicts the assumption that Σ is non-singular.

(ii) In the location–scale model, the scale estimator $\widehat{\vartheta}_{\mathsf{LU}}$ can be written as

$$\begin{aligned}\widehat{\vartheta}_{\mathsf{LU}} &= \frac{1}{\Delta}\big(\mathbb{1}'\Sigma^{-1}\mathbb{1}\mathbf{b}'\Sigma^{-1} - \mathbb{1}'\Sigma^{-1}\mathbf{b}\mathbb{1}'\Sigma^{-1}\big)(\mathbf{X}^{\mathscr{R}} - X_{1:m:n}\mathbb{1}) \\ &= \frac{1}{\Delta}\mathbb{1}'\Sigma^{-1}\big(\mathbb{1}\mathbf{b}' - \mathbf{b}\mathbb{1}'\big)\Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - X_{1:m:n}\mathbb{1}).\end{aligned}$$

11.1.2 Linear Equivariant Estimation

Balakrishnan et al. [139] have presented general results on the relation between best linear unbiased estimation and best linear equivariant estimation. Subsequently, representations of the best (affine) linear equivariant estimators (BLEEs) of the location and scale parameters are given in three different setups, i.e., we present the linear equivariant estimator that minimizes the respective standardized mean squared error in the progressive censoring context. For generalized Pareto distributions, BLEEs and BLUEs are given in Burkschat [231]. Notice that the risk of equivariant estimators (and predictors) is independent of the parameters. The term invariance, which is often used instead of equivariance in the literature (see, for example, Nelson [676] or Takada [833]), refers to this fact. In particular, the BLEEs coincide with the best linear (risk-)invariant estimators (BLIEs) obtained by Mann [635], since every linear risk-invariant estimator is also equivariant (see Bondesson [214]). The following results are taken from Balakrishnan et al. [139].

Location model $\mathscr{F}_l$

The BLEE and BLUE for the location parameter μ coincide, i.e.,

$$\widehat{\mu}_{\mathsf{LE}} = \widehat{\mu}_{\mathsf{LU}} = \left(\mathbb{1}'\Sigma^{-1}\mathbb{1}\right)^{-1}\left(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}} - \mathbb{1}'\Sigma^{-1}\mathbf{b}\right).$$

Scale model $\mathscr{F}_s$

The BLEE of ϑ is given by

$$\widehat{\vartheta}_{\mathsf{LE}} = \mathbf{b}'(\Sigma + \mathbf{b}\mathbf{b}')^{-1}\mathbf{X}^{\mathscr{R}} = (1 + \mathbf{b}'\Sigma^{-1}\mathbf{b})^{-1}\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}.$$

Its expectation and mean squared error are given by $E\widehat{\vartheta}_{\mathsf{LE}} = \frac{\mathbf{b}'\Sigma^{-1}\mathbf{b}}{1+\mathbf{b}'\Sigma^{-1}\mathbf{b}}\vartheta$ and $\mathsf{MSE}(\widehat{\vartheta}_{\mathsf{LE}}) = \frac{\vartheta^2}{1+\mathbf{b}'\Sigma^{-1}\mathbf{b}}$, respectively.

Location–scale model $\mathscr{F}_{ls}$

The BLEEs of the location and scale parameters are given by

$$\widehat{\mu}_{\mathsf{LE}} = \frac{1}{\Delta_1}\left[\left(\mathbf{b}'\Sigma^{-1}\mathbf{b} + 1\right)\left(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\right) - \left(\mathbb{1}'\Sigma^{-1}\mathbf{b}\right)\left(\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\right)\right],$$

$$\widehat{\vartheta}_{\mathsf{LE}} = \frac{1}{\Delta_1}\left[\left(\mathbb{1}'\Sigma^{-1}\mathbb{1}\right)\left(\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\right) - \left(\mathbb{1}'\Sigma^{-1}\mathbf{b}\right)\left(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\right)\right],$$

with

$$E(\widehat{\mu}_{\mathsf{LE}}) = \mu + \frac{\mathbb{1}'\Sigma^{-1}\mathbf{b}}{\Delta_1}\vartheta, \quad E(\widehat{\vartheta}_{\mathsf{LE}}) = \frac{\Delta}{\Delta_1}\vartheta,$$

$$\mathsf{MSE}(\widehat{\mu}_{\mathsf{LE}}) = \frac{\mathbf{b}'\Sigma^{-1}\mathbf{b} + 1}{\Delta_1}\vartheta^2, \quad \mathsf{MSE}(\widehat{\vartheta}_{\mathsf{LE}}) = \frac{\mathbb{1}'\Sigma^{-1}\mathbb{1}}{\Delta_1}\vartheta^2,$$

$$E((\widehat{\mu}_{\mathsf{LE}} - \mu)(\widehat{\vartheta}_{\mathsf{LE}} - \vartheta)) = -\frac{\mathbb{1}'\Sigma^{-1}\mathbf{b}}{\Delta_1}\vartheta^2,$$

where $\Delta_1 = \Delta + \mathbb{1}'\Sigma^{-1}\mathbb{1}$.

11.1.3 First-Order Approximations to BLUEs and BLEEs

In principle, the computation of the linear estimators given above is possible when the second-order moments exist. However, in many cases, the computation of the moments is difficult. Moreover, for large m, the computation of the inverse matrix Σ^{-1} may also cause some difficulties. In such cases, we can use the approximations to the moments given in Sect. 7.6. In order to illustrate this approach, let the cumulative distribution function F and its density function f be as in Sect. 7.6. Then, we have the approximation of the covariance matrix

$$\Sigma \approx \Delta \Sigma_{\mathbf{U}^{\mathscr{R}}} \Delta,$$

with a diagonal matrix $\Delta^{-1} = \mathsf{diag}(f(F^{\leftarrow}(\Pi_1)), \ldots, f(F^{\leftarrow}(\Pi_m)))$, $\Pi_j = EU_{j:m:n}$, $1 \leq j \leq m$, and $\Sigma_{\mathbf{U}^{\mathscr{R}}} = \mathsf{Cov}(\mathbf{U}^{\mathscr{R}})$. Then, the mean vector $\mathbf{b}$ is replaced by $(F^{\leftarrow}(\Pi_1), \ldots, F^{\leftarrow}(\Pi_m))'$ and the covariance matrix Σ is replaced by $\Delta \Sigma_{\mathbf{U}^{\mathscr{R}}} \Delta$. Notice that Δ is a diagonal matrix and that the inverse of $\Sigma_{\mathbf{U}^{\mathscr{R}}}$ is a tridiagonal matrix (see (7.15)). This method provides approximations to the BLUEs and BLEEs given above. The results have been worked out by Balakrishnan and Rao [115]. Expressions can be obtained from the results presented in Sect. 11.1.1.

This approach has been applied to various distributions whose variance–covariance matrix contains analytically intractable expressions. For instance, approximate BLUEs (ABLUES) for the extreme value distribution are presented in Thomas and Wilson [843] and Balakrishnan and Aggarwala [86]. Log-normal distributions have been discussed by Balasooriya and Balakrishnan [158].

11.2 Linear Estimation for Particular Distributions

11.2.1 Exponential Distributions

Suppose that the location–scale family (11.1c) is based on a standard exponential distribution $\mathsf{Exp}(1)$. In order to derive the representations of the BLUEs and BLEEs in the above setups, we need the inverse of the covariance matrix for progressively Type-II censored order statistics $Z_{1:m:n}, \ldots, Z_{m:m:n}$ from an $\mathsf{Exp}(1)$-population. From (7.7), we get

$$\Sigma^{-1} = \Sigma_{\mathbf{Z}^{\mathscr{R}}}^{-1} = \begin{pmatrix} \gamma_1^2+\gamma_2^2 & -\gamma_2^2 & 0 & \cdots & \cdots & 0 \\ -\gamma_2^2 & \gamma_2^2+\gamma_3^2 & -\gamma_3^2 & 0 & \cdots & 0 \\ 0 & -\gamma_3^2 & \gamma_3^2+\gamma_4^2 & -\gamma_4^2 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & -\gamma_{m-1}^2 & \gamma_{m-1}^2+\gamma_m^2 & -\gamma_m^2 \\ 0 & \cdots & \cdots & 0 & -\gamma_m^2 & \gamma_m^2 \end{pmatrix}.$$

Subsequently, we need the following identities, which can be easily checked. Notice that $\mathbf{b} = (b_1, \ldots, b_m)'$ with $b_r = EZ_{r:m:n} = \sum_{j=1}^{r} \frac{1}{\gamma_j}$, $1 \leq r \leq m$, and $\mathbf{e}_1 = (1, 0^{*m-1})'$ denotes the first vector of the canonical standard basis of $\mathbb{R}^m$. Then,

$$\begin{aligned} \Sigma^{-1}\mathbb{1} &= n^2\mathbf{e}_1 \\ \Sigma^{-1}\mathbf{b} &= (\gamma_1-\gamma_2, \gamma_2-\gamma_3, \ldots, \gamma_{m-1}-\gamma_m, \gamma_m)' = \mathscr{R} + \mathbb{1} \\ \mathbf{b}'\Sigma^{-1}\mathbf{b} &= m \\ \mathbf{b}'\Sigma^{-1}\mathbb{1} &= n. \end{aligned}$$

Linear Estimates

Location model $\mathscr{F}_l$

The BLUE (and the BLEE) of μ (with $\vartheta = 1$) is given by

$$\widehat{\mu}_{\mathsf{LU}} = \widehat{\mu}_{\mathsf{LE}} = Z_{1:m:n} - \frac{1}{n}. \tag{11.4}$$

For $\vartheta > 0$, we have $\widehat{\mu}_{\mathsf{LU}} = \widehat{\mu}_{\mathsf{LE}} = Z_{1:m:n} - \frac{\vartheta}{n}$.

Scale model $\mathscr{F}_s$

For $\mu = 0$, the BLUE of ϑ is given by

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{m}\sum_{j=1}^{m}(R_j + 1)Z_{j:m:n}.$$

For $\mu \in \mathbb{R}$, we have $\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{m}\sum_{j=1}^{m}(R_j + 1)(Z_{j:m:n} - \mu)$. The BLEE of ϑ is given by the expression

$$\widehat{\vartheta}_{\mathsf{LE}} = \frac{1}{m+1}\sum_{j=1}^{m}(R_j + 1)(Z_{j:m:n} - \mu). \tag{11.5}$$

Remark 11.2.1. The case of general progressive censoring is discussed in Balakrishnan and Sandhu [123] and Balakrishnan and Aggarwala [86, p. 86].

Location–scale model $\mathscr{F}_{ls}$

Let $m \geq 2$. First, we notice that $\Delta = n^2(m-1) > 0$. Therefore, the BLUEs $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$ are given by

$$\widehat{\mu}_{\mathsf{LU}} = \frac{1}{n^2(m-1)}\Big(mn^2 Z_{1:m:n} - n\sum_{j=1}^{m}(R_j + 1)Z_{j:m:n}\Big) \tag{11.6a}$$

$$= Z_{1:m:n} - \frac{1}{n(m-1)}\sum_{j=2}^{m}(R_j + 1)(Z_{j:m:n} - Z_{1:m:n}) = Z_{1:m:n} - \frac{\widehat{\vartheta}_{\mathsf{LU}}}{n},$$

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{m-1}\sum_{j=2}^{m}(R_j + 1)(Z_{j:m:n} - Z_{1:m:n}). \tag{11.6b}$$

The covariance matrix of $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$ is given by

$$\frac{\vartheta^2}{n^2(m-1)}\begin{pmatrix} m & -n \\ -n & n^2 \end{pmatrix}. \tag{11.7}$$

Noticing that $\Delta_1 = n^2 m$, the BLEEs $\widehat{\mu}_{\mathsf{LE}}$ and $\widehat{\vartheta}_{\mathsf{LE}}$ are given by

$$\widehat{\mu}_{\mathsf{LE}} = Z_{1:m:n} - \frac{1}{nm}\sum_{j=2}^{m}(R_j+1)(Z_{j:m:n} - Z_{1:m:n}) = Z_{1:m:n} - \frac{\widehat{\vartheta}_{\mathsf{LE}}}{n},$$

$$\widehat{\vartheta}_{\mathsf{LE}} = \frac{1}{m}\sum_{j=2}^{m}(R_j+1)(Z_{j:m:n} - Z_{1:m:n}).$$

Remark 11.2.2. Explicit expressions of BLUEs are also available for a general progressive censored sample $Z_{r+1:m:n}, \dots, Z_{m:m:n}$ (see Balakrishnan and Aggarwala [86, Sects. 6.1.2, 6.2.2]), i.e.,

$$\begin{aligned}\widehat{\vartheta}_{\mathsf{LU}} &= \frac{1}{m-r-1}\sum_{j=r+2}^{m}(R_j+1)(Z_{j:m:n} - Z_{r+1:m:n}),\\ \widehat{\mu}_{\mathsf{LU}} &= Z_{r+1:m:n} - \Big(\sum_{j=1}^{r+1}\frac{1}{n-j+1}\Big)\widehat{\vartheta}_{\mathsf{LU}}.\end{aligned} \tag{11.8}$$

The multi-sample case has been addressed in Balakrishnan et al. [130] and Burkschat et al. [236]. They considered s independent progressively Type-II censored samples $Z^{(i)}_{1:m_i:n_i}, \dots, Z^{(i)}_{m_i:m_i:n_i}$, $1 \le i \le s$, with possibly different censoring schemes and sample sizes. For brevity, we present their result only for s samples with identical censoring scheme $\mathscr{R}$. In this case, the BLUEs are given by

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{m-1}\sum_{j=2}^{m}(R_j+1)(\overline{Z}_{j:m:n} - \overline{Z}_{1:m:n}),$$

$$\widehat{\mu}_{\mathsf{LU}} = \frac{1}{(m-1)n}\sum_{j=1}^{m}(R_j+1)(m\overline{Z}_{1:m:n} - \overline{Z}_{j:m:n}),$$

where $\overline{Z}_{j:m:n} = \frac{1}{s}\sum_{i=1}^{s} Z^{(i)}_{j:m:n}$ denotes the average of the jth progressively Type-II censored order statistic in each sample, $1 \le j \le m$. Expressions for the BLEEs are derived in Burkschat [231]. Results for general progressive censoring may be obtained in a similar fashion. For m-generalized order statistics, the problem has been addressed by Ahsanullah [23].

Model	BLUE Location	Scale	BLEE Location	Scale
Scale ($\mu = 0$)	–	9.08625	–	8.07667
Location–scale	−0.32950	9.86857	−0.26447	8.63500

Table 11.1 BLUEs and BLEEs for Nelson's insulating fluid data under exponential assumption

Example 11.2.3. We apply the linear estimators presented above to the progressively censored version of Nelson's insulating data 1.1.5 which has been analyzed under the assumption of exponential distributions in Viveros and Balakrishnan [875]. The results are presented in Table 11.1. Given the same assumptions, the data has also been discussed for general progressive Type-II censoring when the minimum of the sample is censored in Balakrishnan and Aggarwala [86, p. 98].

11.2.2 Generalized Pareto Distributions

Expressions for BLUEs of the parameters of a location–scale family of particular generalized Pareto distributions have been derived by many authors. In the following, we state the BLUEs as presented in Burkschat et al. [236] (see also Burkschat et al. [235] and Burkschat [231]), their variances, and covariance. In order to avoid trivialities, it is supposed throughout that $n \geq 2$. The family of distributions is based on a generalized Pareto distribution as in Definition A.1.11 with cumulative distribution function F, $q \neq 0$, given by

$$F(t) = 1 - (\tau - pt)^{1/q}, \quad t \geq 0, \tau - pt > 0,$$

where $\tau \in \mathbb{R}$ and $p \cdot q > 0$. For $\tau = 1$, the distribution function F corresponds to the generalized Pareto distribution given in Pickands [720]. The choice $\tau = 0$ corresponds to the distribution class defined in Reiss [750, p. 42]. For $p, q > 0$, the support is given by the interval $[(\tau-1)p^{-1}, \tau p^{-1}]$ including the uniform distribution as a particular case. For $\tau = 1$ and $p = q > 0$, the exponential distribution results as a limit if $p \to 0$. In the case $p, q < 0$, the distribution is supported on the interval $[(\tau - 1)p^{-1}, \infty)$ which includes Pareto distributions of the first and second kind (cf. Johnson et al. [483, p. 574–575]).

In the location–scale case, we present a simplified version of the estimators than those presented in Burkschat et al. [236]. This is due to Burkschat [231]. Moreover, we consider only the one-sample case. For details on the multi-sample situation, we refer to Burkschat et al. [236] and Burkschat [231]. In order to simplify the notation, let

$$e_j = \prod_{k=1}^{j}(\gamma_k + 2q),\ d_j = \prod_{k=1}^{j}(\gamma_k + q),\ c_j = \prod_{k=1}^{j}\gamma_k, \quad 1 \leq j \leq m. \tag{11.9}$$

Then, we conclude from Theorems 7.2.2 and 7.2.5 that

$$EY_{j:m:n} = p^{-1}\Big(\tau - \frac{c_j}{d_j}\Big), \quad 1 \le j \le m.$$

Introducing the notation $\mathbf{a} = (a_1, \dots, a_m)'$ with $a_j = c_j/d_j$, $1 \le j \le m$, we get

$$E\mathbf{Y}^{\mathscr{R}} = \mathbf{b} = \frac{1}{p}(\tau\mathbb{1} - \mathbf{a}). \tag{11.10}$$

To present explicit representations of the BLUEs, we define the weights

$$\begin{aligned}
w_1 &= (\Sigma^{-1}\mathbb{1})_1 = \frac{p^2 e_1}{qc_1}(R_1 + 1 - q) + \frac{p^2 e_1 d_1}{q^2},\\
w_j &= (\Sigma^{-1}\mathbb{1})_j = \frac{p^2 e_j}{qc_j}(R_j + 1 - q), \quad 2 \le j \le m-1,\\
w_m &= (\Sigma^{-1}\mathbb{1})_m = \frac{p^2 e_m}{qc_m}(R_m + 1 + q),\\
v_1 &= (\Sigma^{-1}\mathbf{b})_1 = \frac{\tau}{p}w_1 - \frac{p}{q^2}d_1 e_1,\\
v_j &= (\Sigma^{-1}\mathbf{b})_j = \frac{\tau}{p}w_j, \quad 2 \le j \le m,
\end{aligned} \tag{11.11}$$

and

$$\begin{aligned}
\eta_{11} &= \mathbb{1}'\Sigma^{-1}\mathbb{1} = p^2\left[\frac{e_1^2}{q^2} + \Psi\right], \quad \eta_{12} = \mathbb{1}'\Sigma^{-1}\mathbf{b} = \frac{\tau}{p}\eta_{11} - \frac{p}{q^2}d_1 e_1,\\
\eta_{22} &= \mathbf{b}'\Sigma^{-1}\mathbf{b} = \left(\frac{\tau}{p}\right)^2\eta_{11} + \frac{n - 2\tau d_1}{q^2}e_1,
\end{aligned}$$

where $\Psi = \sum_{j=1}^m \frac{e_j}{c_j} = \sum_{j=1}^m \prod_{k=1}^j \Big(1 + \frac{2q}{\gamma_k}\Big)$. Finally, we define

$$\Delta = \eta_{11}\eta_{22} - \eta_{12}^2.$$

Theorem 11.2.4. Suppose $\gamma_m + 2q > 0$.

(i) If ϑ is known, then the BLUE of μ is given by

$$\widehat{\mu}_{\mathsf{LU}} = \frac{1}{\eta_{11}}\sum_{j=1}^m w_j\left[X_{j:m:n} - \frac{\vartheta}{p}\Big(\tau - \frac{c_j}{d_j}\Big)\right],$$

and its variance is $\mathsf{Var}(\widehat{\mu}_{\mathsf{LU}}) = \vartheta^2/\eta_{22}$.

(ii) If μ is known, then the BLUE of ϑ is given by

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{\eta_{22}} \sum_{j=1}^{m} v_j (X_{j:m:n} - \mu),$$

and its variance is $\mathsf{Var}(\widehat{\vartheta}_{\mathsf{LU}}) = \vartheta^2/\eta_{11}$.

(iii) Let $m \geq 2$. If μ and ϑ are to be estimated simultaneously, then the BLUEs are given by

$$\widehat{\mu}_{\mathsf{LU}} = X_{1:m:n} - \frac{\tau(n+q) - n}{p(n+q)} \widehat{\vartheta}_{\mathsf{LU}},$$

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{n+q}{p(n\Psi - n - 2q)} \sum_{j=2}^{m} w_j (X_{j:m:n} - X_{1:m:n}),$$

where w_j, $1 \leq j \leq m$, are given in (11.11). The covariance matrix of the BLUEs is given by

$$\mathsf{Cov}\begin{pmatrix} \widehat{\mu}_{\mathsf{LU}} \\ \widehat{\vartheta}_{\mathsf{LU}} \end{pmatrix} = \vartheta^2 \frac{1}{\Delta} \begin{pmatrix} \eta_{22} & -\eta_{12} \\ -\eta_{12} & \eta_{11} \end{pmatrix}. \tag{11.12}$$

Proof. In the location and scale setup, the results follow directly from the general representations of the BLUEs and (11.11).

In the location–scale situation, the derivation of the above expressions requires more effort. First, we use (11.10) to write the determinant Δ in terms of $\mathbf{a}$:

$$\Delta = \frac{1}{p^2}\Big(\mathbb{1}'\Sigma^{-1}\mathbb{1}\mathbf{a}'\Sigma^{-1}\mathbf{a} - (\mathbb{1}'\Sigma^{-1}\mathbf{a})^2\Big) = \frac{\Delta_{\mathbf{a}}}{p^2}, \quad \text{say}.$$

According to Burkschat et al. [236],

$$\mathbf{a}'\Sigma^{-1}\mathbf{z} = \frac{p^2 e_1 d_1}{q^2} z_1 = z_1 \mathbf{a}'\Sigma^{-1}\mathbb{1}, \quad \mathbf{z} = (z_1, \ldots, z_m)' \in \mathbb{R}^m, \tag{11.13}$$

where $\mathbf{a}$ is given as in (11.10). Furthermore, for $\mathbf{b} = E\mathbf{Y}^{\mathscr{R}}$, we get

$$\mathbf{b}\mathbb{1}' - \mathbb{1}\mathbf{b}' = -\frac{1}{p}\Big(\mathbf{a}\mathbb{1}' - \mathbb{1}\mathbf{a}'\Big).$$

Combining these results, we get from (11.3) and (11.13), after some rearrangements, for the BLUE of ϑ

$$\begin{aligned}
\widehat{\vartheta}_{\mathsf{LU}} &= \frac{1}{\Delta}\big((\mathbb{1}'\Sigma^{-1}\mathbb{1})(\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}) - (\mathbb{1}'\Sigma^{-1}\mathbf{b})(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}})\big)\\
&= -\frac{1}{p\Delta}\mathbb{1}'\Sigma^{-1}(\mathbb{1}\mathbf{a}' - \mathbf{a}\mathbb{1}')\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\\
&= p\frac{\mathbf{a}'\Sigma^{-1}\mathbb{1}}{\Delta_{\mathbf{a}}}\mathbb{1}'\Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - X_{1:m:n}\mathbb{1}). \qquad (11.14)
\end{aligned}$$

Noticing that

$$\begin{aligned}
\Delta_{\mathbf{a}} &= \mathbb{1}'\Sigma^{-1}\mathbf{a}\Big(a_1\mathbb{1}'\Sigma^{-1}\mathbb{1} - \mathbb{1}'\Sigma^{-1}\mathbf{a}\Big)\\
&= \mathbb{1}'\Sigma^{-1}\mathbf{a}\Big(a_1 p^2\Big(\Psi + \frac{e_1^2}{q^2}\Big) - p^2\frac{e_1 d_1}{q^2}\Big)\\
&= \mathbb{1}'\Sigma^{-1}\mathbf{a}\frac{p^2}{d_1}\Big(c_1\Psi + \frac{c_1 e_1^2}{q^2} - \frac{e_1 d_1^2}{q^2}\Big)\\
&= \mathbb{1}'\Sigma^{-1}\mathbf{a}\frac{p^2}{d_1}\Big(c_1\Psi - e_1\Big),
\end{aligned}$$

we get

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{d_1}{p(n\Psi - e_1)}\sum_{j=2}^{m} w_j(X_{j:m:n} - X_{1:m:n}).$$

Recalling that $d_1 = n+q$ and $e_1 = n+2q$ (see (11.9)), this yields the representation of $\widehat{\vartheta}_{\mathsf{LU}}$.

The BLUE of μ has the representation

$$\begin{aligned}
\widehat{\mu}_{\mathsf{LU}} &= \frac{1}{\Delta}\big((\mathbf{b}'\Sigma^{-1}\mathbf{b})(\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}) - (\mathbb{1}'\Sigma^{-1}\mathbf{b})(\mathbf{b}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}})\big)\\
&= \frac{1}{\Delta}\mathbf{b}'\Sigma^{-1}\big(\mathbf{b}\mathbb{1}' - \mathbb{1}\mathbf{b}'\big)\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\\
&= \frac{1}{p^2\Delta}(\tau\mathbb{1} - \mathbf{a})'\Sigma^{-1}\big(\mathbb{1}\mathbf{a}' - \mathbf{a}\mathbb{1}'\big)\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\\
&= -\frac{\tau}{p}\Big(-\frac{1}{p\Delta}\mathbb{1}'\Sigma^{-1}\big(\mathbb{1}\mathbf{a}' - \mathbf{a}\mathbb{1}'\big)\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\Big)\\
&\qquad - \frac{1}{p^2\Delta}\mathbf{a}'\Sigma^{-1}\big(\mathbb{1}\mathbf{a}' - \mathbf{a}\mathbb{1}'\big)\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\\
&\overset{(11.14)}{=} -\frac{\tau}{p}\widehat{\vartheta}_{\mathsf{LU}} - \frac{1}{p^2\Delta}\mathbf{a}'\Sigma^{-1}\big(\mathbb{1}\mathbf{a}' - \mathbf{a}\mathbb{1}'\big)\Sigma^{-1}\mathbf{X}^{\mathscr{R}}. \qquad (11.15)
\end{aligned}$$

The last term can be rewritten using the following two representations:

$$\begin{aligned}
&\mathbf{a}'\Sigma^{-1}\big(\mathbb{1}\mathbf{a}'-\mathbf{a}\mathbb{1}'\big)\Sigma^{-1}(\mathbf{X}^{\mathscr{R}}-X_{1:m:n}\mathbb{1})\\
&\quad=\mathbf{a}'\Sigma^{-1}\mathbb{1}\mathbf{a}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}-\mathbf{a}'\Sigma^{-1}\mathbf{a}\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\\
&\qquad\qquad-\mathbf{a}'\Sigma^{-1}\mathbb{1}\mathbf{a}'\Sigma^{-1}X_{1:m:n}\mathbb{1}+\mathbf{a}'\Sigma^{-1}\mathbf{a}\mathbb{1}'\Sigma^{-1}X_{1:m:n}\mathbb{1}\\
&\quad=\mathbf{a}'\Sigma^{-1}\mathbf{a}\mathbb{1}'\Sigma^{-1}X_{1:m:n}\mathbb{1}-\mathbf{a}'\Sigma^{-1}\mathbf{a}\mathbb{1}'\Sigma^{-1}\mathbf{X}^{\mathscr{R}}\\
&\quad=-a_1\mathbf{a}'\Sigma^{-1}\mathbb{1}\mathbb{1}'\Sigma^{-1}(\mathbf{X}^{\mathscr{R}}-X_{1:m:n}\mathbb{1})\\
&\quad\overset{(11.14)}{=}-a_1 p\Delta\cdot\widehat{\vartheta}_{\mathsf{LU}},\\
&\mathbf{a}'\Sigma^{-1}\big(\mathbb{1}\mathbf{a}'-\mathbf{a}\mathbb{1}'\big)\Sigma^{-1}X_{1:m:n}\mathbb{1}=-\Delta_{\mathbf{a}}\cdot X_{1:m:n}.
\end{aligned}$$

Hence, (11.15) has the representation

$$\widehat{\mu}_{\mathsf{LU}}=X_{1:m:n}-\frac{\tau-a_1}{p}\widehat{\vartheta}_{\mathsf{LU}}.$$

This proves the desired representation of $\widehat{\mu}_{\mathsf{LU}}$ noticing that $\gamma_1=n$ and that $a_1=c_1/d_1=n/(n+q)$ (see (11.9)). □

The multi-sample case has been considered in Burkschat et al. [236] and Burkschat [231]. Expressions for general progressive censoring are presented in Balakrishnan and Aggarwala [86] for uniform and Pareto distributions.

For illustration, we present BLUEs and BLEEs for some important particular distributions.

Uniform Distribution

An important special case of generalized Pareto distributions is the uniform distribution which corresponds to $\tau=p=q=1$. In this case, we get the following simplified expressions. Notice that the condition $\gamma_m+2>0$ is always satisfied so that it can be dropped.

Corollary 11.2.5. Let $\Psi=\sum_{j=1}^m\frac{e_j}{c_j}=\sum_{j=1}^n\prod_{k=1}^j\Big(1+\frac{2}{\gamma_k}\Big)$, where $e_j=\prod_{k=1}^j(\gamma_k+2)$, $c_j=\prod_{k=1}^j\gamma_k$, $1\le j\le m$.

(i) If ϑ is known, then the BLUE of μ is given by

$$\widehat{\mu}_{\mathsf{LU}}=\frac{1}{\Psi+(n+2)^2}\sum_{j=1}^m w_j\Big[X_{j:m:n}-\vartheta\Big(1-\frac{c_j}{d_j}\Big)\Big],$$

where $w_1=\frac{n+2}{n}R_1+(n+2)(n+1)$, $w_j=\frac{e_j}{c_j}R_j$, $2\le j\le m-1$, $w_m=\frac{e_m}{c_m}(R_m+2)$. Its variance is $\mathsf{Var}(\widehat{\mu}_{\mathsf{LU}})=\vartheta^2\big(\Psi+(n+2)^2\big)^{-1}$.

(ii) If μ is known, then the BLUE of ϑ is given by

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{\Psi}\sum_{j=1}^{m} w_j (X_{j:m:n} - \mu),$$

where $w_j = \frac{e_j}{c_j} R_j$, $1 \le j \le m-1$, $w_m = \frac{e_m}{c_m}(R_m + 2)$. Its variance is $\mathsf{Var}(\widehat{\vartheta}_{\mathsf{LU}}) = \vartheta^2 \Psi^{-1}$.

(iii) If μ and ϑ are to be estimated simultaneously, then the BLUEs are given by

$$\widehat{\mu}_{\mathsf{LU}} = X_{1:m:n} - \frac{1}{n+1}\widehat{\vartheta}_{\mathsf{LU}}$$

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{n+1}{\Psi n - n - 2}\sum_{j=2}^{m} w_j (X_{j:m:n} - X_{1:m:n}),$$

which have the covariance matrix

$$\mathsf{Cov}\begin{pmatrix}\widehat{\mu}_{\mathsf{LU}}\\ \widehat{\vartheta}_{\mathsf{LU}}\end{pmatrix} = \frac{\vartheta^2}{(n+2)(\Psi n - n - 2)}\begin{pmatrix}\Psi & -(\Psi + n + 2)\\ -(\Psi + n + 2) & \Psi + (n+2)^2\end{pmatrix}.$$

Related problems for order statistics are discussed in Sarhan [773] and Sarhan and Greenberg [775] (see also Arnold et al. [58, p. 175, 199]).

Pareto Distribution

For the Pareto(α)-distribution, we have $\tau = 0$, $p = -1$, and $q = -1/\alpha > 0$. We present only the location–scale result.

Corollary 11.2.6. Suppose $\gamma_m - 2/\alpha > 0$ and that $m \ge 2$. Then, the BLUEs of μ and ϑ are given by

$$\widehat{\mu}_{\mathsf{LU}} = X_{1:m:n} + \frac{n\alpha}{n\alpha - 1}\widehat{\vartheta}_{\mathsf{LU}},$$

$$\widehat{\vartheta}_{\mathsf{LU}} = -\frac{(n\alpha - 2)(n\alpha - 1)}{\Delta}\sum_{j=2}^{m} w_j (X_{j:m:n} - X_{1:m:n}),$$

where $w_1 = \frac{n\alpha - 1}{n}(R_1 + 1 - 1/\alpha) + (n\alpha - 2)(n\alpha - 1)$, $w_j = \frac{e_j}{c_j}((R_j + 1)\alpha - 1)$, $2 \le j \le m-1$, $w_m = \frac{e_m}{c_m}((R_m + 1)\alpha + 1)$.

The covariance matrix is

$$\mathsf{Cov}\begin{pmatrix}\widehat{\mu}_{\mathsf{LU}}\\ \widehat{\vartheta}_{\mathsf{LU}}\end{pmatrix} = \frac{\vartheta^2}{\Psi n\alpha + n\alpha - 2}\begin{pmatrix}n\alpha & -(n\alpha - 1)\\ -(n\alpha - 1) & \frac{\Psi}{n\alpha - 2} + n\alpha - 2\end{pmatrix}. \qquad (11.16)$$

Example 11.2.7. Balakrishnan and Aggarwala [86] simulated the progressively Type-II censored data

$$5.11073, 5.34932, 5.36434, 5.70137, 5.90067$$

from a Pareto(3)-distribution with location parameter $\mu = 0$, scale parameter $\vartheta = 5$, and censoring plan $\mathscr{R} = (5, 0, 2, 0, 3)$. So, $n = 15$ and $m = 5$ observations result. The linear estimates for the parameters μ and ϑ are computed from the BLUEs (note that these estimators exist since $\gamma_5 = 4 > 2/3$) as $\widehat{\mu}_{\mathsf{LU}} = 1.87675$ and $\widehat{\vartheta}_{\mathsf{LU}} = 3.16211$. Their standard errors are given by

$$\mathsf{SE}(\widehat{\mu}_{\mathsf{LU}}) = 1.83039, \mathsf{SE}(\widehat{\vartheta}_{\mathsf{LU}}) = 1.79116, \mathsf{Cov}(\widehat{\mu}_{\mathsf{LU}}, \widehat{\vartheta}_{\mathsf{LU}}) = -3.27588.$$

These estimates are based on the assumption of a known shape parameter. Since the estimators can be seen as a function of the shape parameter α, Balakrishnan and Aggarwala [86] have conducted a sensitivity analysis w.r.t. the shape parameter α. But, due to the construction of the estimators, explicit representations as a function of the shape parameter are available. In particular, we get for the present data

$$\widehat{\vartheta}_{\mathsf{LU}}(\alpha) = \frac{7.20(15\alpha - 1)(1.76\alpha^4 - 1.45\alpha^3 + 0.69\alpha^2 - 0.16\alpha + 0.01)}{\alpha(160\alpha^3 - 82\alpha^2 + 21\alpha - 2)}.$$

Expanding this expression in a Taylor series of order two around $\alpha = 3$, we obtain the approximation

$$\widetilde{\vartheta}(\alpha) = 3.16211 + 1.17089(\alpha - 3) + 0.00557(\alpha - 3)^2.$$

Similarly, we get

$$\widetilde{\mu}(\alpha) = 1.87675 - 1.17300(\alpha - 3) - 0.00498(\alpha - 3)^2.$$

These expressions provide accurate approximations of the values given in Balakrishnan and Aggarwala [86] for $\alpha \in \{2.6, 2.8, 3.0, 3.2, 3.4\}$. The Taylor expansions tell us that the estimators are almost a linear function of the shape parameter (in a neighborhood of the true value). Further, it can be shown that, for large α, the estimators are almost linear in the shape parameter. In particular,

$$\widehat{\vartheta}_{\mathsf{LU}}(\alpha) \approx 1.18746\alpha - 0.44933, \quad \widehat{\mu}_{\mathsf{LU}}(\alpha) \approx 5.48090 - 1.18746\alpha.$$

0	0.70	1.22	1.72	3.12	3.21	4.13	4.94	5.54	6.61

Table 11.2 Progressively Type-II censored data given in Raqab et al. [746]. The applied censoring scheme was $\mathscr{R} = (3, 0, 1, 2, 0^{*2}, 3, 0, 1, 3)$

Lomax Distribution

In order to obtain the BLUEs in the location–scale model based on a Lomax distribution, we have to choose the parameters $\tau = 1$, $p = -1$, and $q = -1/\alpha$ with $\alpha > 0$. Notice that the different parametrization is important in order to use the correct moments in the weights of the linear estimates.

Corollary 11.2.8. Suppose $\gamma_m - 2/\alpha > 0$ and $m \geq 2$. Then, the BLUEs of μ and ϑ are given by

$$\widehat{\mu}_{\mathsf{LU}} = X_{1:m:n} - \frac{1}{n\alpha - 1}\widehat{\vartheta}_{\mathsf{LU}},$$

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{n\alpha - 1}{n\alpha(1-\Psi) - 2}\sum_{j=2}^{m} w_j (X_{j:m:n} - X_{1:m:n}),$$

where w_j are as given in Corollary 11.2.6. The covariance matrix equals that for the Pareto distribution given in (11.16).

For Type-II right censored samples, this problem has been solved by Vännman [867] (see also Kulldorff and Vännman [556]).

Example 11.2.9. Raqab et al. [746] generated a progressively Type-II censored data from rainfall data recorded at Los Angeles Civic Center (see Table 11.2) in order to predict progressively censored failure times. The resulting predicted values are summarized in Table 16.5. Following Madi and Raqab [626], Raqab et al. [746] assumed a Lomax distribution with shape parameter $\alpha = 1/2$. Notice that this choice is critical since the condition $\gamma_{10} - 2/\alpha > 0$ is violated (here $\gamma_{10} = 4$; see Corollary 11.2.8). The resulting estimates are given by $\widehat{\mu}_{\mathsf{LU}} = -0.30484$ and $\widehat{\vartheta}_{\mathsf{LU}} = 3.20077$. These values differ from those presented in Raqab et al. [746]. This may be due to a wrong use of formulas for the BLUEs presented in Balakrishnan and Aggarwala [86, p. 106]. These results have been established for a different parametrization of Pareto distributions.

11.2.3 Weibull and Extreme Value Distributions

Linear estimation of the model parameters has been studied for Weibull distributions in Mann [636] using a log-transformation of the failure times. Hence, a location–scale model of extreme value distributions as in (11.1c) results. The derivation of linear estimates requires calculation of single and product moments. For the single moments, we refer to (7.16). However, the product moments have to be computed numerically as illustrated in Mann [636] using the results of Lieblein [593]. The computation of moments for progressively Type-II censored order statistics from Weibull distribution has been addressed in Cramer and Kamps [298] for sequential order statistics (see Sect. 7.2.2). They used results of Lieblein [594] to obtain expressions for the product moments which are utilized to derive BLUEs for the distribution parameters. Using the mixture representation in terms of order statistics (see Sect. 10.1), Thomas and Wilson [843] presented also expressions for the BLUEs.

Since the explicit derivations are feasible in these scenarios, approximate BLUEs have been proposed as given in Sect. 11.1.3. Balakrishnan and Aggarwala [86] applied the method to the log-failure times of Nelson's insulating fluid data as given in Table 17.5. Assuming an extreme value distribution, they found the BLUEs $\widehat{\mu}_{\mathsf{LU}} = 2.456$ and $\widehat{\vartheta}_{\mathsf{LU}} = 1.31377$ which compare well to the MLEs as given in Viveros and Balakrishnan [875] ($\widehat{\mu}_{\mathsf{MLE}} = 2.222$ and $\widehat{\vartheta}_{\mathsf{MLE}} = 1.026$).

11.2.4 Laplace Distribution

Best linear estimation for the $\mathsf{Laplace}(\mu, \vartheta)$-distribution has been discussed in, e.g., Sarhan [772,773], Govindarajulu [411], and Balakrishnan et al. [128] (see also Kotz et al. [546, Sect. 2.6.1.5]). As pointed out in Balakrishnan and Aggarwala [86, p. 106], the results presented in Sect. 7.4 for the single and product moments of progressively Type-II censored order statistics from arbitrary continuous symmetric distributions may be utilized to compute the BLUEs for the parameters of scale and location–scale shifted symmetric distributions provided the single and product moments of progressively Type-II right censored and progressively Type-II left withdrawn order statistics from the corresponding folded distribution are known. This applies to the $\mathsf{Laplace}(\mu, \vartheta)$ distribution with cumulative distribution function

$$F(t) = \begin{cases} \frac{1}{2}\mathsf{e}^{(t-\mu)/\vartheta}, & t \leq \mu \\ 1 - \frac{1}{2}\mathsf{e}^{-(t-\mu)/\vartheta}, & t > \mu \end{cases}, \tag{11.17}$$

since the folded distribution is an exponential one. Notice that this distribution has been used to model certain real life-test data (see Bain and Engelhardt [74]) and has many applications in engineering and life sciences (see Kotz et al. [546]).

Explicit expressions for the BLUEs of μ and ϑ or the inverse of the variance–covariance matrix are not available. However, the general formula (11.2) may be used to obtain the required BLUEs as well as their variances and covariance. The moments of Laplace progressively Type-II censored order statistics can be computed easily through the results of Sects. 7.3 and 7.4. Using the results for moments of exponential progressively Type-II censored order statistics given in Theorem 7.2.1

as well as the results presented in Sect. 7.4, single and product moments of Laplace progressively Type-II censored order statistics can be calculated for a given censoring scheme $\mathscr{R}$. The following example is along the lines of Example 6.5 of Balakrishnan and Aggarwala [86].

Example 11.2.10. In the situation of Example 7.4.2, i.e., for the censoring plan $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$ with $m = 10$, and $n = 20$, a progressively Type-II right censored sample from the Laplace$(25, 5)$-distribution in (11.17) has been simulated using Algorithm 8.1.2. The procedure yields the following random numbers:

$$\begin{array}{ccccc} 9.61484196 & 13.59012714 & 21.09368461 & 23.48334002 & 24.00495686 \\ 24.56994841 & 24.74930175 & 24.89373289 & 26.37125510 & 26.53605049. \end{array}$$

Using the results presented in Sect. 7.4, we are now able to compute the BLUEs for μ and ϑ. For instance, the coefficients $a_{i:10:10}$, $1 \le i \le 10$, for the BLUE of μ, to three decimal places, are given by

−0.008	−0.010	−0.010	0.032
0.003	0.038	0.116	0.385
0.235	0.217		

and the coefficients $b_{i:10:20}$, $1 \le i \le 10$, for the BLUE of ϑ are given by

−0.105	−0.099	−0.098	−0.102
−0.098	−0.098	−0.083	0.106
0.073	0.505		

Furthermore, the variance of $\widehat{\mu}_{\mathsf{LU}}$ is $0.084\vartheta^2$, the variance of $\widehat{\vartheta}_{\mathsf{LU}}$ is $0.098\vartheta^2$, and the covariance of the BLUEs is $0.011\vartheta^2$.

Thus, based on the progressively Type-II right censored sample given above, the BLUEs of μ and ϑ and their standard errors are computed as

$$\widehat{\mu}_{\mathsf{LU}} = 25.786, \qquad \widehat{\mathsf{SE}}(\widehat{\mu}_{\mathsf{LU}}) = 1.250,$$

$$\widehat{\vartheta}_{\mathsf{LU}} = 4.315, \qquad \widehat{\mathsf{SE}}(\widehat{\vartheta}_{\mathsf{LU}}) = 1.350.$$

Remark 11.2.11. Obviously, the variance–covariance matrix of the BLUEs can be computed for any censoring scheme with $m = 10$ and $n = 20$ from a Laplace distribution. For Type-II right censoring $\mathscr{O}_{10} = (0^{*9}, 10)$ and for some selected censoring plans, the variances of the BLUEs and the corresponding covariances are given in Table 11.3. These variances are only slightly more favorable than the variances given above for the censoring pattern $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$. However, for this censoring scheme $\mathscr{R}$, items censored early on may be of use to the experimenter. The question of an optimal censoring pattern in terms of the variances of BLUEs will be addressed further in Chap. 26.

$\mathscr{R}$	$\mathsf{Var}(\widehat{\mu}_{\mathsf{LU}})/\vartheta^2$	$\mathsf{Var}(\widehat{\vartheta}_{\mathsf{LU}})/\vartheta^2$	$\mathsf{Cov}(\widehat{\mu}_{\mathsf{LU}}, \widehat{\vartheta}_{\mathsf{LU}})/\vartheta^2$
$(2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$	0.084	0.098	0.011
$\mathscr{O}_{10}$	0.070	0.109	0.013
$\mathscr{O}_1$	0.140	0.086	0.012
$(5, 0^{*8}, 5)$	0.088	0.095	0.006
$(2, 0^{*8}, 8)$	0.074	0.103	0.008
$(0^{*4}, 5^{*2}, 0^{*4})$	0.117	0.098	0.032

Table 11.3 Variances and covariances of the BLUEs for location and scale of a Laplace distribution with progressively Type-II censored order statistics from censoring scheme $\mathscr{R}$

11.2.5 Logistic Distributions

Best linear unbiased estimation for logistic distributions has been addressed by Balakrishnan and Kannan [104] and, in more detail, by Balakrishnan et al. [151]. They discussed a location–scale family $\mathscr{F}_{ls}$ from a standard logistic distribution with cumulative distribution function $F(t) = (1 + \mathrm{e}^{-t})^{-1}$, $t \in \mathbb{R}$. Two-parameter half-logistic distribution is considered in Balakrishnan and Saleh [119], whereas BLUEs for log-logistic distributions are presented in Balakrishnan and Saleh [120] (scale and location–scale families). Furthermore, BLUEs for generalized half-logistic distributions are computed in Balakrishnan and Saleh [121].

11.3 Asymptotic Best Linear Unbiased Estimators for Blocked Progressively Type-II Censored Order Statistics

Hofmann et al. [444] applied the asymptotic result in Theorem 3.4.11 to construct asymptotic BLUEs in a location–scale model as given in (11.1). Proceeding as in Sect. 11.1.1, the asymptotic best linear unbiased estimator (ABLUE) is given by

$$\begin{pmatrix} \widehat{\mu} \\ \widehat{\vartheta} \end{pmatrix} = (B'W^{-1}B)^{-1}B'W^{-1}\mathbf{Y}_{\bullet}^{\mathscr{R}},$$

with covariance matrix $\frac{\vartheta^2}{n}(B'W^{-1}B)^{-1}$, $B = (\mathbb{1}, \mathbf{u})$. The components of $B'W^{-1}B$ are given by

$$K_1 = \mathbb{1}'W^{-1}\mathbb{1} = \sum_{j=1}^{m} v_j^{-1}(\Delta_j^{-1} - \Delta_{j-1}^{-1})^2,$$

$$K_2 = \mathbf{u}'W^{-1}\mathbf{u} = \sum_{j=1}^{m} v_j^{-1}(u_j\Delta_j^{-1} - u_{j-1}\Delta_{j-1}^{-1})^2, \tag{11.18}$$

$$K_3 = \mathbb{1}' W^{-1}\mathbf{u} = \sum_{j=1}^{m} v_j^{-1}(\Delta_j^{-1} - \Delta_{j-1}^{-1})(u_j \Delta_j^{-1} - u_{j-1}\Delta_{j-1}^{-1}).$$

The generalized variance of the ABLUEs is given by

$$\det \mathsf{Cov}\begin{pmatrix} \widehat{\mu} \\ \widehat{\vartheta} \end{pmatrix} = \frac{\vartheta^2}{n}\left(K_1 K_2 - K_3^2\right)^{-1}. \tag{11.19}$$

For order statistics, similar quantities K_j arise which were first presented in Ogawa [694] (see also Ali and Umbach [33, p. 188], David and Nagaraja [327, Sect. 10.4], and Balakrishnan and Cohen [92]).

Chapter 12
Maximum Likelihood Estimation in Progressive Type-II Censoring

In this section, we assume that the sample $X_{1:m:n}, \dots, X_{m:m:n}$ of progressively Type-II censored order statistics is available from a population with an absolutely continuous cumulative distribution function $F_{\boldsymbol{\theta}}$ and density function $f_{\boldsymbol{\theta}}$ with $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)' \in \Theta \subseteq \mathbb{R}^p$. Moreover, we assume that the progressively censored sample is complete. Then, from (2.4), the likelihood function is given by

$$\mathsf{L}(\boldsymbol{\theta}; \mathbf{x}_m) = \prod_{j=1}^{m} \left[\gamma_j f_{\boldsymbol{\theta}}(x_j)(1 - F_{\boldsymbol{\theta}}(x_j))^{R_j}\right] \prod_{j=2}^{m} \mathbb{1}_{[x_{j-1},\infty)}(x_j). \tag{12.1}$$

Given that $x_1 < \cdots < x_m$, the log-likelihood function is

$$\begin{aligned}\ell(\boldsymbol{\theta}; \mathbf{x}_m) &= \log \mathsf{L}(\boldsymbol{\theta}; \mathbf{x}_m) \\ &= \sum_{j=1}^{m} \log \gamma_j + \sum_{j=1}^{m} \log f_{\boldsymbol{\theta}}(x_j) + \sum_{j=1}^{m} R_j \log(1 - F_{\boldsymbol{\theta}}(x_j)). \end{aligned} \tag{12.2}$$

Maximum likelihood estimates may also be computed when some of the progressively Type-II censored order statistics are not observed. In these situations, the estimates have to be obtained numerically in most cases. However, for general progressive censoring, some MLEs are explicitly available. In this case, the likelihood function is given by (see Balakrishnan and Aggarwala [86, pp. 118])

$$\begin{aligned}&\mathsf{L}(\boldsymbol{\theta}; x_{r+1}, \dots, x_m) \\ &= \binom{n}{r} (F_{\boldsymbol{\theta}}(x_{r+1}))^r \prod_{j=r+1}^{m} \left[\gamma_j f_{\boldsymbol{\theta}}(x_j)(1 - F_{\boldsymbol{\theta}}(x_j))^{R_j}\right] \prod_{j=r+2}^{m} \mathbb{1}_{[x_{j-1},\infty)}(x_j).\end{aligned}$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_12,

12.1 Exponential Distribution

According to (12.1), the likelihood function for a progressively censored sample $Z_{1:m:n}, \dots, Z_{m:m:n}$ from an $\mathsf{Exp}(\mu, \vartheta)$-population is given by

$$\mathsf{L}(\mu, \vartheta; \mathbf{z}_m) = \prod_{j=1}^{m} \Big[\frac{\gamma_j}{\vartheta} \mathrm{e}^{-(R_j+1)\frac{z_j - \mu}{\vartheta}} \Big] \prod_{j=2}^{m} \mathbb{1}_{[z_{j-1}, \infty)}(z_j) \mathbb{1}_{[\mu,\infty)}(z_1).$$

Assuming $z_1 < \cdots < z_m$, this expression simplifies to

$$\begin{aligned} \mathsf{L}(\mu, \vartheta; \mathbf{z}_m) &= \frac{c_{m-1}}{\vartheta^m} \exp\Big\{ -\frac{1}{\vartheta} \sum_{j=2}^{m} (R_j + 1)(z_j - z_1) - n \frac{z_1 - \mu}{\vartheta} \Big\} \mathbb{1}_{[\mu,\infty)}(z_1) \\ &= \frac{c_{m-1}}{\vartheta^m} \exp\Big\{ -\frac{m}{\vartheta} \widetilde{\vartheta} \Big\} \cdot \exp\Big\{ -n \frac{z_1 - \mu}{\vartheta} \Big\} \mathbb{1}_{[\mu,\infty)}(z_1), \end{aligned} \tag{12.3}$$

where $\widetilde{\vartheta} = \frac{1}{m} \sum_{j=2}^{m} (R_j + 1)(z_j - z_1)$.

Scale Parameter Unknown

Now, we consider the models given in (11.1). First, let us assume that μ is known. Then, given the assumption $\mu < z_1$, we get with $\widetilde{\mu} = z_1$ the log-likelihood

$$\ell(\vartheta; \mathbf{z}_m) = \log(c_{m-1}) - m \log(\vartheta) - \frac{1}{\vartheta} \Big(m\widetilde{\vartheta} + n(\widetilde{\mu} - \mu) \Big).$$

Obviously, the log-likelihood function $\ell(\cdot; \mathbf{z}_m)$ is strictly concave in $1/\vartheta$ (for any $\mu < z_1$) with $\lim_{\vartheta \to 0} \ell(\vartheta; \mathbf{z}_m) = \lim_{\vartheta \to \infty} \ell(\vartheta; \mathbf{z}_m) = -\infty$ so that the global maximum is an inner point of $(0, \infty)$. Differentiating w.r.t. ϑ results in the likelihood equation

$$-\frac{m}{\vartheta} + \frac{1}{\vartheta^2} \Big(m\widetilde{\vartheta} + n(\widetilde{\mu} - \mu) \Big) = 0$$

which has the unique solution $\vartheta^* = \frac{1}{m} \sum_{j=1}^{m} (R_j + 1)(z_j - \mu)$. This proves that

$$\widehat{\vartheta}^*_{\mathsf{MLE}} = \frac{1}{m} \sum_{j=1}^{m} (R_j + 1)(Z_{j:m:n} - \mu) \tag{12.4}$$

is the maximum likelihood estimator for ϑ if μ is assumed known. Properties of this estimator are presented in Theorem 12.1.1. Since the proof is similar to that of Theorem 12.1.4, it is omitted.

Theorem 12.1.1. The MLE $\widehat{\vartheta}^*_{\mathsf{MLE}}$ is a complete sufficient statistic. It has a $\Gamma(\vartheta/m, m)$-distribution so that its expectation and variance are given by

$$E\widehat{\vartheta}^*_{\mathsf{MLE}} = \vartheta, \quad \mathsf{Var}(\widehat{\vartheta}^*_{\mathsf{MLE}}) = \frac{\vartheta^2}{m}.$$

Sometimes, it is more suitable to consider the distribution of a scaled MLE, i.e.,

$$2m\frac{\widehat{\vartheta}^*_{\mathsf{MLE}}}{\vartheta} \sim \chi^2(2m).$$

Further, it can be shown that the MLE $\widehat{\vartheta}^*_{\mathsf{MLE}}$ attains the Cramér–Rao lower bound (see Cramer and Kamps [299]). Since $\widehat{\vartheta}^*_{\mathsf{MLE}}$ is unbiased and linear, it coincides with both the uniformly minimum variance unbiased estimator (UMVUE) and the BLUE of ϑ (see Sect. 11.2.1).

Remark 12.1.2. For general progressive censoring, an explicit expression of the maximum likelihood estimator is not available unless $r = 1$. It is the solution of the equation

$$(m-r)\vartheta + \frac{r x_{r+1}}{\mathrm{e}^{x_{r+1}/\vartheta} - 1} = w \tag{12.5}$$

with $w = \sum_{j=r+1}^{m}(R_j + 1)x_j$, where $x_{r+1}, \dots, x_m$ is the observed sample. It has been shown by Balakrishnan et al. [130] that the solution of the likelihood equation is unique (even in the multi-sample case). Hence, the solution can be computed easily by a Newton–Raphson procedure. Fernández [364] has established simple bounds on the MLE of ϑ in the scale model with general progressive censoring using Cardano's formula. Using the notation $\rho = 6w - (m-3r)x_{r+1}$, $u = ab/6 - c/2 - (a/3)^3$, $v = b/3 - (a/3)^2$ and $a = [3(m-r)x_{r+1} - 6w]/6m$, $b = [(m-r)x_{r+1}^2 - 3wx_{r+1}]/6m$, $c = -wx_{r+1}^2/6m$, he got the bounds

$$\widehat{\vartheta}_\ell \le \widehat{\vartheta}_{\mathsf{MLE}} \le \widehat{\vartheta}_u,$$

where

$$\widehat{\vartheta}_\ell = \left(u + \sqrt{u^2 + v^3}\right)^{1/3} + \left(u - \sqrt{u^2 + v^3}\right)^{1/3} - \frac{a}{3},$$

$$\widehat{\vartheta}_u = \min\left(\frac{w}{m-r}, \frac{\rho + \sqrt{\rho^2 + 24mwx_{r+1}}}{12m}\right).$$

Remark 12.1.3. Chandrasekar et al. [244] addressed minimum risk equivariant estimation for progressively Type-II censored exponential lifetimes. They showed that the estimator

$$\widehat{\vartheta}_\ell = \frac{\Gamma(m+\ell)}{\Gamma(m+2\ell)}\Big(\sum_{i=1}^m (R_i+1)Z_{i:m:n}\Big)^\ell = m\frac{\Gamma(m+\ell)}{\Gamma(m+2\ell)}\big(\widehat{\vartheta}_{\mathsf{MLE}}\big)^\ell$$

is a minimum risk equivariant (MRE) estimator for ϑ^ℓ. In particular, for $\ell = 1$, $\widehat{\vartheta}_{\mathsf{MRE}} = \frac{m}{m+1}\widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m+1}\sum_{i=1}^m (R_i+1)Z_{i:m:n}$ is a MRE of ϑ. Notice that it equals the BLEE $\widehat{\vartheta}_{\mathsf{LE}}$ given in (11.5).

Location Parameter Unknown

Now, suppose ϑ is known. Then, from (12.3), the likelihood function is proportional to

$$\exp\Big\{-n\frac{z_1-\mu}{\vartheta}\Big\}\mathbb{1}_{[\mu,\infty)}(z_1).$$

For arbitrary ϑ, this expression is zero on $(-\infty, z_1)$ and positive and decreasing on $[z_1,\infty)$. Thus, it is maximized by $\mu = z_1$, and we have thus proved that

$$\widehat{\mu}_{\mathsf{MLE}} = Z_{1:m:n}$$

is the maximum likelihood estimator of μ. Properties of $\widehat{\mu}_{\mathsf{MLE}}$ are presented in Theorem 12.1.4. As in the scale model, an MRE for the location parameter μ can be calculated. According to Chandrasekar et al. [244], it is given by $\widehat{\mu}_{\mathsf{MRE}} = \widehat{\mu}_{\mathsf{MLE}} - \frac{1}{n} = Z_{1:m:n} - \frac{1}{n}$ which equals also the BLUE (and BLEE) $\widehat{\mu}_{\mathsf{LU}}$ of μ [see (11.4)].

Location and Scale Parameters Unknown

Let $m \geq 2$ throughout this section. Now, we combine the above derivations to establish the MLEs in the location–scale model. First, we notice that, for any given value of ϑ, as in the location case, the likelihood in (12.3) is bounded by

$$\mathsf{L}(\mu,\vartheta;\mathbf{z}_m) \leq \frac{c_{m-1}}{\vartheta^m}\exp\Big\{-\frac{m}{\vartheta}\widetilde{\vartheta}\Big\}$$

with equality iff $\mu = z_1$. From the calculations in the scale case, we know that this bound is maximized by $\vartheta = \widetilde{\vartheta}$. Hence, we have proved that

$$\widehat{\mu}_{\mathsf{MLE}} = Z_{1:m:n} \quad \text{and} \quad \widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m}\sum_{j=2}^m (R_j+1)(Z_{j:m:n} - Z_{1:m:n}) \qquad (12.6)$$

are the MLEs of μ and ϑ in the location–scale model. These estimators have the following properties, which will be useful in the construction of confidence intervals. The results of Theorem 12.1.4 can be found in Viveros and Balakrishnan [875] or, in a more general framework, in Cramer and Kamps [299] and Balakrishnan et al. [130].

Theorem 12.1.4. Let $m \geq 2$. The MLEs $\widehat{\mu}_{\mathsf{MLE}}$ and $\widehat{\vartheta}_{\mathsf{MLE}}$ are independent with distributions

$$\widehat{\mu}_{\mathsf{MLE}} \sim \mathsf{Exp}(\mu, \vartheta/n), \quad \widehat{\vartheta}_{\mathsf{MLE}} \sim \Gamma(\vartheta/m, m-1).$$

Their expectations and variances are given by

(i) $E\widehat{\mu}_{\mathsf{MLE}} = \mu + \frac{\vartheta}{n}$, $\mathsf{Var}(\widehat{\mu}_{\mathsf{MLE}}) = \frac{\vartheta^2}{n}$,
(ii) $E\widehat{\vartheta}_{\mathsf{MLE}} = \frac{m-1}{m}\vartheta$, $\mathsf{Var}(\widehat{\vartheta}_{\mathsf{MLE}}) = \frac{m-1}{m^2}\vartheta^2$.

Moreover, $(\widehat{\mu}_{\mathsf{MLE}}, \widehat{\vartheta}_{\mathsf{MLE}})$ is a complete sufficient statistic for (μ, ϑ).

Proof. $\widehat{\vartheta}_{\mathsf{MLE}}$ can be written as

$$\begin{aligned}\widehat{\vartheta}_{\mathsf{MLE}} &= \frac{1}{m}\sum_{j=2}^{m}(R_j+1)(Z_{j:m:n} - Z_{1:m:n}) \\ &= \frac{1}{m}\sum_{j=2}^{m}(\gamma_j - \gamma_{j+1})(Z_{j:m:n} - Z_{1:m:n}) \\ &= \frac{1}{m}\sum_{j=2}^{m}\gamma_j(Z_{j:m:n} - Z_{j-1:m:n}).\end{aligned}$$

Hence, $\widehat{\vartheta}_{\mathsf{MLE}}$ is a function of the normalized spacings $S_j^{\mathscr{R}} = \gamma_j(Z_{j:m:n} - Z_{j-1:m:n})$, $2 \leq j \leq m$ [see (2.9)]. Since these spacings are independent and $\mathsf{Exp}(\vartheta)$-distributed (see Theorem 2.3.2), we find that $\widehat{\vartheta}_{\mathsf{MLE}}$ has a gamma distribution. Moreover, this yields the independence of $\widehat{\vartheta}_{\mathsf{MLE}}$ and $\widehat{\mu}_{\mathsf{MLE}} = S_1^{\mathscr{R}}/n + \mu$. This also implies $\widehat{\mu}_{\mathsf{MLE}} \sim \mathsf{Exp}(\mu, \vartheta/n)$. The expressions for mean and variance are immediate consequences of the distributional results. Finally, we mention that the likelihood function (12.3) can be written as

$$\mathsf{L}(\mu, \vartheta; \mathbf{z}_m) = \frac{c_{m-1}}{\vartheta^m}\exp\Big\{-\frac{m}{\vartheta}\widehat{\vartheta}_{\mathsf{MLE}} - n\frac{\widehat{\mu}_{\mathsf{MLE}} - \mu}{\vartheta}\Big\}\mathbb{1}_{[\mu,\infty)}(\widehat{\mu}_{\mathsf{MLE}}).$$

Applying the Neyman criterion for sufficiency of statistics (see, e.g., Lehmann and Casella [582]), $(\widehat{\mu}_{\mathsf{MLE}}, \widehat{\vartheta}_{\mathsf{MLE}})$ is a sufficient statistic for (μ, ϑ). Finally, the completeness as in Chiou and Cohen [262] (see also Cramer et al. [308]). □

Theorem 12.1.4 yields directly the UMVUEs of μ and ϑ as

$$\begin{aligned}\widehat{\mu}_{\mathsf{UMVUE}} &= Z_{1:m:n} - \frac{\widehat{\vartheta}_{\mathsf{UMVUE}}}{n},\\ \widehat{\vartheta}_{\mathsf{UMVUE}} &= \frac{1}{m-1}\sum_{j=2}^{m}(R_j+1)(Z_{j:m:n} - Z_{1:m:n}).\end{aligned} \tag{12.7}$$

They coincide with the BLUEs given in (11.6).

Remark 12.1.5. The distribution of the MLEs as given in Theorem 12.1.4 yields the distributions

$$n\frac{m-1}{m}\frac{\widehat{\mu}_{\mathsf{MLE}} - \mu}{\widehat{\vartheta}_{\mathsf{MLE}}} \sim \mathsf{F}(2, 2m-2), \quad 2m\frac{\widehat{\vartheta}_{\mathsf{MLE}}}{\vartheta} \sim \chi^2(2m-2).$$

As pointed out by Viveros and Balakrishnan [875], the independence can alternatively be shown by an application of Basu's theorem (see Hogg and Craig [446, pp. 390]).

Example 12.1.6. To illustrate the above results, we assume that Nelson's progressively Type-II censored insulating fluid data 1.1.4 is from an exponential distribution. The following example is presented in Viveros and Balakrishnan [875] for the scale case (see also Example 11.2.3).

For the scale model, the maximum likelihood estimate of the lifetime is given by $\widehat{\vartheta}^*_{\mathsf{MLE}} = 9.086$. The MLE of the reliability $R(t_0) = \mathsf{e}^{-t_0/\vartheta}$ at a given mission time $t_0 > 0$ is given by $\widehat{R}(t_0) = \mathsf{e}^{-t_0/\widehat{\vartheta}^*_{\mathsf{MLE}}}$. For $t_0 = 2$, we get $\widehat{R}(2) = 0.8024$.

In the location–scale model, we get the estimates $\widehat{\mu}_{\mathsf{MLE}} = 0.19$ and $\widehat{\vartheta}_{\mathsf{MLE}} = 9.8686$. The corresponding MLE of the reliability at t_0 is given by $\widehat{R}(t_0) = \mathsf{e}^{-(t_0 - \widehat{\mu}_{\mathsf{MLE}})/\widehat{\vartheta}_{\mathsf{MLE}}}$ provided that $t_0 > \widehat{\mu}_{\mathsf{MLE}}$. This leads to the estimate $\widehat{R}(2) = 0.8324$.

Remark 12.1.7. Balakrishnan et al. [133] compared the MLEs of location and scale parameters in a two-sample setting with identical censoring scheme $\mathscr{R}$. Given that $\mu_1 \le \mu_2$ and $\vartheta_1 \le \vartheta_2$ are the parameters in the samples, they found that the MLEs are stochastically ordered, i.e.,

$$\widehat{\mu}_{1;\mathsf{MLE}} \le_{\mathsf{st}} \widehat{\mu}_{2;\mathsf{MLE}}, \qquad \widehat{\vartheta}_{1;\mathsf{MLE}} \le_{\mathsf{st}} \widehat{\vartheta}_{2;\mathsf{MLE}}.$$

Remark 12.1.8. Large sample results like, e.g., asymptotic normality and consistency, for s independent samples are established in Balakrishnan et al. [130]. Moreover, explicit expressions for the MLEs and UMVUEs are available. For instance, the MLEs in the location–scale model are given by

$$\widehat{\mu}_{\mathsf{MLE}} = \min\{Z_{1,1:m_1:n_1}, \ldots, Z_{s,1:m_s:n_s}\},$$
$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m_\bullet} \sum_{i=1}^{s} \sum_{j=1}^{m_i} (R_{ij} + 1)(Z_{i,j:m_i:n_i} - \widehat{\mu}_{\mathsf{MLE}}), \tag{12.8}$$

where $m_\bullet = \sum_{i=1}^{s} m_i$. It has to be noted that, for $s \geq 2$, both estimators do no longer coincide with the BLUEs. In fact, the location estimator is the minimum of all observed values. Nevertheless, the independence of the MLEs is retained (see Cramer and Kamps [299]). Moreover, the MLEs are complete sufficient statistics.

Remark 12.1.9. For general progressive censoring, explicit expressions for the MLEs are available only in the one sample case (see Balakrishnan and Aggarwala [86]). In particular, for $r + 2 \leq m$, Balakrishnan and Sandhu [123] established the following expressions in the location–scale model:

$$\widehat{\mu}_{\mathsf{MLE}} = Z_{r+1:m:n} + \widehat{\vartheta}_{\mathsf{MLE}} \log\Big(1 - \frac{r}{n}\Big),$$
$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m-r} \sum_{j=r+2}^{m} (R_j + 1)(Z_{j:m:n} - Z_{r+1:m:n}). \tag{12.9}$$

In the multi-sample case, explicit expressions are not available. But, the estimates can be computed as (unique) solutions of the likelihood equations taking some constraints into account. For details, we refer to Balakrishnan et al. [130].

For a doubly Type-II censored sample $X_{r+1:n}, \ldots, X_{n-s:n}$, the expressions in (12.9) reduce to those established by Kambo [493]. Obviously, the MLEs are linear estimators. In fact, the BLUEs are bias-corrected versions of these expressions. This observation leads to expressions for means and variances of the MLEs. In particular, we have for doubly Type-II censored samples

$$\mathsf{Var}(\widehat{\mu}_{\mathsf{MLE}}) = \left\{ \sum_{j=1}^{r+1} \frac{1}{(n-j+1)^2} + (m-r-1)\left[\frac{\log(1-\frac{r}{m})}{m-r}\right]^2 \right\} \vartheta^2,$$
$$\mathsf{Var}(\widehat{\vartheta}_{\mathsf{MLE}}) = \frac{m-r-1}{(m-r)^2} \vartheta^2.$$

Remark 12.1.10. Simultaneous MREs of location and scale are given by

$$\widehat{\mu}_{\mathsf{MRE}} = Z_{1:m:n} - \frac{1}{n} \widehat{\vartheta}_{\mathsf{MRE}} \quad \text{and}$$
$$\widehat{\vartheta}_{\mathsf{MRE}} = \widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m} \sum_{j=2}^{m} (R_j + 1)(Z_{j:m:n} - Z_{1:m:n}),$$

as shown by Chandrasekar et al. [244].

Remark 12.1.11. Let a location–scale family of distributions be defined by

$$F(t) = 1 - \exp\left\{-\frac{d(t)-\mu}{\vartheta}\right\}, \quad t \geq d^{-1}(\mu),\ \mu \in \mathbb{R},\ \vartheta > 0, \tag{12.10}$$

where d is supposed to be strictly increasing and differentiable on $(d^{-1}(\mu), \infty)$ [see (2.12)]. Some important members of this family are exponential, Weibull, Pareto, and Lomax distributions. More details are provided in Cramer and Kamps [300]. Then, the maximum likelihood estimators are obtained directly from the results obtained for the exponential distributions. For instance, the simultaneous MLEs of μ and ϑ are given by

$$\widehat{\mu} = d(X_{1:m:n}), \quad \widehat{\vartheta} = \frac{1}{m}\sum_{j=2}^{m} \gamma_j \left[d(X_{j:m:n}) - d(X_{j-1:m:n})\right]. \tag{12.11}$$

For $\mu = 0$, (12.10) defines a one-parameter exponential family with scale parameter $1/\vartheta$ [see (2.12)]. It includes, e.g., the Weibull distribution with known shape parameter by choosing $d(t) = t^\beta$, $t > 0$. The MLE of ϑ is given by

$$\widehat{\vartheta} = \frac{1}{m}\sum_{j=1}^{m} \gamma_j \left[d(X_{j:m:n}) - d(X_{j-1:m:n})\right]$$

with $d(X_{0:m:n}) = 0$. Notice that the distributional results for the exponential distributions established in Theorems 12.1.1 and 12.1.4 also hold for this family of distributions since the progressively Type-II censored order statistics $d(X_{1:m:n}), \ldots, d(X_{m:m:n})$ can be seen as exponential progressively Type-II censored order statistics.

12.2 Weibull Distribution

As mentioned in Remark 12.1.11, the MLE of the scale parameter of a Weibull(ϑ, β)-distribution with known shape $\beta > 0$ parameter is given by

$$\widehat{\vartheta} = \frac{1}{m}\sum_{j=1}^{m} \gamma_j \left[X_{j:m:n}^{\beta} - X_{j-1:m:n}^{\beta}\right] = \frac{1}{m}\sum_{j=1}^{m} (R_j + 1) X_{j:m:n}^{\beta}. \tag{12.12}$$

The two-parameter setting is discussed in the following theorem.

Theorem 12.2.1. Let $m \geq 2$. Then, for a two-parameter Weibull distribution Weibull(ϑ, β), the maximum likelihood estimator $(\widehat{\vartheta}, \widehat{\beta})$ of (ϑ, β) uniquely exists.

The estimators are given by $\widehat{\vartheta} = \frac{1}{m}\sum_{j=1}^{m}(R_j+1)X_{j:m:n}^{\widehat{\beta}}$ where, for $X_{j:m:n} = x_j$, $1 \leq j \leq m$, $\widehat{\beta}$ is the unique solution of the equation

$$\frac{m}{\beta} + \sum_{j=1}^{m} \log x_j - \frac{\sum_{j=1}^{m}(R_j+1)\log(x_j)x_j^{\beta}}{\sum_{j=1}^{m}(R_j+1)x_j^{\beta}} = 0. \tag{12.13}$$

Proof. Without loss of generality, we can assume $0 < x_1 < \cdots < x_m$. For a two-parameter Weibull(ϑ, β)-distribution, the log-likelihood function is given by

$$\ell(\vartheta, \beta) = \sum_{j=1}^{m} \log \gamma_j + m\log\beta - m\log\vartheta + (\beta-1)\sum_{j=1}^{m}\log x_j - \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j+1)x_j^{\beta}. \tag{12.14}$$

Introducing the quantity $h(\beta) = \frac{1}{m}\sum_{j=1}^{m}(R_j+1)x_j^{\beta}$, we get with the inequality $\log t \geq 1 - \frac{1}{t}$, $t > 0$,

$$\begin{aligned}\ell(\vartheta, \beta) &= \text{const} + m\log\beta - \log h(\beta) + (\beta-1)\sum_{j=1}^{m}\log x_j \\ &\qquad - m\log\frac{\vartheta}{h(\beta)} - m\frac{h(\beta)}{\vartheta} \\ &\leq \text{const} + m\log\beta - m\log h(\beta) + (\beta-1)\sum_{j=1}^{m}\log x_j\end{aligned}$$

with equality iff $\vartheta = h(\beta)$. In order to find the MLE of (ϑ, β), we have to maximize the function

$$H(\beta) = m\log\beta - m\log h(\beta) + (\beta-1)\sum_{j=1}^{m}\log x_j$$

w.r.t. $\beta > 0$. First, notice that H is a continuous function with $\lim_{\beta\to 0} H(\beta) = -\infty$ and that

$$\lim_{\beta\to\infty}\frac{h'(\beta)}{h(\beta)} = \lim_{\beta\to\infty}\frac{\sum_{j=1}^{m}(R_j+1)\log(x_j)x_j^{\beta}}{\sum_{j=1}^{m}(R_j+1)x_j^{\beta}} = \log(x_m). \tag{12.15}$$

Then, using l'Hospital's rule and (12.15), we get

$$\lim_{\beta\to\infty} \frac{H(\beta)}{\beta} = -m \lim_{\beta\to\infty} \frac{\log h(\beta)}{\beta} + \sum_{j=1}^{m} \log x_j$$
$$= -m \log(x_m) + \sum_{j=1}^{m} \log x_j = \sum_{j=1}^{m-1} \log \frac{x_j}{x_m} < 0.$$

Hence, $\lim_{\beta\to\infty} H(\beta) = -\infty$. This proves that the maximum of the likelihood function is attained for some $\beta^* \in (0,\infty)$. Differentiating H w.r.t. β yields the Eq. (12.13). According to Balakrishnan and Kateri [105], this kind of equation has exactly one solution β^* (for a similar argument, see also Cramer and Kamps [295]). This proves the assertion. □

Remark 12.2.2. Wang et al. [890] proposed the so-called inverse estimators (IE) as an alternative to MLEs. The quantity

$$\tau(\mathbf{X}^{\mathscr{R}}, \beta) = 2\sum_{j=1}^{m-1} \log\left(\frac{\sum_{i=1}^{m}(R_i+1)X_{i:m:n}^{\beta}}{\sum_{i=1}^{j-1}(R_i+1)X_{i:m:n}^{\beta} + \gamma_j X_{j:m:n}^{\beta}}\right)$$

is χ^2-distributed with $2(m-1)$ degrees of freedom (see pp. 385). Moreover, $\tau(\mathbf{X}^{\mathscr{R}}, \beta)/(2m-4)$ converges in probability to 1. Then, the inverse estimator is defined as the unique solution of the equation $\tau(\mathbf{X}^{\mathscr{R}}, \beta) = 2m-4$ so that $\widehat{\beta}_{\mathsf{IE}} = \tau^{-1}(\mathbf{X}^{\mathscr{R}}, 2m-4)$. The estimator for the scale parameter ϑ is given by

$$\widehat{\vartheta}_{\mathsf{IE}} = \frac{1}{m-1}\sum_{i=1}^{m}(R_i+1)X_{i:m:n}^{\widehat{\beta}_{\mathsf{IE}}}.$$

Wang et al. [890] claimed that these estimators are consistently superior to the MLEs.

Example 12.2.3. For Nelson's progressively Type-II censored data 1.1.4, Viveros and Balakrishnan [875] computed the MLEs as $\widehat{\beta} = 0.975$ and $\widehat{\vartheta} = 9.225$ (see also Balakrishnan and Kateri [105]). The corresponding inverse estimates are $\widehat{\beta}_{\mathsf{IE}} = 0.7628$ and $\widehat{\vartheta}_{\mathsf{IE}} = 7.1320$.

Ng et al. [691] discussed a three-parameter Weibull distribution introducing an additional location parameter. Beside maximum likelihood estimation, they addressed censored estimation (see Harter and Moore [431] and Smith [809]), corrected and weighted maximum likelihood estimation, least-squares estimation, and other inferential methods for point estimation.

12.3 Reflected Power Distribution

We consider a location–scale family of $\mathsf{RPower}(\beta)$-distributions with shape parameter $\beta > 0$ and location and scale parameters μ and ϑ, respectively. Hence, the density function and cumulative distribution function are given by

$$f(t) = \frac{\beta}{\vartheta}\Big(1 - \frac{t-\mu}{\vartheta}\Big)^{\beta-1}, \quad F(t) = 1 - \Big(1 - \frac{t-\mu}{\vartheta}\Big)^{\beta}, \quad t \in [\mu, \mu+\vartheta],$$

where $\mu \in \mathbb{R}, \vartheta, \beta > 0$. For $\beta = 1$, it reduces to the uniform distribution. The likelihood function is given by (provided $x_1 < \cdots < x_m$)

$$\begin{aligned}
&\mathsf{L}(\mu, \vartheta, \beta; \mathbf{x}_m)\\
&\quad = c_{m-1}\frac{\beta^m}{\vartheta^m}\prod_{j=1}^{m}\Big(1 - \frac{x_j-\mu}{\vartheta}\Big)^{(R_j+1)\beta-1}\mathbb{1}_{[\mu,\mu+\vartheta]}(x_1)\mathbb{1}_{[\mu,\mu+\vartheta]}(x_m).
\end{aligned}$$

In order to get a positive likelihood, the restrictions

$$\mu \le x_1 < x_m \le \mu + \vartheta$$

must hold. This will be assumed subsequently.

Shape Parameter Known

If the parameter μ is known (without loss of generality, let $\mu = 0$), we get the log-likelihood (under the conditions $0 < x_1$ and $x_m \le \vartheta$)

$$\begin{aligned}
\ell(\vartheta) &= \sum_{j=1}^{m}\log\gamma_j + m\log\beta - m\log\vartheta + \sum_{j=1}^{m}\big((R_j+1)\beta - 1\big)\log\Big(1 - \frac{x_j}{\vartheta}\Big)\\
&= \sum_{j=1}^{m}\log\gamma_j + m\log\beta + \sum_{j=1}^{m}\big((R_j+1)\beta - 1\big)\log(\vartheta - x_j) - n\beta\log(\vartheta).
\end{aligned}$$

In order to find the maximum likelihood estimators, we have to consider several parameter setups separately. First, notice that $\lim_{\vartheta\to\infty}\ell(\vartheta) = -\infty$. For $R_m < \frac{1}{\beta} - 1$, the likelihood function is unbounded because $\lim_{\vartheta\to x_m}\ell(\vartheta) = \infty$. Hence, a maximum likelihood estimator of ϑ does not exist in this situation. For $R_m = \frac{1}{\beta} - 1$ and $\max_{1\le j\le m-1} R_j \le \frac{1}{\beta} - 1$, the log-likelihood function is decreasing in ϑ.

Therefore, the MLE is given by $\widehat{\vartheta}_{\mathsf{MLE}} = X_{m:m:n}$. These results are summarized in the following theorem.

Theorem 12.3.1. Given a reflected power distribution with $\beta = \frac{1}{R_m+1}$ and $\max_{1\le j\le m-1} R_j \le R_m$, the MLE of ϑ is given by $\widehat{\vartheta}_{\mathsf{MLE}} = X_{m:m:n}$.

For a reflected power distribution with $\beta > \frac{1}{R_m+1}$, the MLE of ϑ does not exist.

Henceforth, we can assume that $\max_{1\le j\le m-1} R_j > \frac{1}{\beta} - 1$. First, let $R_m > \frac{1}{\beta} - 1$. Then, we have $\lim_{\vartheta\to x_m} \ell(\vartheta) = -\infty$ so that the maximum of the log-likelihood function is attained at an inner point. This value has to solve the likelihood equation given by

$$\sum_{j=1}^{m} \big((R_j + 1)\beta - 1\big)\frac{\vartheta}{\vartheta - x_j} = n\beta, \quad x_m < \vartheta. \tag{12.16}$$

Now, for $\vartheta \to \infty$, the left-hand side of this equation converges to $n\beta - m < n\beta$. For $R_j \ge \frac{1}{\beta} - 1$, $1 \le j \le m-1$, and $R_m > \frac{1}{\beta} - 1$, this equation has a unique solution in (x_m, ∞) because $\lim_{\vartheta\to x_m} \ell(\vartheta) = \infty$ and the function on the right-hand side is strictly decreasing in ϑ. In the other cases, the monotonicity properties may be very complicated. We do only mention that, multiplying Eq. (12.16) by $\prod_{j\in J}(\vartheta - x_j)$ with $J = \{i|(R_i + 1)\beta \ne 1\}$, the resulting equation is a polynomial equation of degree $|J|$ which has at most $|J|$ real roots. Therefore, we have at most $|J|$ candidates for a maximum likelihood estimator.

For $R_m = \frac{1}{\beta} - 1$, Eq. (12.16) reads

$$\sum_{j=1}^{m-1} \big((R_j + 1)\beta - 1\big)\frac{\vartheta}{\vartheta - x_j} = n\beta, \quad x_m \le \vartheta.$$

Here, it may happen that the equation has no solution which implies that $X_{m:m:n}$ is the MLE of ϑ. Notice that in this case the log-likelihood function is bounded from above.

For the location–scale model, we have the restrictions $\mu \le x_1 < x_m \le \mu + \vartheta$. The log-likelihood is given by

$$\ell(\mu,\vartheta) = m\log\beta + \sum_{j=1}^{m}\log\gamma_j - n\beta\log\vartheta + \sum_{j=1}^{m}\big((R_j+1)\beta - 1\big)\log(\vartheta + \mu - x_j). \tag{12.17}$$

This leads to the likelihood equations

$$\begin{aligned}\sum_{j=1}^{m}\big((R_j+1)\beta-1\big)\frac{\vartheta}{\vartheta+\mu-x_j}&=n\beta,\\ \sum_{j=1}^{m}\big((R_j+1)\beta-1\big)\frac{1}{\vartheta+\mu-x_j}&=0.\end{aligned}\tag{12.18}$$

Multiplying the second equation by ϑ, we find that these equations contradict. Hence, (12.18) does not have a solution. This shows that the maximum is attained at the border of the feasible points, i.e., for $\mu = x_1$ or $\mu = x_m - \vartheta$. In order to find the corresponding solutions, many situations depending on the censoring plan $\mathscr{R}$ have to be discussed separately.

Shape Parameter Unknown

If β is assumed unknown, we get for the log-likelihood

$$\begin{aligned}&\ell(\vartheta,\mu,\beta)\\ &=\sum_{j=1}^{m}\log\gamma_j+m(\log\beta-\log\vartheta)+\sum_{j=1}^{m}\big((R_j+1)\beta-1\big)\log\Big(1-\frac{x_j-\mu}{\vartheta}\Big).\end{aligned}$$

Writing $\beta^* = \beta^*(\mu,\vartheta) = -\Big[\frac{1}{m}\sum_{j=1}^{m}(R_j+1)\log\Big(1-\frac{x_j-\mu}{\vartheta}\Big)\Big]^{-1} > 0$ and using the inequality $\log x \le x-1$, $x>0$, we get the upper bound

$$\begin{aligned}\ell(\vartheta,\mu,\beta)&=\sum_{j=1}^{m}\log\gamma_j+m\log\beta-m\log\vartheta-m\frac{\beta}{\beta^*}-\sum_{j=1}^{m}\log\Big(1-\frac{x_j-\mu}{\vartheta}\Big)\\ &\le\sum_{j=1}^{m}\log\gamma_j-m-m\log\vartheta-\sum_{j=1}^{m}\log\Big(1-\frac{x_j-\mu}{\vartheta}\Big)\\ &\qquad\qquad-m\log\Big[-\frac{1}{m}\sum_{j=1}^{m}(R_j+1)\log\Big(1-\frac{x_j-\mu}{\vartheta}\Big)\Big]\\ &=\ell(\vartheta,\mu,\beta^*)\end{aligned}$$

with equality iff $\beta = \beta^*$. This proves that β^* is the maximum likelihood estimator of β given μ and ϑ. In particular, this applies to the situation of known location and scale parameter. If one of these parameters is assumed unknown, the respective log-likelihood $\ell(\vartheta,\mu,\beta^*(\mu,\vartheta))$ has to be maximized.

12.4 Uniform Distribution

The uniform distribution is included in the reflected power distributions choosing $\beta = 1$. Hence, the likelihood function reduces to

$$\mathsf{L}(\mu, \vartheta; \mathbf{x}_m) = c_{m-1} \frac{1}{\vartheta^m} \prod_{j=1}^{m} \Big(1 - \frac{x_j - \mu}{\vartheta}\Big)^{R_j} \mathbb{1}_{[\mu,\mu+\vartheta]}(x_1) \mathbb{1}_{[\mu,\mu+\vartheta]}(x_m).$$

First, we consider the maximum likelihood estimation of μ, the left endpoint of the support. Obviously, the likelihood function is increasing in μ so that the MLE is given by $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$ no matter whether ϑ is known or not.

Location Parameter μ Known

Suppose the location parameter is known. Without loss of generality, let $\mu = 0$. Then, the likelihood function is given by

$$\ell(\vartheta) = \sum_{j=1}^{m} \log \gamma_j + \sum_{j=1}^{m} R_j \log(\vartheta - x_j) - m \log(\vartheta).$$

The likelihood equation is directly obtained from (12.16) by noticing that $\frac{\vartheta}{\vartheta - x_j} = 1 + \frac{x_j}{\vartheta - x_j}$ and that $\sum_{j=1}^{m} R_j = n - m$:

$$\sum_{j=1}^{m} \frac{R_j x_j}{\vartheta - x_j} = m, \quad x_m < \vartheta. \tag{12.19}$$

By analogy with the reflected power distribution, we have to take into account several possibilities depending on the censoring plan $\mathscr{R}$. First, let $R_m > 0$. Then, the left-hand side of the equation is a strictly decreasing function in ϑ with limits ∞ for $\vartheta \to x_m$ and 0 for $\vartheta \to \infty$, respectively. Therefore, we get a unique solution of the likelihood equation. This yields a maximum of the likelihood function since the second derivative of the log-likelihood function is obviously positive for $\vartheta > x_m$. If $R_m = 0$, we have to check whether $\mathscr{R} = 0$ or not. For $\mathscr{R} = 0$, the log-likelihood function is decreasing and bounded from above. Hence, the MLE is given by $X_{m:m:n}$. Notice that this setup corresponds to the situation of a complete sample with sample size $n = m$. For $\mathscr{R} \neq 0$ and $R_m = 0$, the left-hand side of (12.19) is decreasing in ϑ so that the equation has at most one solution. If $\sum_{j=1}^{m-1} \frac{R_j x_j}{x_m - x_j} < m$, the MLE is given by $X_{m:m:n}$.

This result can be extended to general progressive censoring (see, e.g., Aggarwala and Balakrishnan [13]).

Location Parameter μ Unknown

For unknown location parameter, the log-likelihood (12.17) reads with $\beta = 1$

$$\ell(\mu, \vartheta) = \sum_{j=1}^{m} \log \gamma_j - n \log \vartheta + \sum_{j=1}^{m} R_j \log(\vartheta + \mu - x_j).$$

For any $\vartheta \geq x_m - \mu$, this is an increasing function in μ so that $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$ is the maximum likelihood estimator of μ. Therefore, we have to maximize the function $\ell(x_1, \vartheta)$ w.r.t. $\vartheta > x_m - x_1$. Since this corresponds to the problem of a known location parameter with μ-value x_1, we get the following result which summarizes the above findings.

Theorem 12.4.1. The MLE of μ is given by $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$. For a given sample $x_1, \ldots, x_m$, the MLE of ϑ is computed as the unique solution of the equation

$$\sum_{j=1}^{m} \frac{R_j(x_j - \lambda)}{\vartheta + \lambda - x_j} = m, \quad x_m - \lambda < \vartheta, \quad \text{where } \lambda = \begin{cases} \mu, & \mu \text{ known} \\ x_1, & \mu \text{ unknown} \end{cases}.$$

If μ is assumed unknown, the observed sample must have $m \geq 2$ observations.

Notice that for an unknown location the likelihood equation simplifies to

$$\sum_{j=2}^{m} \frac{R_j(x_j - x_1)}{\vartheta + x_1 - x_j} = m.$$

12.5 Pareto Distributions

In this section, we consider Pareto distributions (Lomax distributions) with density function and cumulative distribution function given by

$$f(t) = \frac{\alpha}{\vartheta}\Big(1 + \frac{t-\mu}{\vartheta}\Big)^{-\alpha-1}, \quad F(t) = 1 - \Big(1 + \frac{t-\mu}{\vartheta}\Big)^{-\alpha}, \quad t \in [\mu, \infty),$$

where $\mu \in \mathbb{R}, \vartheta, \alpha > 0$. The likelihood function is given by (provided $x_1 < \ldots < x_m$)

$$\mathsf{L}(\mu, \vartheta, \alpha; \mathbf{x}_m) = c_{m-1} \frac{\alpha^m}{\vartheta^m} \prod_{j=1}^{m} \Big(1 + \frac{x_j - \mu}{\vartheta}\Big)^{-(R_j+1)\alpha - 1} \mathbb{1}_{(-\infty, x_1]}(\mu).$$

Since this function is increasing in $\mu \in (-\infty, x_1]$ for any $\vartheta > 0$ and $\alpha > 0$, the MLE of μ is given by $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$.

Shape Parameter Unknown, Scale Parameter Known

For convenience, we put $\vartheta = 1$. Then, the log-likelihood function is given by

$$\ell(\alpha) = \text{const} + m \log \alpha - \alpha \sum_{j=1}^{m} (R_j + 1) \log \Big(1 + x_j - \mu\Big) - \sum_{j=1}^{m} \log \Big(1 + x_j - \mu\Big),$$
$$\mu \le x_1. \quad (12.20)$$

From this expression, it is immediate that $\widehat{\alpha} = \frac{m}{\sum_{j=1}^{m} (R_j+1) \log \left(1 + X_{j:m:n} - \mu\right)}$ is the MLE of α. Combining the above results, we have proved the following theorem.

Theorem 12.5.1. For a given scale parameter, the MLE of μ is given by $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$. The MLE of α is given by

(i) $\widehat{\alpha} = \frac{m}{\sum_{j=1}^{m} (R_j+1) \log \left(1 + X_{j:m:n} - \mu\right)}$ if μ is known,

(ii) $\widehat{\alpha} = \frac{m}{\sum_{j=2}^{m} (R_j+1) \log \left(1 + X_{j:m:n} - X_{1:m:n}\right)}$ if μ is unknown.

Remark 12.5.2. For $\mu = 1$, the result of Part (i) of Theorem 12.5.1 reads

$$\widehat{\alpha} = \frac{m}{\sum_{j=1}^{m} (R_j + 1) \log \left(X_{j:m:n}\right)}.$$

Noticing that $R_j + 1 = \gamma_j - \gamma_{j+1}$, $1 \le j \le m - 1$ and that $R_m + 1 = \gamma_m$, we obtain the representation

$$\widehat{\eta} = \sum_{j=1}^{m} (R_j + 1) \log \left(X_{j:m:n}\right) = \sum_{j=1}^{m} \log \left(\left(\tfrac{X_{j:m:n}}{X_{j-1:m:n}}\right)^{\gamma_j}\right),$$

where $X_{0:m:n} = 1$. From Corollary 2.3.14, we conclude that the terms in the sum are IID Exp(α)-distributed random variables so that $\widehat{\eta}/(2\alpha)$ has a $\chi^2(2m)$-distribution. Therefore, $2m\alpha/\widehat{\alpha}$ follows a $\chi^2(2m)$-distribution (cf. Theorems 12.1.1 and 12.1.4 for the exponential distribution).

Shape Parameter Known, Scale Parameter Unknown

We can assume μ to be known in order to find the MLE of ϑ. For a given μ and $\mu < x_1 < \cdots < x_m$, the log-likelihood function is given by

$$\ell(\vartheta) = m\log\alpha + \sum_{j=1}^{m}\log\gamma_j - m\log\vartheta - \sum_{j=1}^{m}\big((R_j+1)\alpha+1\big)\log\big(1+\frac{x_j-\mu}{\vartheta}\big).$$

Differentiating ℓ w.r.t. ϑ results in the likelihood equation

$$-m + \sum_{j=1}^{m}\big((R_j+1)\alpha+1\big)\frac{x_j-\mu}{\vartheta+x_j-\mu} = 0$$

which is equivalent to

$$n\alpha - \sum_{j=1}^{m}\big((R_j+1)\alpha+1\big)\frac{\vartheta}{\vartheta+x_j-\mu} = 0. \tag{12.21}$$

Since the left-hand side is strictly decreasing in ϑ, (12.21) has at most one solution. Moreover, the limits for $\vartheta \to 0$ and $\vartheta \to \infty$ are $n\alpha > 0$ and $-m < 0$, so that (12.21) has a unique solution (for any μ). Noticing that $\lim_{\vartheta\to\infty}\ell(\vartheta) = -\infty$ and that, by l'Hospital's rule,

$$\lim_{\vartheta\to 0}\frac{\ell(\vartheta)}{\log\vartheta} = n\alpha > 0,$$

we get $\lim_{\vartheta\to 0}\ell(\vartheta) = -\infty$, so that the MLE of ϑ is the unique solution of (12.21). Combining the above arguments, we have proved the following theorem.

Theorem 12.5.3. Let either μ or ϑ be known. Then, for a Pareto distribution with $\alpha > 0$, the MLE of the other parameter always exists. The MLE of μ is given by $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$. The MLE $\widehat{\vartheta}$ of the scale parameter is defined as the unique solution of Eq. (12.21).

For unknown μ, we have to replace μ by x_1. Hence, (12.21) reads

$$\gamma_2\alpha - 1 - \sum_{j=2}^{m}\Big((R_j+1)\alpha+1\Big)\frac{\vartheta}{\vartheta+x_j-x_1} = 0. \tag{12.22}$$

Using the same arguments as above, we get that the left-hand side is decreasing in ϑ and that the limit for $\vartheta \to \infty$ is given by $-m$. But, for $\vartheta \to 0$, we obtain the limit $\gamma_2\alpha - 1$ which is positive iff $\gamma_2 > \frac{1}{\alpha}$. Hence, a sufficient and necessary condition for the MLE of ϑ to exist is $\gamma_2 > \frac{1}{\alpha}$.

Theorem 12.5.4. Let μ and ϑ be unknown. Then, for a Pareto distribution with $\alpha > 0$, the MLE of μ exists, and it is given by $\widehat{\mu}_{\mathsf{MLE}} = X_{1:m:n}$. The MLE $\widehat{\vartheta}$ of the scale parameter exists iff $m \geq 2$ and $\gamma_2 > \frac{1}{\alpha}$. In this case, it is defined as the unique solution of Eq. (12.22).

Example 12.5.5. For the rainfall data previously analyzed by linear inference in Example 11.2.9, we get the maximum likelihood estimates $\widehat{\mu} = x_1 = 0$ and $\widehat{\vartheta} = 2.56068$ provided that $\alpha = 1/2$. These values fit quite well to the results obtained for the best linear estimators: $\widehat{\mu}_{\mathsf{LU}} = -0.30484$ and $\widehat{\vartheta}_{\mathsf{LU}} = 3.20077$.

Remark 12.5.6. For $m = 2$, Eq. (12.22) can be solved explicitly. Noticing that $n = R_1 + R_2 + 2$, we get

$$(R_2 + 1)\alpha - 1 + \big((R_2 + 1)\alpha + 1\big)\frac{1}{1 + (x_2 - x_1)/\vartheta} = 0.$$

Provided that $\gamma_2 = R_2 + 1 > \frac{1}{\alpha}$, this yields

$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{(R_2 + 1)\alpha - 1}{2}(X_{2:m:n} - X_{1:m:n}).$$

Notice that the condition $R_2 + 1 > 1/\alpha$ ensures that $\widehat{\vartheta}_{\mathsf{MLE}}$ is nonnegative.

Shape and Scale Parameter Unknown

Similar to (12.20), we get for $\mu \leq x_1$ the log-likelihood

$$\ell(\alpha, \mu, \vartheta) = \text{const} + m \log \alpha - m \log \vartheta - \sum_{j=1}^{m} [\alpha(R_j + 1) + 1] \log\Big(1 + \frac{x_j - \mu}{\vartheta}\Big),$$

which is increasing in μ. Thus, $\ell(\alpha, \mu, \vartheta) \leq \ell(\alpha, x_1, \vartheta)$. Furthermore, for any ϑ, this term is bounded from above by

$$\ell(\alpha^*, x_1, \vartheta) = \text{const} - m \log\Big[\sum_{j=1}^{m}(R_j + 1)\log\big(1 + \frac{x_j - x_1}{\vartheta}\big)\Big] - \sum_{j=1}^{m} \log\big(\vartheta + x_j - x_1\big), \quad (12.23)$$

where $\alpha^* = \alpha^*(\vartheta) = \big[\frac{1}{m}\sum_{j=1}^{m}(R_j + 1)\log\big(1 + \frac{x_j - x_1}{\vartheta}\big)\big]^{-1}$. From (12.23), we find that $\lim_{\vartheta \to 0} \ell(\alpha^*, x_1, \vartheta) = -\infty$. Moreover, applying l'Hospital's rule to the ratio

$$q(\vartheta) = \frac{m \log\left[\sum_{j=1}^{m}(R_j+1)\log\left(1+\frac{x_j-x_1}{\vartheta}\right)\right]}{\sum_{j=1}^{m}\log\left(\vartheta+x_j-x_1\right)},$$

we get $\lim_{\vartheta\to\infty} q(\vartheta) = 0$ so that $\lim_{\vartheta\to\infty} \ell(\alpha^*, x_1, \vartheta) = -\infty$, too. This proves that the MLE is given by a solution of the likelihood equation.

Remark 12.5.7. Lomax distributions with known location parameter $\mu = 0$ and the cumulative distribution function $F(t) = 1-(1+\beta t)^{-\theta}$, $t > 0$, have been discussed in Helu et al. [438] when $\beta, \theta > 0$ are supposed to be unknown. As above, it is shown that the solution θ of the likelihood equations can be expressed in terms of β, i.e.,

$$\theta(\beta) = \left[\frac{1}{m}\sum_{j=1}^{m}(R_j+1)\log\left(1+\beta x_j\right)\right]^{-1}.$$

Instead of using a Newton–Raphson procedure, Helu et al. [438] proposed an EM-algorithm-based procedure to compute the MLEs.

Remark 12.5.8. Soliman [813] considered estimation for the Lomax distribution based on general progressively Type-II censored samples.

Shape and Scale Parameter Unknown with Equal Location Scale

Considering the model with unknown shape and scale parameters and assuming additionally $\mu = \vartheta$, we arrive at the Pareto distribution

$$f(t) = \frac{\alpha}{\vartheta}\left(\frac{t}{\vartheta}\right)^{-\alpha-1}, \quad F(t) = 1-\left(\frac{t}{\vartheta}\right)^{-\alpha}, \quad t \in [\vartheta, \infty).$$

For $\vartheta \le x_1$, the log-likelihood function is given by

$$\begin{aligned}\ell(\alpha,\vartheta) &= \text{const} + m\log\alpha - m\log\vartheta - \sum_{j=1}^{m}[\alpha(R_j+1)+1]\log\left(\frac{x_j}{\vartheta}\right)\\ &= \text{const} + m\log\alpha - \alpha\sum_{j=1}^{m}(R_j+1)\log x_j + \alpha\sum_{j=1}^{m}(R_j+1)\log\vartheta.\end{aligned}$$

Obviously, this expression is increasing in ϑ so that the maximum likelihood estimator is given by $\widehat{\vartheta}_{\mathsf{MLE}} = X_{1:m:n}$. Using the same arguments as above, an upper bound of the log-likelihood function is given by $\ell(\alpha^*, x_1)$ with $\alpha^* = \frac{1}{m}\Big[\sum_{j=2}^{m}(R_j +$

$1)\log\left(\frac{x_j}{x_1}\right)\Big]^{-1}$. Then, following the same arguments as in Remark 12.5.2, we conclude that $2m\alpha/\widehat{\alpha}$ follows a $\chi^2(2(m-1))$-distribution. Moreover, $\widehat{\alpha}$ and $\widehat{\vartheta}_{\mathsf{MLE}}$ are independent estimators (see also Balakrishnan and Aggarwala [86, pp. 129–130]).

This model has been considered in Ali Mousa [34] with a slightly different parametrization of the cumulative distribution function of the Pareto distribution:

$$F(t) = 1 - (\lambda t)^{-\vartheta}, \quad t > 1/\lambda, \lambda > 0, \vartheta > 0. \tag{12.24}$$

As above, the MLE of the scale parameter λ is given by $\widehat{\lambda} = X_{1:m:n}^{-1}$, whereas the MLE of the shape parameter is given by

$$\widehat{\vartheta} = \left[\frac{1}{m}\sum_{j=1}^{m}(R_j+1)\log\frac{X_{j:m:n}}{X_{1:m:n}}\right]^{-1}.$$

Ali Mousa [34] additionally presented the MLE $\widehat{R}(t)$ of the reliability $R(t) = \overline{F}(t)$ which is obtained directly from (12.24) by replacing the parameters by its maximum likelihood estimators.

12.6 Laplace Distribution

For a $\mathsf{Laplace}(\mu, \vartheta)$-distribution with density function

$$f(t) = \frac{1}{2\vartheta}\mathrm{e}^{-|t-\mu|/\vartheta}, \quad t \in \mathbb{R}, \tag{12.25}$$

and cumulative distribution function as in (11.17), maximum likelihood estimation based on complete samples has been discussed by a number of authors (see, for example, Johnson et al. [484], Kotz et al. [546]). Balakrishnan and Cutler [94] have discussed maximum likelihood estimation of parameters of the Laplace distribution based on conventionally Type-II censored samples. They considered both symmetric and one-sided (right) censoring. Childs and Balakrishnan [256] utilized these results to develop conditional inference procedures based on conventionally Type-II right censored samples. Proceeding similarly, Childs and Balakrishnan [257] have also derived the MLEs of the parameters μ and ϑ based on general conventionally Type-II censored samples. Recently, Iliopoulos and Balakrishnan [470] presented exact distributional results for the MLEs under Type-II censoring.

In this section, we discuss the MLEs of the location and scale parameters of a Laplace distribution based on progressively Type-II right censored samples as established in Aggarwala and Balakrishnan [14] (see also Balakrishnan and

Aggarwala [86, Sect. 7.3.6]). The results obtained are generalizations of those given in Balakrishnan and Cutler [94], wherein it is shown that for conventionally Type-II right censored samples $X_{1:n}, \dots, X_{m:n}$, the MLE of μ is simply the usual sample median based on the full sample, provided $m \geq \frac{n}{2}$. For $m < \frac{n}{2}$, the MLE of μ turns out to be a linear function of the observed order statistics. In both cases, the MLE of μ is a linear function of the observed order statistics. The results presented in this section for the maximum likelihood estimation based on progressively Type-II right censored samples from the Laplace distribution reduce to those presented by Balakrishnan and Cutler [94] for the special case when $\mathscr{R} = \mathscr{O}_m$.

For a $\mathsf{Laplace}(\mu, \vartheta)$-distribution and a progressively Type-II right censored sample, (12.1) yields the likelihood function

$$\mathsf{L}(\mu, \vartheta) = \mathsf{L}_i(\mu, \vartheta), \quad x_i \leq \mu \leq x_{i+1},\ i = 0, \dots, m,$$

where $x_0 = -\infty$, $x_{m+1} = \infty$ and

$$\mathsf{L}_i(\mu, \vartheta) = \frac{c_{m-1}}{2^n \vartheta^m} \exp\left\{\frac{1}{\vartheta}\Big(\sum_{j=1}^{i}(x_j - \mu) - \sum_{j=i+1}^{m}(R_j + 1)(x_j - \mu)\Big)\right\} \times \prod_{j=1}^{i}\Big(2 - \exp\Big\{\frac{x_j - \mu}{\vartheta}\Big\}\Big)^{R_j}.$$

Denoting by g the function defined by $g(t) = \log(2 - \mathrm{e}^t)$, $t < 0$, we get the piecewise defined log-likelihood function ℓ as

$$\ell_i(\mu, \vartheta) = \text{const} - m \log \vartheta + \sum_{j=1}^{i} \tfrac{x_j - \mu}{\vartheta} - \sum_{j=i+1}^{m} (R_j + 1) \tfrac{x_j - \mu}{\vartheta} + \sum_{j=1}^{i} R_j\, g\big(\tfrac{x_j - \mu}{\vartheta}\big),$$
$$x_i \leq \mu \leq x_{i+1},\ i = 0, \dots, m. \qquad (12.26)$$

Notice that g is a concave, decreasing, and nonnegative function on $(-\infty, 0]$. In order to find the MLEs, we distinguish three different situations.

Location Parameter $\mu \in \mathbb{R}$ Known and Scale Parameter $\vartheta > 0$ Unknown

Since μ is supposed to be known, we can find an interval $[x_i, x_{i+1}]$ such that $\mu \in [x_i, x_{i+1}]$. Therefore, we have to maximize the log-likelihood function ℓ_i. It follows from (12.26) that ℓ_i is a strictly concave function in $\eta = \frac{1}{\vartheta}$ so that we have a unique maximum of ℓ_i in $[x_i, x_{i+1}]$. Based on the identified interval $[x_i, x_{i+1}]$, we can provide some additional information on the MLE of ϑ. We consider the following situations:

(i) $i = 0$, i.e., $\mu < x_1$. Now, the log-likelihood function is given by

$$\ell_0(\mu, \vartheta) = \text{const} - m \log \vartheta - \sum_{j=i+1}^{m} (R_j + 1)\frac{x_j - \mu}{\vartheta}.$$

Obviously, it is maximized by

$$\widehat{\vartheta} = \frac{1}{m} \sum_{i=1}^{m} (R_i + 1)\,(x_i - \mu) > 0$$

which proves that $\widehat{\vartheta}$ is the MLE of ϑ provided that $\mu < x_1$.

(ii) Suppose $x_i \le \mu \le x_{i+1}$ for some $i \in \{1, \ldots, m-1\}$. Then, ℓ_i is given by (12.26). Since the log-likelihood function is concave w.r.t. $1/\vartheta$, we find that the likelihood equation

$$-m\vartheta - \sum_{j=1}^{i}(x_j - \mu) + \sum_{j=1}^{i} R_j \frac{x_j - \mu}{2\mathrm{e}^{-\frac{x_j - \mu}{\vartheta}} - 1} + \sum_{j=i+1}^{m} (R_j + 1)(x_j - \mu) = 0$$

has at most one solution. Noticing that

$$\ell_i(\mu, \vartheta) \overset{\vartheta \to 0}{\sim} -\frac{1}{\vartheta}\Big(\sum_{j=1}^{i} |x_j - \mu| + \sum_{j=i+1}^{m} (R_j + 1)|x_j - \mu|\Big),$$

we obtain $\lim_{\vartheta \to \infty} \ell_i(\mu, \vartheta) = -\infty$ and $\lim_{\vartheta \to 0} \ell_i(\mu, \vartheta) = -\infty$ so that we have exactly one solution of the likelihood equation. Using the inequality

$$0 < \frac{1}{2e^{-\frac{x_j - \mu}{\vartheta}} - 1} \le 1, \quad x_j \le \mu, 1 \le j \le i,$$

we conclude that the solution ϑ satisfies the inequalities

$$\begin{aligned} \vartheta &\le \frac{1}{m}\Big(\sum_{j=1}^{i} |x_j - \mu| + \sum_{j=i+1}^{m} (R_j + 1)|x_j - \mu|\Big), \\ \vartheta &\ge \frac{1}{m}\Big(\sum_{j=1}^{i}(1 - R_j)|x_j - \mu| + \sum_{j=i+1}^{m} (R_j + 1)|x_j - \mu|\Big). \end{aligned} \tag{12.27}$$

(iii) For $\mu > x_m$, the maximization problem is given by $\max_{\vartheta>0} \ell_m(\mu,\vartheta)$ with

$$\ell_m(\mu,\vartheta) = \text{const} - m\log\vartheta + \sum_{j=1}^{m}\frac{x_j-\mu}{\vartheta} + \sum_{j=1}^{m} R_j \log\Big[2 - \mathrm{e}^{\frac{x_j-\mu}{\vartheta}}\Big].$$

By analogy with the previous case, this is a concave function in $1/\vartheta$ with $\lim_{\vartheta\to\infty} \ell_m(\mu,\vartheta) = -\infty$ and $\lim_{\vartheta\to 0} \ell_m(\mu,\vartheta) = -\infty$. Moreover, we get by similar arguments the following bounds for the MLE of ϑ:

$$\frac{1}{m}\sum_{j=1}^{m}(1-R_j)|x_j-\mu| \le \widehat{\vartheta} \le \frac{1}{m}\sum_{j=1}^{m}|x_j-\mu|.$$

Remark 12.6.1. In the special case $\mathscr{R} = \mathscr{O}_m$, i.e., Type-II right censoring, the above results simplify considerably. For $\mu \le x_m$, the lower and upper bounds in (12.27) coincide so that the MLE is given by

$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m}\Big(\sum_{j=1}^{m-1}|X_{j:n}-\mu| + (n-m+1)|X_{m:n}-\mu|\Big).$$

Location Parameter $\mu \in \mathbb{R}$ Unknown and Scale Parameter $\vartheta > 0$ Known

First, we prove that the log-likelihood function (12.26) is concave in μ. We make use of the following theorem which is given in Hiriart-Urruty and Lemaréchal [441, pp. 35] for convex functions.

Theorem 12.6.2. Let f be a continuous function on an open interval (a,b) with $a,b \in \mathbb{R}$, $a = -\infty$, or $b = \infty$. Moreover, let $a = x_0 < x_1 < \cdots < x_m < x_{m+1} = b$ such that f is twice differentiable and concave on (x_i, x_{i+1}), $i = 0,\ldots,m$. Moreover, let

$$D_- f(x_i) \ge D_+ f(x_i), \quad 1 \le i \le m, \tag{12.28}$$

where $D_- f$ and $D_+ f$ denote the left and right derivative of f. Then, f is concave on (a,b).

In order to establish the concavity of the log-likelihood function in the location parameter μ, we have to verify the assumptions of the preceding theorem. It is easy to see that the log-likelihood function is continuous as well as twice differentiable and concave on (x_i, x_{i+1}), $0 \le i \le m$. Thus, it remains to show that the left and right derivatives w.r.t. μ satisfy the relation (12.28). Let $i \in \{1,\ldots m\}$. From (12.26), we have that

$$D_{-}\ell_{i-1}(x_i, \vartheta) = \frac{1}{\vartheta}\Big[\gamma_i - i + 1 - \sum_{j=1}^{i-1} R_j g'\Big(\frac{x_j - x_i}{\vartheta}\Big)\Big],$$

$$D_{+}\ell_i(x_i, \vartheta) = \frac{1}{\vartheta}\Big[\gamma_{i+1} - i - \sum_{j=1}^{i-1} R_j g'\Big(\frac{x_j - x_i}{\vartheta}\Big) - R_i g'(0)\Big].$$

Since $g'(0) = -1$ and $\gamma_i - \gamma_{i+1} - R_i = 1$, we get

$$\vartheta\Big(D_{-}\ell_{i-1}(x_i, \vartheta) - D_{+}\ell_i(x_i, \vartheta)\Big) = 2 \geq 0,$$

so that we arrive at the desired representation. In particular, we conclude from this result that the log-likelihood function is not differentiable in x_i, $i \in \{1, \ldots, m\}$. This yields the following theorem.

Theorem 12.6.3. The log-likelihood function is a strictly concave function in μ. Given data $x_1 < \cdots < x_m$, the MLE of μ can be obtained as follows: For $j \in \{1, \ldots, m\}$, let $a_j = D_{+}\ell_j(x_j, \vartheta)$.

(i) If $a_1 < 0$, then $\widehat{\mu}_{\mathsf{MLE}} = x_1$;
(ii) If $a_1 \geq \cdots \geq a_{i-1} > 0 \geq a_i$ with $2 \leq i \leq m$, then the MLE of μ is located in the interval $[x_{i-1}, x_i]$. For, $D_{-}\ell_{i-1}(x_i, \vartheta) \geq 0$, x_i denotes the MLE of μ. Otherwise, it is given by a solution of the likelihood equation

$$\gamma_i - i + 1 = \sum_{j=1}^{i-1} R_j g'\Big(\frac{x_j - \mu}{\vartheta}\Big). \tag{12.29}$$

If $R_j > 0$ for some $j \in \{1, \ldots, i-1\}$, the solution is unique;
(iii) If $a_m > 0$, then the MLE of μ is located in the interval $[x_m, \infty)$. It is the unique solution of the likelihood equation

$$-m = \sum_{j=1}^{m} R_j g'\Big(\frac{x_j - \mu}{\vartheta}\Big). \tag{12.30}$$

Proof. First, notice that $\ell_0(\cdot, \vartheta)$ is a linear increasing function for any ϑ. For $a_1 < 0$, $l_i(\cdot, \vartheta)$, $2 \leq i \leq m$ is decreasing so that x_1 yields the MLE of μ.

Let $a_{i-1} > 0 > a_i$ with $2 \leq i \leq m$. Then, for $D_{-}\ell_{i-1}(x_i, \vartheta) \geq 0$, x_i defines the MLE of μ. If $D_{-}\ell_{i-1}(x_i) < 0$, the MLE is an inner point of the interval (x_{i-1}, x_i). It is given by the solution of the likelihood equation $\frac{\partial \ell_{i-1}}{\partial \mu}(\mu, \vartheta) = 0$, which simplifies to (12.29). Notice that g is a strictly concave function so that g' is strictly decreasing on $[x_{i-1}, x_i]$. Hence, for $\sum_{j=1}^{i-1} R_j > 0$, the solution is unique.

For $a_m > 0$, we conclude that $\ell_m(\cdot, \vartheta)$ must be increasing–decreasing on $[x_m, \infty)$ because $\lim_{\mu\to\infty} \ell_m(\mu, \vartheta) = -\infty$ for any $\vartheta > 0$. Notice that $\lim_{t\to-\infty} g(t) = \log(2)$. Hence, the maximum is attained at an inner point which is specified by the likelihood Eq. (12.30). □

Remark 12.6.4. Equation (12.29) can be seen as a condition that the MLE is located in the ith interval $[x_{i-1}, x_i]$. Applying some inequalities, we can provide some simple conditions which exclude some intervals. It is important that the log-likelihood function is increasing–decreasing for every $\vartheta > 0$ and its derivative is first positive and then negative.

(i) The interval $(-\infty, x_i]$, $i \in \{1, \ldots, m\}$, cannot contain the MLE if the derivative $D_-\ell(\mu, \vartheta)$ is positive at the right endpoint x_i. Hence, we get the condition

$$D_+\ell(x_i, \vartheta) = \frac{\partial}{\partial\mu}\ell_i(\mu, \vartheta)\Big|_{\mu=x_i} = \frac{1}{\vartheta}\Big[\gamma_{i+1} - i - \sum_{j=1}^{i} R_j g'\Big(\frac{x_j - x_i}{\vartheta}\Big)\Big] > 0.$$

Since $g'\Big(\frac{x_j-\mu}{\vartheta}\Big) \le 0$, $1 \le j \le i$, and $g'(0) = -1$, this expression can be bounded from below by

$$D_+\ell(x_i, \vartheta) \ge \frac{1}{\vartheta}(\gamma_{i+1} - i - 1).$$

If $\gamma_{i+1} - i - 1 > 0$, then the maximum cannot be attained in the interval $(-\infty, x_i]$. Accordingly, a necessary condition for $(-\infty, x_i]$ to contain the MLE of μ is given by $\gamma_{i+1} \le i + 1$.

(ii) On the other hand, the interval (x_i, ∞), $i \in \{1, \ldots, m\}$, cannot contain the MLE if the derivative $D_+\ell(\mu, \vartheta)$ is negative at the left endpoint x_i. As above and using that $g'(0) = -1$, we get

$$D_+\ell(x_i, \vartheta) = \frac{\partial}{\partial\mu}\ell_i(\mu, \vartheta)\Big|_{\mu=x_i} = \frac{1}{\vartheta}\Big[\gamma_i - i - 1 - \sum_{j=1}^{i-1} R_j g'\Big(\frac{x_j - x_i}{\vartheta}\Big)\Big].$$

From $g'(t) \ge -1$, $t \le 0$, we find the upper bound

$$D_+\ell(x_i, \vartheta) \le \frac{1}{\vartheta}\Big[\gamma_i - (i + 1) + \sum_{j=1}^{i-1} R_j\Big] = n - 2i.$$

This expression is negative for $\frac{n}{2} < i$. Hence, for $\frac{n}{2} < i$, the MLE cannot be included in the interval (x_i, ∞).

Hence, we get the following conditions for the index $i \in \{1, \ldots, m\}$ of the interval containing the MLE of μ:

$$\gamma_{i+1} - 1 \leq i < \frac{n}{2}. \tag{12.31}$$

Notice that the upper bound is only useful for $\frac{n}{2} \leq m$.

This result is also helpful when the scale parameter is unknown. Since the bounds in (12.31) are independent of the particular value of the scale parameter ϑ, this reduces the number of maximization problems to be considered in the case when both parameters are unknown as well.

Summarizing the above findings, we get the following results for the MLE of μ. The value k is chosen as the minimum value of $i \in \{0, \ldots, m-1\}$ satisfying Eq. (12.31). Notice that this value always exists because $n = \gamma_1 > \gamma_2 > \cdots > \gamma_m \geq 1$:

$$\widehat{\mu}_{\mathrm{MLE}} \in \begin{cases} [X_{k+1:m:n}, \infty), & m < \frac{n}{2} \\ [X_{k+1:m:n}, X_{m:m:n}], & m = \frac{n}{2} \\ \left[X_{k+1:m:n}, X_{\lfloor n/2 \rfloor+1:m:n}\right], & m > \frac{n}{2} \end{cases}. \tag{12.32}$$

In particular, this result shows that $\widehat{\mu}_{\mathrm{MLE}} \geq X_{1:m:n}$.

Remark 12.6.5. It is obvious from (12.29) that the likelihood equation has a solution for $R_1 = \cdots = R_{i-1} = 0$ iff $\gamma_i = i + 1$. Now, given this condition $n = \gamma_i + i - 1$ we get the equation $n = 2i$ proving that this is only possible for an even sample size. In fact, we conclude that in this setting each value in $[x_{i-1}, x_i]$ yields the MLE of μ.

This result is known in the setting of order statistics (see Balakrishnan and Cutler [94]). In this situation, the log-likelihood function is a piecewise linear function of μ. Given a realization $x_1, \ldots, x_{n-m}$ of the order statistics $X_{1:n}, \ldots, X_{n-m:n}$, one has

$$a_j = D_+\ell_j(x_j, \vartheta) = \frac{\gamma_{j+1} - j}{\vartheta} = \frac{n - 2j}{\vartheta}, \quad j \in \{1, \ldots, m-1\}.$$

Thus, $a_j > 0 \iff \frac{n}{2} > j$. We distinguish the cases of n as odd or even.

Suppose n is odd. Then, $a_j > 0$, $j \in \{1, \ldots, \min(\frac{n-1}{2}, m-1)\}$. For $\frac{n-1}{2} \geq m-1$, the MLE is given by $X_{m:n}$. Otherwise, we have $a_j < 0$, $j \in \{\frac{n+1}{2}, \ldots, m\}$, and the MLE is given by $X_{(n+1)/2:n}$.

Suppose n is even. For $\frac{n}{2} \geq m-1$, the MLE is given by $X_{m:n}$. In the other case, we have the log-likelihood function to be constant in the interval $[x_{n/2}, x_{n/2+1}]$. Thus, we conclude from the concavity property that any value in this interval yields an MLE of μ. Summing up, we get the following result of Balakrishnan and Cutler [94] that the MLE of μ is given by

$$\widehat{\mu}_{\mathsf{MLE}} = \begin{cases} X_{m:n}, & \lfloor n/2 \rfloor \geq m-1 \\ X_{(n+1)/2:n}, & \lfloor n/2 \rfloor < m-1, n \text{ odd} \\ \alpha X_{n/2:n} + (1-\alpha) X_{n/2+1:n}, & \lfloor n/2 \rfloor < m-1, n \text{ even}, \alpha \in [0,1] \end{cases}.$$

Childs and Balakrishnan [257] established a generalization to general Type-II censored samples.

Location Parameter $\mu \in \mathbb{R}$ and Scale Parameter $\vartheta > 0$ Unknown

The preceding results can now be combined to derive the maximum likelihood estimators of location and scale parameters when both are unknown. We present a slight modification of an algorithm due to Aggarwala and Balakrishnan [14]. The procedure is based on the bounds for the MLE of μ given in (12.32).

Procedure 12.6.6 (Computation of MLEs for Laplace parameters). Let $k \in \{0, \ldots, m-1\}$ be as defined in (12.32).

(i) For $m < \frac{n}{2}$, the MLEs are given by the solutions of the maximization problem

$$\max_{k+1 \leq j \leq m} \max_{x_j \leq \mu \leq x_{j+1}, \vartheta > 0} \ell_j(\mu, \vartheta);$$

(ii) For $m > \frac{n}{2}$, the MLEs are given by the solutions of the maximization problem

$$\begin{matrix} \max\limits_{k+1 \leq j \leq n/2} \max\limits_{x_j \leq \mu \leq x_{j+1}, \vartheta > 0} \ell_j(\mu, \vartheta), & \text{if } n \text{ is even} \\ \max\limits_{k+1 \leq j \leq (n-1)/2} \max\limits_{x_j \leq \mu \leq x_{j+1}, \vartheta > 0} \ell_j(\mu, \vartheta), & \text{if } n \text{ is odd} \end{matrix};$$

(iii) For $m = \frac{n}{2}$, the MLEs are given by the solutions of the maximization problem

$$\max_{k+1 \leq j \leq n/2-1} \max_{x_j \leq \mu \leq x_{j+1}, \vartheta > 0} \ell_j(\mu, \vartheta).$$

Remark 12.6.7. (i) In case (ii) of Procedure 12.6.6, we may have $k = \frac{n-1}{2}$. This simply means that the log-likelihood function is increasing for $\mu < x_{(n+1)/2}$ and decreasing for $\mu > x_{(n+1)/2}$. Thus, the MLE of μ is $X_{(n+1)/2:m:n}$, which we can use to solve for the MLE of ϑ. The resulting likelihood equation to be solved for ϑ is given by

i	1	2	3	4	5	6	7	8	9	10
R_i	2	0	0	2	0	0	0	2	0	4
γ_i	20	17	16	15	12	11	10	9	6	5

Table 12.1 Censoring scheme $\mathscr{R}$ and related γ's for the situation of Example 11.2.10

$$-m\vartheta - \sum_{j=1}^{(n+1)/2-1}(x_j - x_{(n+1)/2}) + \sum_{j=1}^{(n+1)/2-1} R_j \frac{x_j - x_{(n+1)/2}}{2\mathrm{e}^{-\frac{x_j - x_{(n+1)/2}}{\vartheta}} - 1} + \sum_{j=(n+1)/2+1}^{m}(R_j+1)(x_j - x_{(n+1)/2}) = 0.$$

(ii) In case (iii) of Procedure 12.6.6, notice that here $k = \frac{n}{2} - 1 = m - 1$ is possible. This means that for $\mu < X_{m:m:n}$, the log-likelihood function is increasing and for $\mu > X_{m:m:n}$, the log-likelihood function is decreasing. Therefore, the MLE of μ is $X_{m:m:n}$. This can be used to solve for the MLE of ϑ. The resulting likelihood equation to be solved w.r.t. ϑ is given by

$$-m\vartheta - \sum_{j=1}^{m-1}(x_j - x_m) + \sum_{j=1}^{m-1} R_j \frac{x_j - x_m}{2\mathrm{e}^{-\frac{x_j - x_m}{\vartheta}} - 1} = 0.$$

Example 12.6.8. Let us consider the data in Example 11.2.10. In this case, we have a progressively Type-II right censored sample of size $m = 10$ from a sample of size $n = 20$ from the Laplace distribution with $\mu = 25$ and $\vartheta = 5$, with censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$, and the progressively Type-II right censored sample observed is as follows:

12.99290868	18.39049456	22.71250514	22.86934464	23.52886140
23.65727057	24.29938590	24.30197858	25.17997875	25.54438754

Now, we calculate the value of k given in Procedure 12.6.6. From Table 12.1, we conclude that $k = 7$. Thus, we have to solve two maximization problems:

$$\max_{x_8 \le \mu \le x_9, \vartheta > 0} \ell_8(\mu, \vartheta), \qquad \max_{x_9 \le \mu \le x_{10}, \vartheta > 0} \ell_9(\mu, \vartheta).$$

Using Maple 16, the maximum value of the log-likelihood function is obtained when we maximize $\ell_8(\mu, \vartheta)$ over the region specified above. The corresponding MLEs are $\widehat{\mu}_{\mathsf{MLE}} = 24.89573$ and $\widehat{\vartheta}_{\mathsf{MLE}} = 2.76910$. Recall that in Example 11.2.10, we determined the best linear unbiased estimates of μ and ϑ and their standard errors as

623	709	732	773	785	824	893	932	934	938
985	995	998	1002	1029	1079	1080	1096	1118	1122
1137	1162	1197	1225	1255	1292	1385	1385	1430	1485

Table 12.2 Random sample of size $n = 30$ from the incandescent lamps data by Davis [329]

623	709	732	773	785	893	932
934	938	985	995	998	1029	1079
1096	1122	1137	1162	1255	1385	1385

Table 12.3 Progressively Type-II censored sample of size $m = 21$ from the incandescent lamps data of Davis [329] generated by Childs and Balakrishnan [258]

$$\widehat{\mu}_{\mathsf{LU}} = 24.86538065, \qquad \widehat{\mathsf{SE}}(\widehat{\mu}_{\mathsf{LU}}) = 0.7100109565,$$

$$\widehat{\vartheta}_{\mathsf{LU}} = 2.910007086, \qquad \widehat{\mathsf{SE}}(\widehat{\vartheta}_{\mathsf{LU}}) = 0.8284699580.$$

These values agree well with the MLEs we have just obtained.

Example 12.6.9. Childs and Balakrishnan [258] also considered a progressively Type-II censored data generated from lifetimes of 417 incandescent lamps presented in Davis [329]. According to the analysis of the data by Davis [329], the Laplace distribution might be appropriate for this data. Childs and Balakrishnan [258] found that it may be used in favor of a normal distribution. In order to illustrate the above approach, Childs and Balakrishnan [258] generated the random sample of size $n = 30$ given in Table 12.2 from these observations. Employing the censoring plan $\mathscr{R} = (1, 0^{*2}, 2, 0, 1, 0^{*2}, 1, 0^{*3}, 1, 0^{*3}, 2, 0^{*3}, 1)$ to this sample, $m = 21$ failure times given in Table 12.3 result. According to the above procedure, the MLEs of μ and ϑ are given by $\widehat{\mu} = 1033.81667$ and $\widehat{\vartheta} = 182.86675$. Since $F^{\leftarrow}(t) = \mu - \vartheta \log(2 - 2t)$, $t \in (0, 1)$, this yields the maximum likelihood estimate of $\widehat{\xi}_{0.9} = 1328.13$ for the 90th quantile.

Remark 12.6.10. Similar results have been developed for record values by Cramer and Naehrig [304]. In this setting, an explicit representation for the MLEs of the parameters is available in the location–scale case.

12.7 Some Other Location–Scale Families

Suppose the progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$ is based on a general location–scale family of distributions $\mathscr{F}_{ls}$ as defined in (11.1). F denotes the absolutely continuous standard member of $\mathscr{F}_{ls}$ with density function f. We consider the location–scale case only. Similar results can be established in the location or scale cases as well. Given observations $x_1, \ldots, x_m$, the log-likelihood function $\ell(\mu, \vartheta) = \ell(\mu, \vartheta; \mathbf{x}_m)$ [see (12.2)] has derivatives

$$\frac{\partial \ell}{\partial \mu}(\mu, \vartheta) = -\frac{1}{\vartheta}\sum_{i=1}^{m}\frac{f'((x_i-\mu)/\vartheta)}{f((x_i-\mu)/\vartheta)} + \frac{1}{\vartheta}\sum_{i=1}^{m} R_i \frac{f((x_i-\mu)/\vartheta)}{1-F((x_i-\mu)/\vartheta)}, \tag{12.33a}$$

$$\frac{\partial \ell}{\partial \vartheta}(\mu, \vartheta) = -\frac{m}{\vartheta} - \frac{1}{\vartheta^2}\sum_{i=1}^{m} x_i \frac{f'((x_i-\mu)/\vartheta)}{f((x_i-\mu)/\vartheta)} + \frac{1}{\vartheta^2}\sum_{i=1}^{m} R_i x_i \frac{f((x_i-\mu)/\vartheta)}{1-F((x_i-\mu)/\vartheta)}. \tag{12.33b}$$

The resulting likelihood equations cannot be generally solved explicitly (some examples have been presented in the previous sections). In the following, we show results for some of the most important distributions.

12.7.1 Weibull Distributions

Cohen [270] has considered maximum likelihood estimation in the three-parameter Weibull distribution (related results are presented in Wingo [898] and Lemon [583]). For a sample $x_1 \leq \cdots \leq x_m$, the likelihood function [see (12.1)] reads

$$\mathsf{L}(\mu, \vartheta, \beta) = \prod_{j=1}^{m}\left[\frac{\gamma_j \beta}{\vartheta}\Big(\frac{x_j-\mu}{\vartheta}\Big)^{\beta-1}\right]\exp\Big\{-\sum_{j=1}^{m} R_j \Big(\frac{x_j-\mu}{\vartheta}\Big)^{\beta}\Big\}\mathbb{1}_{[\mu,\infty)}(x_1). \tag{12.34}$$

For $\beta < 1$ and any value of ϑ, the likelihood function is unbounded and tends to infinity when $\mu \to x_1$. Thus, formally, an MLE of μ does not exist. But, it is reasonable to choose the estimate $\widehat{\mu} = X_{1:m:n}$. In fact, Cohen [270] proposed the estimate $\widehat{\mu} = X_{1:m:n} - \frac{\eta}{2}$, where η denotes the unit of precision of measurements made. This ensures a finite likelihood. After plugging in this estimate, the problem reduces to the scale or scale-shape problem which has been addressed in Sect. 12.2.

For $\beta > 1$, the likelihood function is bounded. Thus, three likelihood equations have to be solved. Here again, the problem can be reduced to a system of two equations by eliminating the scale parameter ϑ. Rayleigh distributions (i.e., $\beta = 2$) are discussed in Ali Mousa and Al-Sagheer [36].

For the scale-shape model, a log-transformation of the sample yields progressively Type-II censored order statistics from a location–scale family of extreme value distributions (Type I). Here, Ng et al. [688] have proposed an EM-algorithm to compute the MLEs (see Sect. 12.7.4). This setting has also been discussed by Wu [904].

Kim and Han [529] considered maximum likelihood estimation for the scale parameter of a Rayleigh distribution based on a general progressively Type-II cen-

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
$x_{i:35:50}$	0.1	0.2	1	1	1	1	1	2	3	6	7	11	18	18	18	18	21	32
R_i	0	0	0	3	0	0	0	0	0	0	3	0	0	0	0	0	0	3
i	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	
$x_{i:35:50}$	36	45	47	50	55	60	63	63	67	67	75	79	82	84	84	85	86	
R_i	0	0	0	0	0	0	3	0	0	0	0	0	0	3	0	0	0	

Table 12.4 Generated progressively Type-II censored data with $m = 35$ and $n = 50$ from failure time data presented in Aarset [1]

sored sample. Ng [681] proposed a modified Weibull distribution with cumulative distribution function

$$F(t) = 1 - \mathrm{e}^{-\alpha t^{\beta} \mathrm{e}^{\lambda t}}, \quad t \geq 0,\ \alpha, \beta, \lambda > 0. \tag{12.35}$$

In order to illustrate his results, he generated progressively Type-II censored failure time data from a sample reported by Aarset [1]. The data and the censoring scheme $\mathscr{R}$ are presented in Table 12.4. It is used to illustrate bathtub-shaped distributions. Further inferential results on this distribution with progressively Type-II censored data are presented in Soliman et al. [818].

12.7.2 Normal Distributions

Balakrishnan et al. [134] discussed the maximum likelihood estimation for progressively Type-II censored normal samples. In this case, the likelihood equations, taken from (12.33), read

$$0 = \sum_{i=1}^{m} z_i + \sum_{i=1}^{m} R_i \cdot \frac{\varphi(z_i)}{1 - \Phi(z_i)}, \tag{12.36a}$$

$$0 = -m + \sum_{i=1}^{m} z_i^2 + \sum_{i=1}^{m} R_i z_i \cdot \frac{\varphi(z_i)}{1 - \Phi(z_i)}, \tag{12.36b}$$

where $z_i = (x_i - \mu)/\vartheta$, $1 \leq i \leq m$, and φ and Φ denote the density function and cumulative distribution function of a standard normal distribution. Cohen [267] has also discussed the solution of these equations in terms of progressive Type-I censoring. The likelihood equations have the same structure with x_i replaced by the censoring time T_i in the second sum in each equation. Showing log-concavity properties of the likelihood function, Balakrishnan and Mi [113] established the existence and uniqueness of the MLEs. This also holds for general progressive Type-II censoring.

Ng et al. [688] propose the EM-algorithm to compute the MLEs. Since this approach has been widely used for other distributions, we present the details for this particular case. We use the notation introduced in Sect. 9.1.2. The log-likelihood function based on the complete sample $(\mathbf{X}^{\mathscr{R}}, \mathbf{W})$ can be written as

$$\ell(\mu,\vartheta) = \text{const} - n\log\vartheta - \frac{1}{2\vartheta^2}\sum_{i=1}^{m}(x_i-\mu)^2 - \frac{1}{2\vartheta^2}\sum_{j=1}^{m}\sum_{k=1}^{R_j}(w_{jk}-\mu)^2. \tag{12.37}$$

To perform the EM-algorithm, W_{jk}^{ν} has to be replaced by $E(W_{jk}^{\nu}|W_{jk} > x_j)$, $\nu = 1, 2$, in the log-likelihood function (12.37), and the expression

$$-n\log\vartheta - \frac{1}{2\vartheta^2}\sum_{i=1}^{m}(x_i-\mu)^2 \\ - \frac{1}{2\vartheta^2}\sum_{j=1}^{m}\sum_{k=1}^{R_j}\Big[E(W_{jk}^2|W_{jk} > x_j) - 2\mu E(W_{jk}|W_{jk} > x_j) + \mu^2\Big]$$

has to be computed. Therefore, the conditional expectations $E(W_{jk}^{\nu}|W_{jk} > x_j)$, $\nu = 1, 2$, must be available. From Theorem 9.1.8, it follows that W_{jk}, given $X_{j:m:n} = x_j$, has a left-truncated normal distribution. Therefore,

$$E(W_{jk}|W_{jk} > x_j) = \mu + \vartheta h_j = \tau_j^{(1)}(\mu,\vartheta), \tag{12.38a}$$

$$E(W_{jk}^2|W_{jk} > x_j) = \vartheta^2(1+\xi_j h_j) + 2\vartheta\mu h_j + \mu^2 = \tau_j^{(2)}(\mu,\vartheta), \tag{12.38b}$$

where $\xi_j = (x_j - \mu)/\vartheta$, $h_j = \varphi(\xi_j)/(1-\Phi(\xi_j))$, $1 \le j \le m$ (see Cohen [272]). In the maximization step, the log-likelihood function (12.37) has to be maximized. The solution is given by mean and empirical standard deviation of the data. Hence, the $(\ell+1)$th iteration of the EM-algorithm is given by

$$\widehat{\mu}^{(\ell+1)} = \frac{1}{n}\Big[\sum_{j=1}^{m} x_j + \sum_{j=1}^{m} R_j \tau_j^{(1)}(\mu^{(\ell)},\vartheta^{(\ell)})\Big],$$

$$\widehat{\vartheta}^{(\ell+1)} = \frac{1}{\sqrt{n}}\Big[\sum_{j=1}^{m} x_j^2 + \sum_{j=1}^{m} R_j \tau_j^{(2)}(\mu^{(\ell)},\vartheta^{(\ell)}) - \widehat{\mu}^{(\ell+1)}\Big]^{1/2},$$

where $\tau_j^{(\nu)}(\mu^{(\ell)},\vartheta^{(\ell)})$, $\nu = 1, 2$ are taken from (12.38).

As mentioned by Ng et al. [688], the EM-algorithm converges often rather slowly in comparison to the Newton–Raphson procedure (in particular, if the proportion of missing data is large). However, it provides a measure of information in the censored data through the missing information principle (see Sect. 9.1.2). For the normal case, the information matrix for the complete sample is given by

$$\mathscr{I}(\mathbf{X};\mu,\vartheta) = \frac{n}{\vartheta^2}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

From the logarithm of the truncated normal density function

$$\log f^{W_{jk}|X_{j:m:n}}(w|x_j) = \text{const} - \log\vartheta - \log(1-\Phi(x_j)) - \frac{1}{2\vartheta^2}(w-\mu)^2,$$

we get via differentiation w.r.t. μ and ϑ:

$$\frac{\partial}{\partial\mu}\log f^{W_{jk}|X_{j:m:n}}(w|x_j) = \frac{1}{\vartheta}\left(\frac{w-\mu}{\vartheta} - h_j\right),$$

$$\frac{\partial}{\partial\vartheta}\log f^{W_{jk}|X_{j:m:n}}(w|x_j) = \frac{1}{\vartheta}\left(\frac{(w-\mu)^2}{\vartheta^2} - (1+\xi_j h_j)\right).$$

Calculating the conditional expectations $E((W_{jk}-\mu)^\nu|W_{jk}>x_j)$, $\nu = 1,2,3,4$, as a function of μ and ϑ, Ng et al. [688] get components of the conditional Fisher information matrix $\mathscr{I}^{(j)}(W_{jk}|X_{j:m:n};\mu,\vartheta)$ given in (9.8) to be

$$E\left[\left(\frac{\partial}{\partial\mu}\log f^{W_{jk}|X_{j:m:n}}(W_{jk}|X_{j:m:n})\right)^2\right] = \frac{1}{\vartheta^2}(1+\xi_j h_j - h_j^2),$$

$$E\left[\left(\frac{\partial}{\partial\vartheta}\log f^{W_{jk}|X_{j:m:n}}(W_{jk}|X_{j:m:n})\right)^2\right] = \frac{1}{\vartheta^2}(2+\xi_j h_j(1-\xi_j h_j+\xi_j^2)),$$

$$E\left[\frac{\partial}{\partial\vartheta}\log f^{W_{jk}|X_{j:m:n}}(W_{jk}|X_{j:m:n})\frac{\partial}{\partial\mu}\log f^{W_{jk}|X_{j:m:n}}(W_{jk}|X_{j:m:n})\right]$$
$$= \frac{1}{\vartheta^2}(h_j+\xi_j h_j(\xi_j - h_j)).$$

Thus, the expected Fisher information can be obtained from (9.9). Inversion of the Fisher information matrix $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\mu,\vartheta)$ yields the variance-covariance matrix of the maximum likelihood estimator $(\widehat{\mu}_{\mathsf{MLE}},\widehat{\vartheta}_{\mathsf{MLE}})$.

12.7.3 Log-Normal Distributions

Lifetimes are often modeled by log-normal distributions because the logarithm of the lifetime variables is normal. Moreover, a location–scale model with location parameter μ and scale parameter ϑ results. Since the log-transformation preserves the order of the random variables, a progressively Type-II censored sample from a log-normal distribution leads to progressively Type-II censored order statistics from a normal population with mean μ and variance ϑ^2. Supposing $X_{1:m:n}, \dots, X_{m:m:n}$ are the log-lifetimes in a progressively censored experiment, the likelihood equations are given by (12.36). They can be solved only by numerical methods. Early attempts are presented in Cohen [267] and Gajjar and Khatri [387]. Balakrishnan et al. [134] used the IMSL nonlinear equation solver using approximate MLEs as initial values for the iterative procedure. Ng et al. [688] applied the EM-algorithm (see Sect. 12.7.2; see also Singh et al. [804] who used the same account for progressively Type-II censored log-normal data). Basak et al. [179] have considered the application of the EM-algorithm for a three-parameter log-normal distribution based on progressively censored data.

12.7.4 Extreme Value Distribution (Type I)

Maximum likelihood estimation for progressively censored extreme value data has been addressed by Balakrishnan et al. [136]. The likelihood equations to be solved for μ and ϑ are

$$m = \sum_{j=1}^{m} (R_j + 1)\mathrm{e}^{\xi_j}, \quad m = \sum_{j=1}^{m} (R_j + 1)\xi_j \mathrm{e}^{\xi_j} - \sum_{j=1}^{m} \xi_j,$$

where $\xi_j = (x_j - \mu)/\vartheta$, $1 \le j \le m$. Obviously, these equations cannot be solved explicitly. Balakrishnan et al. [136] used the IMSL nonlinear equation solver to compute the estimates. The method has been applied to the log-failure times of Nelson's insulating fluid data as given in Table 17.5. The resulting estimates are given by $\widehat{\mu}_{\mathsf{MLE}} = 2.222$ and $\widehat{\vartheta}_{\mathsf{MLE}} = 1.026$ (see also Viveros and Balakrishnan [875]).

Numerical approaches to obtain the MLEs have also been discussed by Wingo [898], Lemon [583], and Cohen [270, 272] for progressive Type-I censoring. This setting leads to similar equations. Ng et al. [688] applied the EM-algorithm to compute the estimates. Moreover, they presented expressions for the observed Fisher information.

12.7.5 Logistic Distribution

Balakrishnan and Kannan [104] have considered the maximum likelihood estimation based on progressively Type-II censored data from location–scale families of logistic distributions with standard cumulative distribution function

$$F(t) = \frac{1}{1+\mathrm{e}^{-\pi t/\sqrt{3}}}, \quad t \in \mathbb{R}. \tag{12.39}$$

The standard member F of this family has mean 0 and variance 1. For more information on logistic distributions, we refer to Balakrishnan [81]. In this case, the likelihood equations become

$$m = \sum_{j=1}^{m} \frac{R_j + 2}{1+\mathrm{e}^{-\pi\xi_j/\sqrt{3}}}, \quad m\frac{\sqrt{3}}{\pi} = \sum_{j=1}^{m} \frac{(R_j+2)\xi_j}{1+\mathrm{e}^{-\pi\xi_j/\sqrt{3}}} - \sum_{j=1}^{m} \xi_j,$$

where $\xi_j = (x_j - \mu)/\vartheta$, $1 \le j \le m$. These equations can be solved numerically by a Newton–Raphson procedure. Expressions for the observed and expected Fisher information are also provided by Balakrishnan and Kannan [104]. Furthermore, we refer to Gajjar and Khatri [387]. Asgharzadeh [61] and Balakrishnan and Hossain [100] have addressed Type-I and Type-II generalized logistic distributions, respectively (for the classification, see Balakrishnan and Leung [109]). A scaled half-logistic distribution is considered in Balakrishnan and Asgharzadeh [87] (see also Kim and Han [530]). Exponentiated half-logistic distributions are addressed in Kang and Seo [506] and Rastogi and Tripathi [748].

12.8 Other Distributions

Maximum likelihood estimation based on progressively Type-II censored samples has been discussed for many other distributions. For completeness, we just mention briefly these works. Three-parameter Weibull distributions are addressed in Ng et al. [691]. Three-parameter gamma distributions are analyzed in Cohen and Norgaard [274] and Basak and Balakrishnan [176]. In the latter, an iterative procedure is proposed that can be used to compute the MLEs provided that the shape parameter exceeds 1. However, as pointed out in Laumen and Cramer [568], the likelihood equations presented in Cohen and Norgaard [274] seem to be in error which affects the follow-up papers, too. For further details, we refer to Laumen and Cramer [568]. Generalized gamma distributions as defined by Stacy [823] have been investigated in Chen and Lio [251]. In Wingo [901], progressively Type-I and Type-II data is assumed to follow a Burr-XII population distribution (see Sect. 13.5 and Lio et al. [615]). Progressive Type-II censoring in the Burr model is also discussed in Ali Mousa and Jaheen [38] and Soliman [811] who provided details about uniqueness and existence of the maximum likelihood estimates. Type-II censored data is considered in Wingo [900]. Log-gamma distributions with density function

$$f(t) = \frac{1}{\Gamma(\alpha)} \exp\left\{\alpha t - \mathrm{e}^t\right\}, \quad t \in \mathbb{R}, \alpha > 0, \tag{12.40}$$

are considered in Lin et al. [603] (for known α, see also Lin et al. [602]). The authors applied both a Newton–Raphson procedure and the EM-algorithm to compute the maximum likelihood estimates. Pradhan and Kundu [727] used the same approaches for generalized exponential distributions with cumulative distribution function

$$F(t) = (1 - \mathrm{e}^{-\lambda t})^{\alpha}, \quad t > 0, \tag{12.41}$$

and parameters $\alpha, \lambda > 0$. They illustrated their results by the data given in Table 12.5 which was generated from 36 failure times of appliances reported in Lawless [575]. It was shown by Pradhan and Kundu [727] that the original data fit a generalized exponential distribution very well. Scaled generalized exponential distributions are addressed in Asgharzadeh [62].

Inverse Weibull or Fréchet distributions with cumulative distribution function $F(t) = \mathrm{e}^{-\lambda t^{-\alpha}}$, $t > 0$, are discussed in Sultan et al. [827].

Two-parameter Gompertz distributions with cumulative distribution function

$$F(t) = 1 - \exp\Big\{ -\frac{\lambda}{c}(\mathrm{e}^{ct} - 1)\Big\}, \quad t \ge 0, \tag{12.42}$$

have been considered by Wu et al. [914] (for a generalization of Gompertz distributions (12.42), see Wu et al. [923]). Ghitany et al. [400] established a necessary and sufficient condition for the existence and uniqueness of the MLEs of the shape and scale parameters using a slightly different parametrization than that in (12.42).

A two-parameter bathtub-shaped lifetime distribution with cumulative distribution function

$$F(t) = 1 - \exp\Big\{\lambda\big(1 - \mathrm{e}^{t^{\beta}}\big)\Big\}, \quad t > 0, \tag{12.43}$$

has been considered by Wu [906] and Wu et al. [923] (see also Rastogi et al. [749], Sarhan et al. [778], and Ahmed [20]). Rastogi et al. [749] illustrated their results by the data given in Table 12.4.

Maxwell distributions are discussed in Krishna and Malik [552]. Results on likelihood inference for inverse Gaussian distributions can be found in Basak and Balakrishnan [177]. Bivariate normal distributions are addressed in Balakrishnan and Kim [107]. Birnbaum–Saunders distributions with density function

$(15, 5, 4, 0^{*9})$	11	35	49	329	1062	1167	1594	1990	2451	2471	2551	3059
$(0^{*11}, 24)$	11	35	49	170	329	381	708	958	1062	1167	1594	1925
$(24, 0^{*11})$	11	35	49	329	381	958	1062	1594	1925	2223	2451	2471

Table 12.5 Progressively Type-II censored data sets generated by Pradhan and Kundu [727] (see also Kundu and Pradhan [562])

$$f(t;\alpha,\beta) = \frac{1}{2\sqrt{2\pi}\alpha\beta}\sqrt{\frac{\beta}{t}}\Big[1+\frac{\beta}{t}\Big]\exp\Big\{-\frac{1}{2\alpha^2}\Big(\frac{t}{\beta}+\frac{\beta}{t}\Big)\Big\}, \quad t>0,$$

are addressed in Pradhan and Kundu [728]. The MLEs are computed by the EM-algorithm using the missing information principle.

Exponentiated distributions have been discussed by Ghitany et al. [398, 399] for a general class of exponentiated distributions having the cumulative distribution function F with

$$F(t) = \Big(1-\exp\big\{-\lambda Q(t)\big\}\Big)^{\theta}, \quad t \geq 0$$

where Q is an increasing function with $Q(0) = 0$ and $\lim_{t\to\infty} Q(t) = \infty$. The family includes exponentiated (inverted) exponential, exponentiated Rayleigh distribution, and exponentiated Pareto distribution. Using the transformation $Q^{\leftarrow}(Z_{j:m:n})$, $1 \leq j \leq m$, where $Z_{1:m:n}, \ldots, Z_{m:m:n}$ are exponential progressively Type-II censored order statistics, the problem can be embedded into exponentiated exponential distributions. The special case of inverted exponential distributions is discussed in Krishna and Kumar [551]. Klakattawi et al. [534] addressed an exponentiated version of the modified Weibull distribution given in (12.35).

Lindley distributions, i.e., a mixture of exponential and $\Gamma(2,\vartheta)$-distributions with the same scale parameter ϑ and mixture probabilities $1/(1+\vartheta)$ and $\vartheta/(1+\vartheta)$, are discussed in Krishna and Kumar [550].

Recently, Ahmadi et al. [19] addressed generalized half-normal distributions with cumulative distribution function

$$F(t;\alpha,\vartheta) = 1-2\Phi\Big(-\Big(\frac{t}{\vartheta}\Big)^{\alpha}\Big), \quad t \geq 0,\ \alpha,\vartheta > 0. \tag{12.44}$$

They presented the likelihood equations and proposed an EM-algorithm type procedure to compute the maximum likelihood estimators.

Given a cumulative distribution function $G(\cdot;\boldsymbol{\theta})$ with parameter $\boldsymbol{\theta}$, Teimouri et al. [842] discussed likelihood inference for beta-kernel distributions, i.e., population cumulative distribution functions

$$F(t;\alpha,\beta,\boldsymbol{\theta}) = \frac{1}{\mathsf{B}(\alpha,\beta)}\mathsf{B}_{G_{\boldsymbol{\theta}}(t)}(\alpha,\beta), \quad t \in \mathbb{R}, \tag{12.45}$$

with kernel $G_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_p) \in \Theta \subseteq \mathbb{R}^p$.

12.9 Related Methods

In some cases, the likelihood equations are difficult to solve in the sense that iterative procedures like the EM-algorithm or Newton–Raphson procedures do not converge.

Sometimes, they do not have solutions. For these situations, alternative approaches have been proposed. Subsequently, we illustrate some approaches by example and mention the corresponding references where these methods have been applied.

12.9.1 Modified Maximum Likelihood Estimation

As mentioned above in the Weibull case, the likelihood function may be unbounded so that formally maximum likelihood estimators do not exist or have regularity problems. In such cases, Cohen [270] has applied the method of modified maximum likelihood estimation to get estimates for the three-parameter Weibull distribution [see (12.34)]. Cohen and Whitten [277] favored modified maximum likelihood estimates (MMLE) over modified moment estimates for censored data.

In particular, the likelihood equation for the threshold parameter μ has to be replaced by the moment equation $EX_{1:m:n} = x_1$. For the three-parameter Weibull distribution, this equation reads

$$\mu + \Big(\frac{\vartheta}{n}\Big)^{1/\beta}\Gamma\Big(1+\frac{1}{\beta}\Big) = x_1.$$

Notice that $EX_{1:m:n}$ equals the expectation of the minimum in an IID sample of size n. Solving for μ yields the identity $\mu = x_1 - \big(\frac{\vartheta}{n}\big)^{1/\beta}\Gamma\big(1+\frac{1}{\beta}\big)$. This ensures a finite modified likelihood function which has to be maximized w.r.t. ϑ and β. Other possibilities to replace the mentioned likelihood equation are discussed in Cohen and Whitten [276]. For $\beta = 1$, the MMLEs for the two-parameter exponential distribution result.

It has to be noted that several possibilities for the moment equation have been discussed in the literature (see, e.g., Cohen and Whitten [276]). This may be of interest if the moment of $X_{1:m:n}$, or, more generally, of the rth progressively Type-II censored order statistic $X_{r:m:n}$, is difficult to obtain. Instead of this equation, we may impose the condition $EF(X_{r:m:n}) = F(x_r)$ for some $r \in \{1,\ldots,m\}$, where F is the population cumulative distribution function. Notice that $F(X_{r:m:n}) \stackrel{d}{=} U_{r:m:n}$ so that its value is given by (see Theorem 7.2.3)

$$EF(X_{r:m:n}) = EU_{r:m:n} = 1 - \prod_{j=1}^{r}\frac{\gamma_j}{1+\gamma_j}, \quad 1 \le r \le m.$$

Hence, applying the quantile function of the baseline distribution, a simpler equation may result. For the three-parameter Weibull distribution

$$F(t) = 1 - \exp\Big\{-\big(\tfrac{t-\mu}{\vartheta}\big)^{\beta}\Big\}, \quad t \ge \mu,$$

this leads to the alternative equations

$$\mu = x_r - \big(-\vartheta \log(1 - EU_{r:m:n})\big)^{1/\beta}, \quad r \in \{1, \dots, m\}.$$

The three-parameter gamma distribution has been addressed by Cohen and Norgaard [274]. Basak et al. [179] apply this approach for a three-parameter log-normal distribution (see also Cohen [271]). For the three-parameter inverse Gaussian distribution, we refer to Cohen and Whitten [277].

12.9.2 *Approximate Maximum Likelihood Estimation*

Approximate maximum likelihood estimators (AMLE) are calculated by expanding parts of the likelihood equations in Taylor series so that the resulting equations are simple functions of the unknown parameters (see Balakrishnan and Varadan [127]). The AMLE method often yields explicit estimators which is advantageous over the MLE method. Moreover, the AMLEs may be used as initial values for the iterative solution of the likelihood equations to compute the MLEs.

As an example, we present the method for the two-parameter extreme value distribution as presented in Balakrishnan et al. [136] and for the Weibull distribution as applied in Balasooriya et al. [163]. Notice that the approaches are slightly different.

Extreme Value Distribution

First, the partial derivatives of the log-likelihood function are given by

$$\begin{aligned}
\frac{\partial \ell}{\partial \mu}(\mu, \vartheta) &= -\frac{m}{\vartheta} + \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j + 1)\mathrm{e}^{\xi_j}, \\
\frac{\partial \ell}{\partial \vartheta}(\mu, \vartheta) &= -\frac{m}{\vartheta} - \frac{1}{\vartheta}\sum_{j=1}^{m}\xi_j + \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j + 1)\xi_j\mathrm{e}^{\xi_j},
\end{aligned} \tag{12.46}$$

where $\xi_j = (x_j - \mu)/\vartheta$, $1 \le j \le m$. Now, we denote by $\nu_{r:m:n} = 1 - \prod_{j=1}^{r} \frac{\gamma_j}{1+\gamma_j}$ the mean of the rth uniform progressively Type-II censored order statistic $U_{r:m:n}$ and expand $h(\xi_j) = \mathrm{e}^{\xi_j}$ around $\nu_{j:m:n}$ in a Taylor series of order 1 to get

$$h(\xi_j) \doteq h(\nu_{j:m:n}) + (\xi_j - \nu_{j:m:n})h'(\nu_{j:m:n}).$$

Hence, we get

$$e^{\xi_j} \doteq \alpha_j + \beta_j \xi_j$$

where $\alpha_j = e^{\nu_{j:m:n}}(1 - \nu_{j:m:n})$ and $\beta_j = e^{\nu_{j:m:n}}$, $1 \le j \le m$. This yields the approximations of (12.46) as

$$\begin{aligned}
\frac{\partial \ell}{\partial \mu}(\mu, \vartheta) &\doteq -\frac{m}{\vartheta} + \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j + 1)\alpha_j + \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j + 1)\beta_j\xi_j = 0,\\
\frac{\partial \ell}{\partial \vartheta}(\mu, \vartheta) &\doteq -\frac{m}{\vartheta} - \frac{1}{\vartheta}\sum_{j=1}^{m}\xi_j + \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j + 1)\alpha_j\xi_j + \frac{1}{\vartheta}\sum_{j=1}^{m}(R_j + 1)\beta_j\xi_j^2 = 0,
\end{aligned} \tag{12.47}$$

which is a system of linear and quadratic equations in μ and ϑ. First, we get

$$\mu = K + L\vartheta$$

with $K = \frac{\sum_{j=1}^m (R_j+1)\beta_j x_j}{\sum_{j=1}^m (R_j+1)\beta_j}$ and $L = \frac{\sum_{j=1}^m (R_j+1)\alpha_j - m}{\sum_{j=1}^m (R_j+1)\beta_j}$. ϑ is the unique positive solution of the quadratic equation

$$m\vartheta^2 - A\vartheta - B = 0$$

with $A = \sum_{j=1}^m [(R_j+1)\alpha_j - 1](x_j - K)$ and $B = \sum_{j=1}^m (R_j+1)\beta_j(x_j - K)^2 \ge 0$. Hence, we get the AMLEs as

$$\widehat{\mu}_{\mathsf{AMLE}} = K + L\widehat{\vartheta}_{\mathsf{AMLE}}, \quad \widehat{\vartheta}_{\mathsf{AMLE}} = \frac{1}{2m}\Big(A + \sqrt{A^2 + 4mB}\Big). \tag{12.48}$$

For a Type-II censored sample, i.e., $\mathscr{R} = \mathscr{O}_m$, the AMLEs derived by Balakrishnan and Varadan [127] result.

Weibull Distribution

Balasooriya et al. [163] have established approximate MLEs for a Weibull distribution parametrized as $\mathsf{Weibull}(\sigma^\beta, \beta)$. Transforming the distribution to an extreme value distribution, log-lifetimes are used in the analysis. Thus, the log-likelihood function is given by (12.46) with $\xi_j = (x_j - \mu)/\vartheta$, $j = 1, \ldots, m$, where $\mu = \log\sigma$ and $\vartheta = 1/\beta$. Denoting by $\mu_{j:m:n}$, $j = 1, \ldots, m$, the moments of progressively Type-II censored order statistics from a standard extreme value distribution given in (7.16), Balasooriya et al. [163] expand the exponential function around $\mu_{j:m:n}$ in a Taylor series of order 1. Since the approximation proceeds only by a different expansion point, this yields similar approximated likelihood equations as in (12.47), where now $\alpha_j = \mathrm{e}^{\mu_{j:m:n}}(1-\mu_{j:m:n})$ and $\beta_j = \mathrm{e}^{\mu_{j:m:n}}$, $1 \le j \le m$. Hence, the estimators $\widehat{\mu}, \widehat{\vartheta}$ have the same form as in (12.48). Transforming back to the Weibull distribution, this yields the estimators $\widehat{\sigma}_{\mathsf{AMLE}} = \mathrm{e}^{\widehat{\mu}}$ and $\widehat{\beta}_{\mathsf{AMLE}} = 1/\widehat{\vartheta}$.

Using arguments of Balakrishnan and Varadan [127], Balasooriya et al. [163] proposed approximate bias-corrected estimators of $\widehat{\mu}, \widehat{\sigma}$ as

$$\widehat{\mu}_c = \widehat{\mu} - a\widehat{\vartheta}, \quad \widehat{\vartheta}_c = \frac{\widehat{\vartheta}}{b}$$

with some constants a and b. The resulting estimators for the Weibull distribution are given by

$$\widehat{\sigma}^*_{\mathrm{AMLE}} = \mathrm{e}^{\widehat{\mu_c}}, \quad \widehat{\beta}^*_{\mathrm{AMLE}} = \frac{1}{\widehat{\vartheta}_c}. \tag{12.49}$$

Example 12.9.1. As mentioned in Example 12.2.3, Viveros and Balakrishnan [875] computed the MLEs $\widehat{\beta} = 0.975$ and $\widehat{\vartheta} = 9.226$ for Nelson's progressively Type-II censored data 1.1.4. For the estimators given in (12.49), Balasooriya et al. [163] obtained the scale estimate 9.107 (9.806 with bias correction) and the shape estimate 0.973 (1.013 with bias correction).

Balasooriya et al. [163] presented also Monte Carlo experiments to study the finite-sample properties of the AMLEs.

Other Distributions

Lin et al. [603] considered approximate maximum likelihood estimates for a three-parameter log-gamma distribution with density function

$$f(t) = \frac{\beta^{\beta-1/2}}{\Gamma(\beta)} \exp\left\{\sqrt{\beta}\frac{t-\mu}{\vartheta} - \beta\exp\left(\frac{t-\mu}{\vartheta\sqrt{\beta}}\right)\right\},\ t \in \mathbb{R}, \beta > 0, \vartheta > 0, \mu \in \mathbb{R},$$

based on a progressively Type-II censored sample. This parametrization includes the location–scale family of the extreme value distribution for $\beta = 1$. They extended the results of Balakrishnan and Chan [90, 91] for Type-II censored data. Some other distributions that have been studied are logistic distributions (see Balakrishnan and Kannan [104]), half-logistic distributions (see Balakrishnan and Asgharzadeh [87] and Kim and Han [530]), Gaussian distributions (see Balakrishnan et al. [134]), generalized half-normal distributions as in (12.44) (see Ahmadi et al. [19]), bivariate normal distributions (see Balakrishnan and Kim [107]), and Type-II generalized logistic distributions (see Balakrishnan and Hossain [100]).

12.10 *M*-Estimation

Given progressively Type-II censored data $X_{1:m:n}, \ldots, X_{m:m:n}$ with realizations $x_1, \ldots, x_m$, Basak and Balakrishnan [174] discussed the robust estimation of a location parameter. According to (12.2), the likelihood equation reads, with the notation $\rho_1(\cdot;\theta) = \log f_\theta$ and $\rho_2(\cdot;\theta) = \log \overline{F}_\theta$,

$$\sum_{j=1}^{m} \rho_1'(x_j;\theta) + \sum_{j=1}^{m} R_j \rho_2'(x_j;\theta) = 0. \tag{12.50}$$

Since MLEs are sensitive to outliers in the data, the derivatives ρ_j', $j = 1, 2$, are replaced by appropriately chosen functions ψ_1 and ψ_2 in robust *M*-estimation as proposed by Huber [464] (see also Huber [465]). Hence, we have to solve the equation

$$\sum_{j=1}^{m} \psi_1(x_j;\theta) + \sum_{j=1}^{m} R_j \psi_2(x_j;\theta) = 0 \tag{12.51}$$

instead of (12.50). Noticing that $\rho_2'(x;\theta) = E(\rho_1'(X;\theta)|X > x)$, Basak and Balakrishnan [174] considered a class of robust *M*-estimators obtained from (12.51) with ψ_2 given by $\psi_2(x;\theta) = E(\psi_1(X;\theta)|X > x)$. The resulting estimators are called James-type *M*-estimators (see James [478]).

In order to establish an expression for the influence function, Basak and Balakrishnan [174] considered a block censoring model as described by Hofmann et al. [444] with $R_0 = 0$ and censoring scheme

$$\mathscr{R}_m = (0^{*\overline{R}_1}, R_1, 0^{*\overline{R}_2}, R_2, \ldots, 0^{*\overline{R}_m}, R_m)$$

(see also Sect. 3.4.6). By similarity with (3.19), we use the following notation of Basak and Balakrishnan [174]:

(i) $p_i^* = \lim_{n\to\infty} \frac{\overline{R}_i}{n}$, $q_i = \lim_{n\to\infty} \frac{R_i}{n}$, $i = 1, \ldots, m$,
(ii) p_i denotes the (asymptotic) proportion of uncensored observations up to step i, $i = 1, \ldots, m$,
(iii) $\xi_p = F^{\leftarrow}(p)$ denotes the pth quantile of F, $p \in (0, 1)$.

Then, Basak and Balakrishnan [174] found the following representation of the influence function $\mathsf{IF}(\cdot; T, F)$ of $T(F)$ which denotes the functional form of the *M*-estimator

$$\mathsf{IF}(x; T, F) = -\frac{N(x; T, F)}{D_1(x; T, F)},$$

where

$$N(x;T,F) = \psi_1(x;T(F))\Big(\mathbb{1}_{(-\infty,\xi_{p_1})}(x) + \sum_{i=1}^{m-1} \frac{1-p^*_{\bullet i}-q_{\bullet i}}{1-p_i}\mathbb{1}_{(\xi_{p_i},\xi_{p_{i+1}})}(x)\Big) + \sum_{i=1}^{m} q_i\psi_2(\xi_{p_i};T(F)),$$

$$D_1(x;T,F) = \int_{-\infty}^{\xi_{p_1}} \psi_1'(y,T(F))dF(y) + \sum_{i=1}^{m-1}\frac{1-p^*_{\bullet i}-q_{\bullet i}}{1-p_i}\int_{\xi_{p_i}}^{\xi_{p_{i+1}}}\psi_1'(y,T(F))dF(y) + \sum_{i=1}^{m} q_i\psi_2'(\xi_{p_i};T(F)).$$

If the estimator is Fisher consistent for $F = F_\theta$, the influence function reads

$$\mathsf{IF}(x;T,F_\theta) = -\frac{N(x;T,F_\theta)}{D_2(x;T,F_\theta)}$$

with $N(x;T,F_\theta)$ as above (replace $T(F)$ by θ) and

$$D_2(x;T,F_\theta) = \int_{-\infty}^{\xi_{p_1}} \psi_1'(y,\theta)\Lambda_\theta(y)dF_\theta(y) + \sum_{i=1}^{m-1}\frac{1-p^*_{\bullet i}-q_{\bullet i}}{1-p_i}\int_{\xi_{p_i}}^{\xi_{p_{i+1}}}\psi_1'(y,\theta)dF_\theta(y) + \sum_{i=1}^{m} q_i\psi_2'(\xi_{p_i};\theta),$$

where $\Lambda_\theta(y) = \frac{\partial}{\partial\theta}\log f_\theta(y) = \frac{f_\theta'(y)}{f_\theta(y)}$.

Applying Theorem 2.6 of Huber [465, pp. 54], Basak and Balakrishnan [174] obtained results for the breakdown point of $T(F,\psi_1,\psi_2)$ in the location case. Supposing that ψ_i is increasing and takes values of both signs, the breakdown point is given by

$$\min\Big\{\frac{-\psi_1(-\infty)}{\psi_1(\infty)-\psi_1(-\infty)}, \frac{\psi_1(\infty)p_{\bullet m}+\psi_2(\infty)(1-p_{\bullet m})}{\psi_1(\infty)-\psi_1(-\infty)}\Big\}.$$

For James-type M-estimators of location, the breakdown point simplifies to

$$\frac{\min\{-\psi_1(-\infty), \psi_1(\infty)\}}{\psi_1(\infty) - \psi_1(-\infty)}.$$

The gross-error sensitivity $\gamma_F = \sup_x |\mathsf{IF}(x; T, F)|$ serves as a measure of robustness. The most robust estimator minimizes γ_F. For a location parameter, this leads to the median, $\widehat{\theta} = \text{med}(F)$. For Type-II right censored data, Akritas et al. [25] proved that the most robust estimator results from the choice $\psi_1^*(x) = \text{sgn}(x)$ and $\psi_2^*(x) = 1$. If the proportion of uncensored observations exceeds 0.5, the most robust estimator is the median, too. For progressive censoring, Basak and Balakrishnan [174] stated that, given $p_1^* = \cdots = p_m^*$, the most robust estimator is obtained from the choice $\psi_1^*(x) = \text{sgn}(x)$ and $\psi_2^*(x) = 0$. For some i with $\xi_{p_i} \le \theta < \xi_{p_{i+1}}$, the most robust estimator is given by $\widehat{\theta} = F^{\leftarrow}(p_i)$.

Assuming a parametric family $(F_\theta)_{\theta \in \Theta}$ with some regularity conditions imposed on f_θ and Θ, Basak and Balakrishnan [174] discussed optimal robust estimation for two choices of ψ-functions. They obtained results for the optimal choice of these functions which lead to optimal robust estimators solving (12.51). The results were illustrated by modifying Nelson's insulating fluid data 1.1.4.

12.11 Order Restricted Inference

Bhattacharya [198] and Beutner and Kamps [195] have considered order restricted inference in the k-sample case with censored data. Bhattacharya [198] discussed general progressive censoring from two-parameter exponential distributions, whereas Beutner and Kamps [195] addressed sequential order statistics from location–scale families with cumulative distribution function in (12.10). Notice that $d \equiv \text{id}$, i.e., $d(t) = t$, $t \in \mathbb{R}$, leads to exponential distributions.

Suppose we have k independent samples $X_{i;1:m_i:n_i}^{\mathscr{R}_i}, \ldots, X_{i;m_i:m_i:n_i}^{\mathscr{R}_i}$ of progressively Type-II censored order statistics with censoring schemes $\mathscr{R}_i$, baseline cumulative distribution function as in (12.10) with parameters (μ_i, ϑ_i), $1 \le i \le k$. Then, from the independence assumption and (12.11), the (unrestricted) MLEs of $(\mu_i, \vartheta_i)_{1 \le i \le k}$ are given by

$$\widehat{\mu}_i = d(X_{i;1:m_i:n_i}^{\mathscr{R}_i}), \quad \widehat{\vartheta}_i = \frac{1}{m} \sum_{j=2}^{m} \gamma_j(\mathscr{R}_i) \Big[d(X_{i;j:m_i:n_i}^{\mathscr{R}_i}) - d(X_{i;j-1:m_i:n_i}^{\mathscr{R}_i}) \Big], \quad 1 \le i \le k. \tag{12.52}$$

Then, adapting the result of Beutner and Kamps [195] to the setting of progressively Type-II censored order statistics, we get the following result. It follows from the representation in (12.52) and, as pointed out by Beutner and Kamps [195], from the MLEs under order restrictions of exponential parameters as presented in Barlow et al. [170, pp. 45].

Theorem 12.11.1. Subject to the constraint $\vartheta_1 \leq \cdots \leq \vartheta_k$, the MLEs of $(\mu_i, \vartheta_i)_{1\leq i\leq k}$ are given by

$$\widetilde{\vartheta}_i = \max_{1\leq \nu\leq i} \min_{i\leq \ell\leq k} \frac{\sum_{j=\nu}^{\ell} m_j \widehat{\vartheta}_j}{\sum_{j=\nu}^{\ell} m_j},$$
$$\widehat{\mu}_i = d(X_{i;1:m_i:n_i}^{\mathscr{R}_i}), \quad 1\leq i\leq k.$$

For general progressive censoring with samples $X_{i;r_i+2:m_i:n_i}^{\mathscr{R}_i}, \ldots, X_{i;m_i:m_i:n_i}^{\mathscr{R}_i}$, $r_i + 2 \leq m_i$, $1 \leq i \leq k$, the MLEs result from Remark 12.1.9 as

$$\widehat{\mu}_i = d(X_{i;r_i+2:m_i:n_i}^{\mathscr{R}_i}) + \widehat{\vartheta}_i \log\Big(1 - \frac{r_i}{m_i}\Big),$$
$$\widehat{\vartheta}_i = \frac{1}{m}\sum_{j=r_i+2}^{m} \gamma_j(\mathscr{R}_i)\Big[d(X_{i;j:m_i:n_i}^{\mathscr{R}_i}) - d(X_{i;j-1:m_i:n_i}^{\mathscr{R}_i})\Big], \quad 1\leq i\leq k. \tag{12.53}$$

As in Theorem 12.11.1, the MLEs of $(\vartheta_i)_{1\leq i\leq k}$ under order restriction $\vartheta_1 \leq \cdots \leq \vartheta_k$ are now given by

$$\widetilde{\vartheta}_i = \max_{1\leq \nu\leq i} \min_{i\leq \ell\leq k} \frac{\sum_{j=\nu}^{\ell} (m_j - r_j)\widehat{\vartheta}_j}{\sum_{j=\nu}^{\ell} (m_j - r_j)}, \quad 1\leq i\leq k.$$

The MLEs $\widetilde{\mu}_i$ of the location parameters μ_i result by replacing $\widehat{\vartheta}_i$ by $\widetilde{\vartheta}_i$ in (12.53), $1 \leq i \leq k$. Finally, Beutner and Kamps [195] stated that the MLEs are unique because they result as a solution of a generalized isotonic regression problem (see Robertson et al. [755]).

Furthermore, denoting by $\mathsf{L}(\boldsymbol{\mu}_k, \boldsymbol{\vartheta}_k)$ the likelihood function, a likelihood ratio test

$$\mathsf{T} = -\log \frac{\sup_{\vartheta_1=\cdots=\vartheta_k} \mathsf{L}(\boldsymbol{\mu}_k, \boldsymbol{\vartheta}_k)}{\sup_{\vartheta_1\leq\cdots\leq\vartheta_k} \mathsf{L}(\boldsymbol{\mu}_k, \boldsymbol{\vartheta}_k)}$$

has been proposed to test the hypothesis

$$H_0 : \vartheta_1 = \cdots = \vartheta_k$$

versus the alternative

$$H_1 : \vartheta_1 \leq \cdots \leq \vartheta_k \quad \text{and } \vartheta_i < \vartheta_{i+1} \text{ for some } i \in \{1, \ldots, k-1\}.$$

It turns out that, under some regularity conditions, the asymptotic distribution of T (for $m_i - r_i \to \infty$, $1 \le i \le k$) is a mixture of χ^2-distributions. For details, we refer to Bhattacharya [198] and Beutner and Kamps [195]. Furthermore, Bhattacharya [198] illustrated the approach by simulation results. The method is applied to survival data of patients with squamous carcinoma of the oropharynx taken from Kalbfleisch and Prentice [490].

Chapter 13
Point Estimation in Progressive Type-I Censoring

Progressive Type-I censoring poses some problems in developing both exact distribution theory and exact inferential procedures. This is mainly due to the random nature of the failures occurring within each time interval $[T_{\ell-1}, T_\ell]$, which allow for both a termination of the life test before the final censoring time T_k and no observation. Nevertheless, it has been noticed by many authors that the likelihood function deduced from the density function given in (4.7) yields explicit representations for the maximum likelihood estimates conditionally on the event that at least one failure has been observed. Despite this, exact inferential analysis of the estimators is available only in the case of an exponential distribution (see Balakrishnan et al. [150]). Thus, most of the available inferential analysis for progressively Type-I censored data is approximation based and numerical in nature. Inferential results for various distributions have been obtained by Cohen [267, 269, 270, 271, 272], Ringer and Sprinkle [754], Wingo [898, 901], Cohen and Norgaard [274], Nelson [676], Gibbons and Vance [403], Cohen and Whitten [277], Balakrishnan and Cohen [92], and Wong [902]. We shall discuss maximum likelihood inference for several distributions important in lifetime modeling. The inference is carried out, given that $D_{\bullet k} \geq 1$, ensuring that one failure has been observed in the progressively censored experiment. We consider a family of absolutely continuous lifetime distributions given by a cumulative distribution function $F_{\boldsymbol{\theta}}$ with density function $f_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^p$, for some $p \in \mathbb{N}$. From (4.7) and (4.8), we get the likelihood function

$$\mathsf{L}(\boldsymbol{\theta} | \mathbf{x}, \mathbf{d}) = C_I \prod_{i=1}^{m} f_{\boldsymbol{\theta}}(x_i) \prod_{i=1}^{k} [1 - F_{\boldsymbol{\theta}}(T_i)]^{R_i}, \tag{13.1}$$

where $\mathbf{x} = (x_1, \ldots, x_m)$, $\mathbf{d} = (d_1, \ldots, d_k) \in \mathfrak{D}$ with $m = d_{\bullet k} \geq 1$, and $\mathscr{R} = (R_1, \ldots, R_k)$ is the effectively applied censoring scheme.

Notice that the likelihood function in (13.1) has the same structure as the likelihood function (12.1) in the case of progressive Type-II censoring. The expressions

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_13,

differ only by the normalizing constant (which does not depend on the parameter and, thus, is not relevant for the maximization) and the argument of the cumulative distribution function $F_{\boldsymbol{\theta}}$. Here, the random censoring times x_i are replaced by the prefixed censoring times T_i. Therefore, the resulting expressions for the likelihood equations and, thus, for the MLEs are quite similar. However, the distributions of the resulting estimators are different and much more complicated in the case of progressive Type-I censoring.

13.1 Exponential Distribution

Maximum likelihood estimation for exponential progressively Type-I censored order statistics is discussed in Cohen [267], Nelson [676, pp. 317], Cohen [272, pp. 128], Balakrishnan and Aggarwala [86, pp. 122], Balakrishnan et al. [150], and Cramer and Tamm [306]. Subsequently, we present results for both scale and location–scale models. Throughout, we assume that $T_1 < \cdots < T_k$ and that the initially planned censoring scheme is given by $\mathscr{R}^0 = (R_1^0, \ldots, R_{k-1}^0)$ with $n > R_{\bullet k-1}^0$. The notation is as introduced in Chap. 4.

13.1.1 One-Parameter Exponential Distribution

Suppose the underlying random variables are distributed according to an $\mathsf{Exp}(\vartheta)$-distribution, $\vartheta > 0$. Then, the likelihood function reads

$$\mathsf{L}(\vartheta|\mathbf{x}, \mathbf{d}) = \frac{C_I}{\vartheta^m} \exp\Big\{-\frac{1}{\vartheta}\sum_{j=1}^{k}\Big[\sum_{i=d_{\bullet j-1}+1}^{d_{\bullet j}} x_i + R_j T_j\Big]\Big\}, \tag{13.2}$$

where $d_{\bullet 0} = 0$. The log-likelihood function is given by

$$\ell(\vartheta|\mathbf{x}, \mathbf{d}) = \text{const} - m\log\vartheta - \frac{1}{\vartheta}\sum_{j=1}^{k}\Big[\sum_{i=d_{\bullet j-1}+1}^{d_{\bullet j}} x_i + R_j T_j\Big] \tag{13.3}$$

showing that the MLE of ϑ does not exist for $m = d_{\bullet k} = 0$. In this case, ℓ is a strictly increasing function. For $m \geq 1$, a maximization w.r.t. $\vartheta > 0$ yields the maximum likelihood estimator

$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m}\sum_{j=1}^{k}\Big[\sum_{i=d_{\bullet j-1}+1}^{d_{\bullet j}} x_i + R_j T_j\Big] = \frac{1}{m}\sum_{j=1}^{k} U_j, \tag{13.4}$$

where $U_j = \sum_{i=d_{\bullet j-1}+1}^{d_{\bullet j}}(x_i - T_{j-1}) + (n - d_{\bullet j} - R_{\bullet j-1})\Delta_j$, $\Delta_j = T_j - T_{j-1}$, is the *total time on test* in the jth interval, $1 \leq j \leq k$, $T_0 = 0$. Notice that the effectively applied censoring plan $\mathscr{R} = (R_1, \ldots, R_k)$, and not the initially planned censoring scheme $\mathscr{R}^0$, is used in the estimate. The particular case of Type-I censoring ($k = 1$)

has been considered by Epstein [353]. Cohen [267] established an estimate for the asymptotic variance of the MLE via the observed Fisher information:

$$\widehat{\mathsf{Var}}(\widehat{\vartheta}_{\mathsf{MLE}}) = \left(-\frac{\partial^2 \ell(\vartheta)}{\partial \vartheta^2}\bigg|_{\vartheta = \widehat{\vartheta}_{\mathsf{MLE}}} \right)^{-1}.$$

A simple calculation shows that $\widehat{\mathsf{Var}}(\widehat{\vartheta}_{\mathsf{MLE}}) = \widehat{\vartheta}^2_{\mathsf{MLE}}/m$.

Example 13.1.1. Based on the failure data of diesel engine fans given in Data B.2.1, the question of an 8,000-hour warranty is addressed (see also Cohen [272]). We are interested in the estimation of the mean time to failure. The data possess $m = 12$ observed failure times and 58 units censored at $k = 27$ censoring times. Assuming an Exp(ϑ)-distribution, the MLE of ϑ has the value

$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{344\,440}{12} = 28\,703.3 \text{ hours}.$$

Then, the maximum likelihood estimate of the cumulative distribution function at $x = 8\,000$ is given by

$$\widehat{P}_{\widehat{\vartheta}_{\mathsf{MLE}}}(X \le 8\,000) = 0.2787,$$

showing that we can expect about 27.9 % of the fans to fail within the warranty period.

Using the decomposition (13.4), Balakrishnan et al. [150] established the conditional density function of $\widehat{\vartheta}_{\mathsf{MLE}}$, given $D_{\bullet k} \ge 1$.

Theorem 13.1.2. The conditional density function of $\widehat{\vartheta}_{\mathsf{MLE}}$, given $D_{\bullet k} \ge 1$, is given by

$$f^{\widehat{\vartheta}_{\mathsf{MLE}} | D_{\bullet k} \ge 1}(x) = \sum_{m=1}^{n} \sum_{\mathbf{d} \in \mathfrak{D}_m} \sum_{\nu_1=0}^{d_1} \cdots \sum_{\nu_k=0}^{d_k} C^{[\vartheta]}_{\mathbf{d},\boldsymbol{\nu}} \cdot \gamma\left(x - \tau_{\mathbf{d},\boldsymbol{\nu}}; m, \frac{m}{\vartheta}\right), x \in \mathbb{R},$$

where

$$\tau_{\mathbf{d},\boldsymbol{\nu}} = \frac{1}{m} \sum_{i=1}^{k} (n - d_{\bullet i} - R_{\bullet i-1} + \nu_i)\, \Delta_i,$$

$$C^{[\vartheta]}_{\mathbf{d},\boldsymbol{\nu}} = \frac{(-1)^{\sum_{i=1}^k \nu_i}}{1 - \exp\left(-\frac{\sum_{i=1}^k R_i T_i}{\vartheta}\right)} \times \left\{ \prod_{i=1}^{k} \binom{n - d_{\bullet i-1} - R_{\bullet i-1}}{d_i} \binom{d_i}{\nu_i} \right\} \exp\left\{-\frac{m}{\vartheta} \tau_{\mathbf{d},\boldsymbol{\nu}}\right\},$$

$$\mathfrak{D}_m = \{\mathbf{d} \in \mathfrak{D} : d_{\bullet k} = m\},$$

and $\gamma(\cdot; \alpha, \lambda)$ denotes the density function of a gamma distribution $\Gamma(\alpha, \lambda)$ (see (A.1.7)).

Theorem 13.1.2 reveals that the conditional distribution of $\widehat{\vartheta}_{\mathsf{MLE}}$, given $D_{\bullet k} \geq 1$, is a generalized mixture of gamma distributions. This allows for explicit expressions for conditional moments.

Corollary 13.1.3. The conditional expectation and conditional variance of $\widehat{\vartheta}_{\mathsf{MLE}}$ are given by

$$E(\widehat{\vartheta}_{\mathsf{MLE}} | D_{\bullet k} \geq 1) = \vartheta + b(\widehat{\vartheta}_{\mathsf{MLE}})$$

and

$$\mathsf{Var}(\widehat{\vartheta}_{\mathsf{MLE}} | D_{\bullet k} \geq 1) = \mathrm{MSE}(\widehat{\vartheta}_{\mathsf{MLE}}) - b^2(\widehat{\vartheta}_{\mathsf{MLE}}),$$

where

$$b(\widehat{\vartheta}_{\mathsf{MLE}}) = \sum_{m=1}^{n} \sum_{\mathbf{d}\in\mathfrak{D}_m} \sum_{\nu_1=0}^{d_1} \cdots \sum_{\nu_k=0}^{d_k} C_{\mathbf{d},\boldsymbol{\nu}}^{[\vartheta]} \tau_{\mathbf{d},\boldsymbol{\nu}},$$

$$\mathsf{MSE}(\widehat{\vartheta}_{\mathsf{MLE}}) = \sum_{m=1}^{n} \sum_{\mathbf{d}\in\mathfrak{D}_m} \sum_{\nu_1=0}^{d_1} \cdots \sum_{\nu_k=0}^{d_k} C_{\mathbf{d},\boldsymbol{\nu}}^{[\vartheta]} \left(\frac{\vartheta^2}{m} + \tau_{\mathbf{d},\boldsymbol{\nu}}^2\right)$$

are bias and mean squared error.

From the generalized mixture representation, an expression for the conditional survival function $P_\vartheta(\widehat{\vartheta}_{\mathsf{MLE}} > x | D_{\bullet k} \geq 1)$ can be directly obtained as

$$P_\vartheta(\widehat{\vartheta}_{\mathsf{MLE}} > x | D_{\bullet k} \geq 1) = \sum_{m=1}^{n} \sum_{\mathbf{d}\in\mathfrak{D}_m} \sum_{\nu_1=0}^{d_1} \cdots \sum_{\nu_k=0}^{d_k} \left\{ C_{\mathbf{d},\boldsymbol{\nu}}^{[\vartheta]} \cdot \mathrm{e}^{-\frac{m}{\beta}(x-\tau_{\mathbf{d},\boldsymbol{\nu}})} \cdot \sum_{i=1}^{m} \frac{\left[\frac{m}{\vartheta}(x-\tau_{\mathbf{d},\boldsymbol{\nu}})\right]^{i-1}}{(i-1)!} \right\}$$

for $x > \tau_{\mathbf{d},\boldsymbol{\nu}}$. Notice that the occurring gamma distributions are Erlang distributions and, thus, their survival function has an explicit sum representation. It will be used to construct conditional confidence intervals. The method is applied to Data B.2.2 (see Balakrishnan et al. [150, pp. 349]).

Example 13.1.4. For the data given in Table B.2, the MLE of ϑ is given by $\widehat{\vartheta}_{\mathsf{MLE}} = 10.861$. Using this result, the estimated values of the bias, standard error, and mean squared error of ϑ are obtained as

$$\widehat{b}(\widehat{\vartheta}_{\mathsf{MLE}}) = 0.946, \widehat{\mathsf{Var}}(\widehat{\vartheta}_{\mathsf{MLE}}) = 24.108, \text{ and } \widehat{\mathsf{MSE}}(\widehat{\vartheta}_{\mathsf{MLE}}) = 25.004,$$

respectively. The estimated asymptotic variance is given by 14.746. For the data given in Table B.3, we get the estimates $\widehat{\vartheta}_{\mathsf{MLE}} = 11.497$,

$$\widehat{b}(\widehat{\vartheta}_{\mathsf{MLE}}) = 0.633, \widehat{\mathsf{Var}}(\widehat{\vartheta}_{\mathsf{MLE}}) = 20.027, \text{ and } \widehat{\mathsf{MSE}}(\widehat{\vartheta}_{\mathsf{MLE}}) = 19.626.$$

The estimated asymptotic variance is given by 14.686. Notice that the estimated asymptotic variance is considerably smaller in both cases so that it underestimates the variation.

13.1.2 Two-Parameter Exponential Distribution

Now, we turn to the two-parameter case $\mathsf{Exp}(\mu, \vartheta)$ which is considered, e.g., in Cohen [272, pp. 128] and Cohen and Whitten [277, pp. 110]. It has to be noted that the maximum likelihood estimator of ϑ presented in Cramer and Tamm [306] differs from their expression in some cases. It was overlooked in the earlier papers that the first failure may occur after the first censoring time T_1. Therefore, the first observation may be larger than T_1 which leads to a change in the likelihood function. We assume that μ and ϑ are unknown and have to be estimated. In the two-parameter exponential distribution, the location parameter causes the support to depend on μ (see Definition A.1.5) so that the corresponding survival function is given by

$$1 - F_{\mu,\vartheta}(T_i) = \begin{cases} \exp\left(-\frac{T_i-\mu}{\vartheta}\right), & \mu \le T_i \\ 1, & \mu > T_i \end{cases} = \exp\left(-\frac{[T_i - \mu]_+}{\vartheta}\right).$$

Hence, the likelihood function is given by

$$\mathsf{L}(\mu, \vartheta) = \frac{C_I}{\vartheta^m} \exp\left(-\sum_{i=1}^{m} \frac{x_i - \mu}{\vartheta} - \sum_{i=1}^{k} \frac{R_i[T_i - \mu]_+}{\vartheta}\right) \mathbb{1}_{(-\infty, x_1]}(\mu).$$

Conditional on $m \ge 1$, the maximum likelihood estimates need not exist in any case. For instance, given $m = 1$, it is possible that either the likelihood function is unbounded or unique maximum likelihood estimators for μ and ϑ exist. For details, we refer to Cramer and Tamm [306]. Assuming $m \ge 2$ and that $x_2 > x_1$, both MLEs exist. Then, for any $\vartheta > 0$, the term

$$\sum_{i=1}^{m} \frac{x_i - \mu}{\vartheta} + \sum_{i=1}^{k} \frac{R_i[T_i - \mu]_+}{\vartheta}$$

is increasing in $\mu \le x_1$. Therefore, the upper bound

$$\mathsf{L}(\vartheta) = \frac{C_I}{\vartheta^m} \exp\left(-\frac{1}{\vartheta}\left\{\sum_{i=2}^{m}(x_i - x_1) + \sum_{i=1}^{k} R_i[T_i - x_1]_+\right\}\right)$$

of $\mathsf{L}(\mu, \vartheta)$ results which is attained for $\mu = x_1$ only. Hence, the MLEs of μ and ϑ are given by

$$\widehat{\mu}_{\mathsf{MLE}} = X_{1:M:n}^{\mathscr{R},T},$$

$$\widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{M}\sum_{i=2}^{M}(X_{i:M:n}^{\mathscr{R},T} - X_{1:M:n}^{\mathscr{R},T}) + \frac{1}{M}\sum_{i=1}^{k} R_i[T_i - X_{1:M:n}^{\mathscr{R},T}]_+.$$

The MLE for ϑ differs from that presented by Cohen [272] and Cohen and Whitten [277]

$$\widehat{\vartheta}^* = \frac{1}{M}\sum_{i=2}^{M}(X_{i:M:n}^{\mathscr{R},T} - X_{1:M:n}^{\mathscr{R},T}) + \frac{1}{M}\sum_{i=1}^{k} R_i(T_i - X_{1:M:n}^{\mathscr{R},T})$$

when the first observation exceeds the first censoring time T_1. Cramer and Tamm [306] illustrated this effect by a data set presented in Wingo [901], where the first censoring time precedes the first observation (see Data B.2.4).

Example 13.1.5. Using Data B.2.4, we get the location estimate $\widehat{\mu}_{\mathsf{MLE}} = 0.529$. The corresponding estimates for the scale parameter are $\widehat{\vartheta}_{\mathsf{MLE}} = 0.333$ and $\widehat{\vartheta}^* = 0.262$. The latter estimate leads to a smaller value since it incorporates the censoring times T_1 and T_2. The estimated pain relief times are 0.862 and 0.791, respectively. Notice that Wingo [901] fitted a Burr-XII model to this data. He obtained an estimated mean pain relief time of 0.836 which is quite close to the maximum likelihood estimate based on the two-parameter exponential distribution.

Remark 13.1.6. The MLE $\widehat{\vartheta}_{\mathsf{MLE}}$ includes only these censoring times which exceed the first observed failure time. This is quite natural in the sense that the left endpoint of the support is estimated by $X_{1:M:n}^{\mathscr{R},T}$ and, thus, values less than $X_{1:M:n}^{\mathscr{R},T}$ do not contain valuable information.

Similar problems arise for other distributions having a finite left endpoint of support. This includes, e.g., three-parameter Weibull distributions (see Cohen [270]), three-parameter log-normal distributions (see Cohen [271]), and three-parameter gamma distributions (see Cohen and Norgaard [274]). In these studies, the problem has not been taken into account, too.

Cramer and Tamm [306] mentioned that the MLEs of both the location and scale parameter are biased which has also been frequently observed in similar settings (see Cohen [273]). They illustrated this observation by a simulation study and proposed the bias-adjusted MLEs

$$\widehat{\vartheta}_{\text{adj}} = \frac{1}{M-1}\sum_{i=2}^{M}(X_{i:M:n}^{\mathscr{R},T} - X_{1:M:n}^{\mathscr{R},T}) + \frac{1}{M}\sum_{j=1}^{k} R_j[T_j - X_{1:M:n}^{\mathscr{R},T}]_+,$$

$$\widehat{\mu}_{\text{adj}} = X_{1:M:n}^{\mathscr{R},T} - \frac{1}{n}\widehat{\vartheta}_{\text{adj}},$$

which seem to provide better estimates especially in small samples. More details as well as details on density estimates can be found in Cramer and Tamm [306].

13.1.3 Modified Moment Estimation

Cohen [272, pp. 130] proposed modified maximum likelihood estimators for the $\mathsf{Exp}(\mu, \vartheta)$-model. He introduced an additional equation

$$E(X_{1:n}) = x_1 \quad \text{or, equivalently,} \quad \mu = x_1 - \frac{\vartheta}{n}$$

which obviously makes sense only when the minimum of the data is included in the progressively Type-I censored sample. Therefore, we assume for this section that the first observation is given by the minimum and that $x_1 < T_1$. Replacing μ by $x_1 - \frac{\vartheta}{n}$ in the likelihood equation $\frac{\partial}{\partial\vartheta}\ell(\mu, \vartheta) = 0$ leads to the estimators

$$\widehat{\mu}_{\mathsf{MMLE}} = X_{1:M:n}^{\mathscr{R},T} - \frac{\widehat{\vartheta}_{\mathsf{MMLE}}}{n},$$

$$\widehat{\vartheta}_{\mathsf{MMLE}} = \frac{1}{M-1}\sum_{i=1}^{M}(X_{i:M:n}^{\mathscr{R},T} - X_{1:M:n}^{\mathscr{R},T}) + \frac{1}{M-1}\sum_{i=1}^{k} R_i(T_i - X_{1:M:n}^{\mathscr{R},T}).$$

Remark 13.1.7. Notice that $P(X_{1:n} > T_1) = \exp\{-\frac{n[T_1-\mu]_+}{\vartheta}\} > 0$ for any $\mu \in \mathbb{R}$ and $\vartheta > 0$. In particular, this probability is one for $\mu > T_1$. Moreover, the minimum of the sample may be progressively censored with positive probability so that it is not observed. Suppose that $X_{1:n} > T_1$ and $R_1^0 > 0$. Then, the probability of censoring $X_{1:n}$ at T_1 is given by $\frac{R_1^0}{n}$. Therefore, we find that the probability not to observe $X_{1:n}$ is positive:

$$\begin{aligned} P(X_{1:n} \text{ not observed}) &= P(X_{1:n} \text{ censored} | X_{1:n} > T_1) P(X_{1:n} > T_1) \\ &\geq \frac{R_1^0}{n} \exp\Big\{-\frac{n[T_1-\mu]_+}{\vartheta}\Big\} > 0. \end{aligned}$$

On the other hand, assuming that $T_1 > \mu$, we get $\lim_{n\to\infty} P(X_{1:n} > T_1) = 0$. For large samples and an appropriately chosen first censoring time T_1, we can proceed on the assumption that the minimum is observed before the first censoring time. However, for small samples which are commonly encountered in reliability, the problem is still present and must not be neglected.

Example 13.1.8. Data B.2.1 is considered with an $\mathsf{Exp}(\mu, \vartheta)$-model (see also Example 13.1.1). Cohen [272, pp. 136] applied the modified maximum likelihood estimates to obtain

$$\widehat{\mu}_{\mathsf{MMLE}} = 43.6 \quad \text{and} \quad \widehat{\vartheta}_{\mathsf{MMLE}} = 28\,449.$$

Now, we get an estimate for a failure within the 8,000-hour warranty period of $\widehat{P}_{\mathsf{MMLE}} = 0.244$. Using the MLEs, we get

$$\widehat{\mu}_{\mathsf{MLE}} = 450, \quad \widehat{\vartheta}_{\mathsf{MLE}} = 26\,078, \quad \text{and} \quad \widehat{P}_{\mathsf{MLE}} = 0.251.$$

13.2 Weibull Distributions

Maximum likelihood estimation for two-parameter Weibull distributions with progressively Type-I censored data has been considered in Cohen [268, 269], Gibbons and Vance [403], Cohen [272, pp. 88], and Balakrishnan and Aggarwala [86, pp. 125]. Explicit expressions for the MLEs are not available and the estimates have to be computed by numerical procedures. Balakrishnan and Kateri [105] have established the existence and uniqueness of the MLEs. Three-parameter Weibull distributions are considered in Cohen [270], Wingo [898], and Lemon [583].

We consider the $\mathsf{Weibull}(\vartheta, \beta)$-distribution as defined in Definition A.1.6. For a progressively Type-I censored sample $X_{1:M:n}^{\mathscr{R},T}, \ldots, X_{M:M:n}^{\mathscr{R},T}$ with observations $x_1, \ldots, x_m$, the log-likelihood function is given by

$$\ell(\vartheta, \beta) = \log C_I + m \log \beta - m \log \vartheta + (\beta - 1) \sum_{i=1}^{m} \log x_i - \frac{1}{\vartheta} \sum_{i=1}^{m} x_i^{\beta} - \frac{1}{\vartheta} \sum_{i=1}^{k} R_i T_i^{\beta}.$$

Differentiating ℓ w.r.t. ϑ and β results in the likelihood equations

$$\frac{\partial}{\partial \beta} \ell(\vartheta, \beta) = \frac{m}{\beta} + \sum_{i=1}^{m} \log x_i - \frac{1}{\vartheta} \sum_{i=1}^{m} x_i^{\beta} \log x_i - \frac{1}{\vartheta} \sum_{i=1}^{k} R_i T_i^{\beta} \log T_i = 0, \tag{13.5a}$$

$$\frac{\partial}{\partial \vartheta} \ell(\vartheta, \beta) = -\frac{m}{\vartheta} + \frac{1}{\vartheta^2} \sum_{i=1}^{m} x_i^{\beta} + \frac{1}{\vartheta^2} \sum_{i=1}^{k} R_i T_i^{\beta} = 0. \tag{13.5b}$$

A rearrangement of (13.5b) leads to the expression

$$\widehat{\vartheta} = \frac{1}{m} \Big(\sum_{i=1}^{m} x_i^{\beta} + \sum_{i=1}^{k} R_i T_i^{\beta} \Big).$$

Then, equation (13.5a) can be rewritten as

$$\frac{1}{m} \sum_{i=1}^{m} \log x_i = \frac{\sum_{i=1}^{m} x_i^{\beta} \log x_i + \sum_{i=1}^{k} R_i T_i^{\beta} \log T_i}{\sum_{i=1}^{m} x_i^{\beta} + \sum_{i=1}^{k} R_i T_i^{\beta}} - \frac{1}{\beta}. \tag{13.6}$$

Equation (13.6) has to be solved numerically. Therefore, it is important to know whether it has a (unique) solution. This result was established by Balakrishnan and Kateri [105]. This shows that (13.6) can be solved by some numerical procedure leading to the desired MLEs. As mentioned before, equation (13.6) is very similar to the corresponding equation for progressive Type-II censoring given in (12.13).

The asymptotic variance–covariance matrix can be estimated by the inverse of the observed Fisher information matrix as

$$\begin{pmatrix} \widehat{\mathsf{Var}}(\widehat{\beta}) & \widehat{\mathsf{Cov}}(\widehat{\beta},\widehat{\vartheta}) \\ \widehat{\mathsf{Cov}}(\widehat{\vartheta},\widehat{\beta}) & \widehat{\mathsf{Var}}(\widehat{\vartheta}) \end{pmatrix} = \begin{pmatrix} -\dfrac{\partial^2 \ell(\vartheta,\beta)}{\partial\beta^2}\Big|_{\widehat{\beta},\widehat{\vartheta}} & -\dfrac{\partial^2 \ell(\vartheta,\beta)}{\partial\beta\partial\vartheta}\Big|_{\widehat{\beta},\widehat{\vartheta}} \\ -\dfrac{\partial^2 \ell(\vartheta,\beta)}{\partial\vartheta\partial\beta}\Big|_{\widehat{\beta},\widehat{\vartheta}} & -\dfrac{\partial^2 \ell(\vartheta,\beta)}{\partial\vartheta^2}\Big|_{\widehat{\beta},\widehat{\vartheta}} \end{pmatrix}^{-1}. \tag{13.7}$$

The required second partial derivatives are given by

$$\frac{\partial^2}{\partial\beta^2}\ell(\vartheta,\beta) = -\frac{m}{\beta^2} - \frac{1}{\vartheta}\left(\sum_{i=1}^m x_i^\beta(\log x_i)^2 + \sum_{i=1}^k R_i T_i^\beta (\log T_i)^2\right),$$

$$\frac{\partial^2}{\partial\beta\partial\vartheta}\ell(\vartheta,\beta) = \frac{1}{\vartheta^2}\left(\sum_{i=1}^m x_i^\beta \log x_i + \sum_{i=1}^k R_i T_i^\beta \log T_i\right),$$

$$\frac{\partial^2}{\partial\vartheta^2}\ell(\vartheta,\beta) = \frac{m}{\vartheta^2} - \frac{2}{\vartheta^3}\left(\sum_{i=1}^m x_i^\beta + \sum_{i=1}^k R_i T_i^\beta\right).$$

Remark 13.2.1. Gibbons and Vance [403] have conducted a simulation study which compares the MLEs with estimators resulting from a graphical procedure (the so-called least-squares median rank estimators (LSMRE)) described in Johnson [481]. The mean squared errors were computed for both procedures and for several sample sizes and proportions of censored units.

Example 13.2.2. We illustrate the approach by Montanari and Cacciari's [655] XLPE-isolated cable data (see Data B.2.5) who assumed Weibull lifetimes to analyze the data. Notice that the data have been discussed in Montanari and Cacciari [655] for a different parametrization of the Weibull distribution. Solving the likelihood equations yields the estimates

$$\widehat{\beta}_{\mathsf{MLE}} = 3.077, \quad \widehat{\vartheta}_{\mathsf{MLE}} = 1.802697652 \cdot 10^9.$$

The α-quantile of the two-parameter Weibull distribution is given by $\xi_\alpha = [-\vartheta\log(1-\alpha)]^{1/\beta}$, so that the estimated quantile is given by $\widehat{\xi}_{0.5} = 903$. This means that we have to expect 50 % of the failures within the first 903 h of operation.

Remark 13.2.3. Cohen [270], Wingo [898], and Lemon [583] considered three-parameter Weibull distributions with a left endpoint of support μ. As mentioned above for the two-parameter exponential distribution with location parameter μ, the authors have not taken into account that the minimum may not be observed. Therefore, their formulas also need to be corrected in this case.

13.3 Extreme Value Distributions

Cohen [272, pp. 144] considered progressively Type-I censored order statistics from an extreme value distribution with location and scale parameters μ and ϑ, respectively. Since these data can be transformed by an exponential function, i.e., $Y_{j:M:n}^{\mathscr{R},T} = \exp(X_{j:M:n}^{\mathscr{R},T})$, to get progressively Type-I censored order statistics from a Weibull distribution, the results for the Weibull case can be directly applied. The MLEs are given by the solutions of the equations

$$\mu = \vartheta \log\left[\frac{1}{m}\left(\sum_{i=1}^{m} \mathrm{e}^{x_i/\vartheta} + \sum_{i=1}^{k} R_i \mathrm{e}^{T_i/\vartheta}\right)\right],$$

$$\frac{1}{m}\sum_{i=1}^{m} x_i = \frac{\sum_{i=1}^{m} \mathrm{e}^{x_i/\vartheta} x_i + \sum_{i=1}^{k} R_i \mathrm{e}^{T_i/\vartheta} T_i}{\sum_{i=1}^{m} \mathrm{e}^{x_i/\vartheta} + \sum_{i=1}^{k} R_i \mathrm{e}^{T_i/\vartheta}} - \vartheta.$$

Notice that the first equation is explicit in μ and the second equation has a unique solution in ϑ.

Example 13.3.1. The logarithmic failure times of Nelson's progressively Type-I censored insulating fluid data 1.1.9 are used to illustrate the preceding approach (see Table 13.1). The effectively applied censoring scheme is given by $\mathscr{R} = (2^{*3})$. The transformed censoring times are $T_1 = 1.09861$, $T_2 = 2.19722$, and $T_3 = 2.89037$. The resulting estimates are given by $\widehat{\mu} = 2.20241, \widehat{\vartheta} = 1.07550$. The estimated mean lifetime is given by $\mathrm{e}^{\widehat{\mu}} \cdot \Gamma(1 + \widehat{\vartheta}) = 9.35717$.

13.4 Normal Distribution

For a location–scale model from a normal distribution, Cohen [267] derived maximum likelihood estimators (see also Cohen [272], Nelson [676], and Balakrishnan and Aggarwala [86]). Denoting by φ and Φ the density function and the cumulative distribution function of the standard normal distribution, respectively, the resulting partial derivatives of the log-likelihood function are

−1.66073	−0.24846	−0.04082	0.27003	1.02245	1.15057	1.42311
1.54116	1.57898	1.87180	1.99470	2.11263	2.48989	

Table 13.1 Logarithmic values of Nelson's progressively Type-I censored insulating fluid data 1.1.9

$$\frac{\partial}{\partial\mu}\ell(\mu,\vartheta) = \frac{m}{\vartheta}\left(\frac{\overline{x}-\mu}{\vartheta} + \sum_{i=1}^{k}\frac{R_i}{m}h_i\right),$$

$$\frac{\partial}{\partial\vartheta}\ell(\mu,\vartheta) = \frac{m}{\vartheta}\left(\frac{s^2+(\overline{x}-\mu)^2}{\vartheta^2} - 1 + \sum_{i=1}^{k}\frac{\xi_i R_i}{m}h_i\right),$$

where $\xi_i = \frac{T_i-\mu}{\vartheta}$ and $h_i = h(\xi_i) = \frac{\varphi(\xi_i)}{1-\Phi(\xi_i)}$, $1 \le i \le k$. The corresponding likelihood equations are given by

$$\overline{x} = \mu - \vartheta\sum_{i=1}^{k}\frac{R_i}{m}h_i, \tag{13.8a}$$

$$s^2 = \vartheta^2\left[1 - \frac{1}{m}\sum_{i=1}^{k}\xi_i R_i h_i - \left(\sum_{i=1}^{k}\frac{R_i}{m}h_i\right)^2\right], \tag{13.8b}$$

where $\overline{x} = \frac{1}{m}\sum_{i=1}^{m} x_i$ and $s^2 = \frac{1}{m}\sum_{i=1}^{m}(x_i - \overline{x})^2$ denote the sample mean and sample variance, respectively. Equations (13.8a) and (13.8b) can only be solved numerically (e.g., by a Newton–Raphson procedure; see Cohen [267]). Notice that the likelihood equations for progressive Type-II censored data are similar. The censoring times T_i have to be replaced by the corresponding failure times x_i at which progressive censoring takes place.

The asymptotic variance–covariance matrix can be estimated by the inverse observed Fisher information matrix [see (13.7)]. The required partial derivatives of the log-likelihood function are given by

$$\frac{\partial^2}{\partial\mu^2}\ell(\mu,\vartheta) = -\frac{m}{\vartheta^2}\left[1 - \frac{1}{m}\sum_{i=1}^{k}R_i\kappa_i\right],$$

$$\frac{\partial^2}{\partial\mu\partial\vartheta}\ell(\mu,\vartheta) = -\frac{m}{\vartheta^2}\left[\frac{2(\overline{x}-\mu)}{\vartheta} - \frac{1}{m}\sum_{i=1}^{k}R_i\,(\xi_i\kappa_i - h_i)\right],$$

$$\frac{\partial^2}{\partial\vartheta^2}\ell(\mu,\vartheta) = -\frac{m}{\vartheta^2}\left[3\frac{s^2+(\overline{x}-\mu)^2}{\vartheta^2} - 1 + \frac{1}{m}\sum_{i=1}^{k}R_i\left(2h_i\xi_i - \xi_i^2\kappa_i\right)\right],$$

where $\kappa_i = h_i(\xi_i - h_i)$, $1 \le i \le k$.

Log-normal distributions are investigated in Gajjar and Khatri [387].

13.5 Burr-XII Distribution

Burr-XII distributions are sometimes proposed as an alternative to normal, log-normal, gamma, logistic, and exponential distributions since the shapes of these densities are qualitatively depicted by those of Burr-XII density functions. For $\alpha, \beta > 0$, these are given by

$$f_{\alpha,\beta}(t) = \frac{\alpha}{\beta} \frac{t^{\alpha-1}}{(t^\alpha + 1)^{1/\beta+1}}, \quad t > 0. \tag{13.9}$$

The corresponding cumulative distribution function is given by $F_{\alpha,\beta}(t) = 1 - (t^\alpha + 1)^{-1/\beta} \mathbb{1}_{(0,\infty)}(t)$, $t \in \mathbb{R}$. Details on Burr-XII distributions are presented in Tadikamalla [830]. Further comments can be found in, e.g., Wingo [899], Gupta et al. [424], and Jalali and Watkins [476].

Progressively Type-I censored data from Burr-XII distributions has been addressed by Wingo [901]. The resulting log-likelihood function is given by

$$\ell(\alpha,\beta) = \log C_I + m \log \alpha - m \log \beta + (\alpha - 1) \sum_{i=1}^m \log x_i - \sum_{i=1}^m \log(x_i^\alpha + 1) - \frac{\delta(\alpha)}{\beta}, \tag{13.10}$$

where $\delta(\alpha) = \sum_{i=1}^m \log(x_i^\alpha + 1) + \sum_{i=1}^k R_i \log(T_i^\alpha + 1)$.

For given α, the MLE of β is given by $\widehat{\beta} = \frac{1}{m}\delta(\alpha)$. Using the inequality $\log t \le t - 1$, $t > 0$, this can be easily seen from the bound

$$\ell(\alpha,\beta) \le \log C_I - m + m \log \alpha - m \log \widehat{\beta} + (\alpha - 1) \sum_{i=1}^m \log x_i - \sum_{i=1}^m \log(x_i^\alpha + 1) = \ell(\alpha), \tag{13.11}$$

which is valid for any $\beta > 0$. Equality holds iff $\beta = \widehat{\beta}$.

Supposing that β is known, the log-likelihood function given in (13.10) has to be maximized w.r.t. α. Noticing that $\log(x^\alpha + 1)$ is strictly concave in α, we conclude that $\delta(\alpha)$ and $\ell(\alpha, \beta)$ are strictly concave functions in α. This proves that the log-likelihood function has at most one global maximum. Now, it is sufficient to prove that the log-likelihood is bounded from above and the maximum is attained. Obviously, $\lim_{\alpha\to 0} \ell(\alpha,\beta) = -\infty$. Consider the sets

$$\mathcal{X}_> = \{i | x_i > 1,\ i \in \{1,\ldots,m\}\}, \quad \mathcal{X}_< = \{i | x_i < 1, i \in \{1,\ldots,m\}\},$$
$$\mathcal{T}_> = \{i | T_i > 1, R_i > 0, i \in \{1,\ldots,k\}\},$$

and suppose that $\mathcal{X}_> \cup \mathcal{X}_< \cup \mathcal{T}_> \neq \emptyset$. Now, we prove that the log-likelihood function is strictly decreasing for sufficiently large α. With $g_t(\alpha) = \frac{t^\alpha \log t}{t^\alpha+1}$, the partial derivative is given by

$$\frac{\partial}{\partial\alpha}\ell(\alpha,\beta) = \frac{m}{\alpha} + \sum_{i=1}^m \log x_i - \Big(1 + \frac{1}{\beta}\Big)\sum_{i=1}^m g_{x_i}(\alpha) - \sum_{i=1}^m R_i g_{T_i}(\alpha).$$

Since $\lim_{\alpha\to\infty} g_t(\alpha) = \mathbb{1}_{(1,\infty)}(t)$, we arrive at

$$\lim_{\alpha\to\infty} \ell(\alpha,\beta) = \sum_{i\in\mathcal{X}_<} \log x_i - \frac{1}{\beta}\sum_{i\in\mathcal{X}_>} \log x_i - \sum_{i\in\mathcal{T}_>} R_i \log T_i.$$

From the assumption $\mathcal{X}_> \cup \mathcal{X}_< \cup \mathcal{T}_> \neq \emptyset$, we conclude that this limit is negative. Notice that $P(\mathcal{X}_> \cup \mathcal{X}_< = \emptyset) = P(X_{i:M:n}^{\mathscr{R},T} = 1 \text{ for any i}) = 0$ because the Burr-XII distribution is continuous.

In the case of unknown parameters α and β, Wingo [901] has shown that ℓ in (13.11) has a unique global maximum for $\alpha \in (0,\infty)$ provided that $\mathcal{X}_< \neq \emptyset$. The solution $\widehat{\alpha}$ can be derived by solving the equation $\frac{\partial}{\partial\alpha}\ell(\alpha) = 0$, $\alpha > 0$. The estimate for β is given by $\widehat{\beta} = \delta(\widehat{\alpha})$. Details on the proof can also be found in Gupta et al. [424] and Jalali and Watkins [476]. Wang and Cheng [886] proposed an EM-algorithm to compute the estimates. Moreover, they conducted a simulation study comparing the performance of the Newton–Raphson procedure and the EM-algorithm. They recommended to use the EM-algorithm to compute the MLEs since it is better than the Newton–Raphson method in terms of bias and root mean squared error.

Example 13.5.1. For Wingo's pain relief data B.2.4, we get the estimates $\widehat{\alpha} = 6.560$ and $\widehat{\beta} = 2.597$. Using that the expectation of a Burr-XII distribution is given by $\frac{\Gamma(\beta-1/\alpha)\Gamma(1+1/\alpha)}{\Gamma(\beta)}$, an estimate of the mean pain relief time is obtained as 0.836.

13.6 Logistic Distributions

Logistic distributions from a location–scale family with standard member $F(t) = (1 + \mathrm{e}^{-t})^{-1}$, $t \in \mathbb{R}$, are investigated in Gajjar and Khatri [387]. The resulting likelihood equations are closely connected to those in Type-II progressive censoring as presented in Sect. 12.7.5.

Chapter 14
Progressive Hybrid and Adaptive Censoring and Related Inference

Type-I hybrid censoring was originally proposed by Epstein [352], while that of Type-II hybrid censoring was introduced by Childs et al. [259]. Barlow et al. [169] called this sampling procedure *truncated sampling*. According to the constructions illustrated in detail in Sect. 1.1.3, the number of observations is random. In particular, it is possible to have fewer than m observations in case of Type-I hybrid censoring, whereas in the case of Type-II hybrid censoring, we will have at least m observations. Thus, in Type-I hybrid censoring, we are faced with the problem of no observations so that results are normally formulated conditionally on the number D of observed failures.

Type-I hybrid censoring has been used in designing reliability acceptance tests (see MIL-STD-781-C [648]). Many attempts have been made to find the exact distribution of the MLEs in the case of exponential populations. Here, we only sketch the developments in this regard. For a detailed review, we refer to the recent survey paper by Balakrishnan and Kundu [108]. For Type-I hybrid censoring, Chen and Bhattacharyya [249] employed the method of conditional moment generating function to derive the density function of the MLE of ϑ in the scale model. The same approach has been utilized to simplify this expression of the density function by Childs et al. [259]. Childs et al. [260] as well as Kundu and Joarder [561] have utilized the same method for Type-I progressive hybrid censoring (for the two-parameter exponential, see Childs et al. [261]). Cramer and Balakrishnan [292] have used a spacings approach leading to more compact forms of the density functions. Similar results for Type-II (progressive) hybrid censored data have been established in Childs et al. [259], Childs et al. [260], Ganguly et al. [392], and Cramer et al. [315].

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_14,

14.1 Likelihood Inference for Type-I Progressive Hybrid Censored Data

Given $X_1^{(\mathrm{I})}, \ldots, X_m^{(\mathrm{I})}$ with $m \geq 2$ and D as in (2.40), the data $(\mathbf{x}_d, T^{*(m-d)})$ has been observed, where $x_1 \leq \cdots \leq x_d < T$ and $D = d$. Then, with (5.2), it follows that the joint distribution of $X_1^{(\mathrm{I})}, \ldots, X_m^{(\mathrm{I})}$ and D has a density function $f_{\boldsymbol{\theta}}^{(d)}$ w.r.t. the measure $\lambda^d \otimes \bigotimes_{j=1}^{m-d} \varepsilon_T$. This is given by

$$f_{\boldsymbol{\theta}}^{(d)}(\mathbf{x}_d) = C_d (1 - F_{\boldsymbol{\theta}}(T))^{\gamma_{d+1}} f_{1,\ldots,d:d:n-\gamma_{d+1};\boldsymbol{\theta}}^{\mathscr{R}_d}(\mathbf{x}_d), \quad x_1 \leq \cdots \leq x_d < T. \tag{14.1}$$

14.1.1 Likelihood Inference for Two-Parameter Exponential Distributions

For an $\mathsf{Exp}(\mu, \vartheta)$-distribution and a random sample $Z_1^{(\mathrm{I})}, \ldots, Z_m^{(\mathrm{I})}, D$, the likelihood function (14.1) can be written as

$$\mathsf{L}(\mu, \vartheta; \mathbf{z}_d) = \frac{\prod_{j=1}^d \gamma_j}{\vartheta^d} \exp\Big\{ -\frac{1}{\vartheta}\Big[\sum_{j=1}^d (R_j + 1)(z_j - \mu) + \gamma_{d+1}(T - \mu)\Big]\Big\},$$
$$\mu \leq z_1 \leq \cdots \leq z_d \leq T. \tag{14.2}$$

Location Parameter Known

For brevity, let $\mu = 0$. Then, as shown in Childs et al. [260], the MLE of ϑ exists provided $D > 0$ and is given by

$$\widehat{\vartheta} = \frac{1}{D}\Big[\sum_{j=1}^D \gamma_j (Z_j^{(\mathrm{I})} - Z_{j-1}^{(\mathrm{I})}) + \gamma_{d+1} T\Big] = \frac{1}{D} S_D,$$

where S_D is the total time on test statistic. Then, by following the lines of Theorem 5.1.4, we readily obtain the following distributional result for $\widehat{\vartheta}$ in this case:

$$f^{\widehat{\vartheta}|D\geq 1}(t) = \frac{1}{P(D \geq 1)} \sum_{d=1}^m f^{\widehat{\vartheta}|D=d}(t) P(D = d)$$
$$= \frac{1}{1 - \mathsf{e}^{-nT/\vartheta}} \sum_{d=1}^m \Big[\prod_{j=1}^d \gamma_j\Big] \frac{T^d}{(d-1)!\vartheta^d} B_{d-1}(dt|\gamma_{d+1}T, \ldots, \gamma_1 T) \mathsf{e}^{-dt/\vartheta},$$
$$t \geq 0. \tag{14.3}$$

Notice that, by definition, $\gamma_{m+1} \equiv 0$. Due to the properties of the B-splines, the support of $P^{\widehat{\vartheta}|D\geq 1}$ is given by the union

$$\bigcup_{d=1}^{m} \Big[\frac{\gamma_{d+1}T}{d}, \frac{\gamma_1 T}{d}\Big] \subseteq [0, nT], \tag{14.4}$$

which may have gaps depending on $\gamma_1, \ldots, \gamma_m$.

For order statistics, we observe $Z_{1:n}^{(\mathrm{I})}, \ldots, Z_{m:n}^{(\mathrm{I})}$ so that (14.3) reads as

$$f^{\widehat{\vartheta}|D\geq 1}(t) = \frac{n}{1-\mathrm{e}^{-nT/\vartheta}}\Big(\sum_{d=1}^{m-1}\binom{n-1}{d-1}\frac{T^d}{\vartheta^d}B_{d-1}(dt|(n-d)T, \ldots, nT)\mathrm{e}^{-dt/\vartheta}$$
$$+\binom{n-1}{m-1}\frac{T^m}{\vartheta^m}B_{m-1}(mt|0, (n-m+1)T, \ldots, nT)\mathrm{e}^{-mt/\vartheta}\Big), \quad t \geq 0.$$

For $n = m$, this expression simplifies to

$$f^{\widehat{\vartheta}|D\geq 1}(t) = \frac{n}{1-\mathrm{e}^{-nT/\vartheta}}\sum_{d=1}^{n}\binom{n-1}{d-1}\frac{T^d}{\vartheta^d}B_{d-1}(dt|(n-d)T, \ldots, nT)\mathrm{e}^{-dt/\vartheta},$$
$$t \geq 0,$$

which is the density function of $\widehat{\vartheta}$, given $D \geq 1$, in the case of Type-I censored data.

Figure 14.1 gives plots of $f^{\widehat{\vartheta}|D\geq 1}$ for $T = 1$, $\mu = 0$, and $\vartheta = 1$ and $m = n \in \{2, 3, \ldots, 9\}$. It illustrates that the density function $f^{\widehat{\vartheta}|D\geq 1}$ is multimodal and that its support has gaps (see (14.4)). This can also be seen from the fact that S_d has support $[\frac{n}{d} - 1, \frac{n}{d}]$ in this setting, $1 \leq d \leq n$. Hence, the supports of S_1 and S_d, $d \geq 2$, do not overlap for $n \geq 3$.

From (5.15), the following expression for the (conditional) mean of the MLE results:

$$E(\widehat{\vartheta}|D \geq 1) = \vartheta - \frac{nT\mathrm{e}^{-nT/\vartheta}}{(1-\mathrm{e}^{-T/\vartheta})^n} + \frac{\vartheta T}{(1-\mathrm{e}^{-T/\vartheta})^n}\sum_{d=2}^{m}\frac{1}{d(d-1)}f_{\vartheta}^{X_{d:m:n}}(T). \tag{14.5}$$

Confidence Intervals

Many authors have proposed approaches to construct confidence intervals for ϑ. A standard method to construct exact confidence interval for a parameter ϑ is based on the property that an estimator $\widehat{\vartheta}$ is stochastically increasing in ϑ, i.e.,

$$P_{\vartheta}(\widehat{\vartheta} > t) \text{ is increasing in } \vartheta.$$

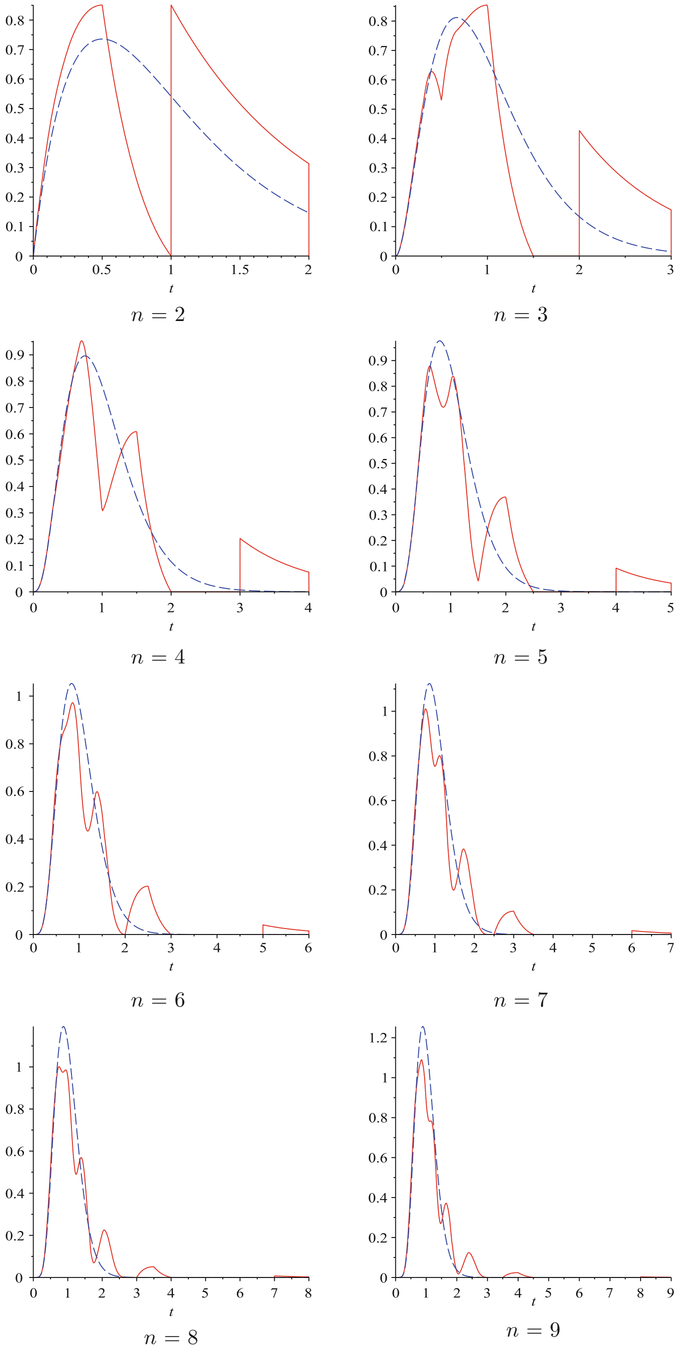

Fig. 14.1 Plots of $f^{\widehat{\vartheta}|D\geq 1}$ (*solid line*) for $T = 1$, $\mu = 0$, and $\vartheta = 1$ and $m = n \in \{2, 3, \dots, 9\}$. The *dashed line* represents the density function of $\widehat{\vartheta}$ in the uncensored case

Then, given $\alpha_1, \alpha_2 \geq 0$ with $\alpha_1 + \alpha_2 = \alpha \in (0,1)$, this result can be used to construct exact confidence intervals with level α as follows:

① Solve the equations

$$P_\vartheta(\widehat{\vartheta} > \vartheta_{\text{obs}}) = \alpha_i, \quad i = 1, 2, \tag{14.6}$$

for ϑ where ϑ_{obs} is the observed value of $\widehat{\vartheta}$;

② Denote the solutions by $\vartheta_L(\vartheta_{\text{obs}})$ and $\vartheta_U(\vartheta_{\text{obs}})$;

③ $[\vartheta_L(\vartheta_{\text{obs}}), \vartheta_U(\vartheta_{\text{obs}})]$ is the realization of a $100(1-\alpha)\%$ confidence interval for ϑ.

The method is described in detail in Casella and Berger [239] and may have been proposed first by Barlow et al. [169] to construct confidence intervals for Type-I censored data (see also Chen and Bhattacharyya [249]).

However, as pointed out in Balakrishnan et al. [156], the procedure may fail for Type-I (progressive hybrid) censored data. In such a setting, the equations (14.6) may have no solution for given values of α_i. Therefore, the exact confidence interval does not exist. Balakrishnan et al. [156] presented an extension to overcome this drawback. But, as a consequence, the resulting confidence intervals may have infinite width.

Balakrishnan and Iliopoulos [101] have established the stochastic monotonicity property for the MLE $\widehat{\vartheta}$ in the setting of hybrid Type-I censored data. Assuming that the survival function of $\widehat{\vartheta}$ has the form

$$P_\vartheta(\widehat{\vartheta} > t) = \sum_{d \in \mathbb{D}} P_\vartheta(D = d) P_\vartheta(\widehat{\vartheta} > t | D = d)$$

with a finite set $\mathbb{D}$, they proved the following result called *Three Monotonicities Lemma* .

Lemma 14.1.1 (Balakrishnan and Iliopoulos [101]). Suppose the following three properties are satisfied:

(M1) For all $d \in \mathbb{D}$, the conditional distribution of $\widehat{\vartheta}$, given $D = d$, is stochastically increasing in ϑ, i.e., the function $P_\vartheta(\widehat{\vartheta} > t | D = d)$ is increasing in ϑ for all t and $d \in \mathbb{D}$;

(M2) For all t and $\vartheta > 0$, the conditional distribution of $\widehat{\vartheta}$, given $D = d$, is stochastically decreasing in d, i.e., the function $P_\vartheta(\widehat{\vartheta} > t | D = d)$ is decreasing in $d \in \mathbb{D}$;

(M3) D is stochastically decreasing in ϑ.

Then, $\widehat{\vartheta}$ is stochastically increasing in ϑ.

However, the monotonicity result for Type-I progressive hybrid censored data remains an open problem.

Open problem 14.1.2. For Type-I progressive hybrid censoring, the conditional survival function $P_\vartheta(\widehat{\vartheta} > t | D \geq 1)$ is increasing in $\vartheta > 0$.

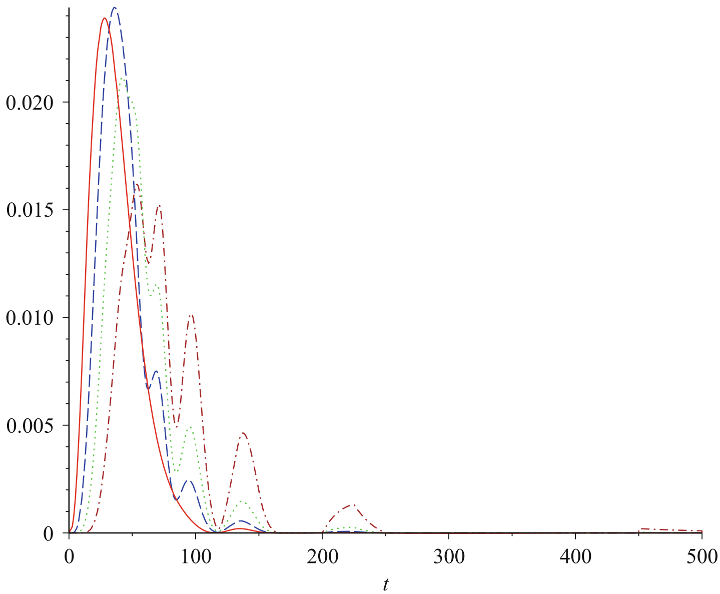

Fig. 14.2 Plots of estimated density functions $f_{m,\widehat{\vartheta}}$ for the data of Barlow with $m = 4$ (*solid red line*), $m = 6$ (*dashed blue line*), $m = 8$ (*dotted green line*), and $m = 10$ (*dashed-dotted brown line*)

Under the assumption that Open Problem 14.1.2 is true, Childs et al. [260] used the above method to construct one- and two-sided confidence intervals.

Example 14.1.3. For illustration, Cramer and Balakrishnan [292] considered data from Barlow et al. [169] (see also Chen and Bhattacharyya [249] and Childs et al. [259, 261]). The data consist of the first six ordered values $4, 9, 11, 18, 27, 38$ of a Type-I censored sample with termination time $T = 50$. Let $m \in \{4, 6, 8, 10\}$. Then, the estimates of ϑ are given by

m	4	6	8	10
$\widehat{\vartheta}$	37.50	43.17	51.17	68.33

The density estimates are given in Fig. 14.2.

Location Parameter Unknown

First, it can be seen that the MLEs of μ and ϑ do not exist when $d = 0$ (the likelihood function is a decreasing function in ϑ). But, when $d > 0$ or equivalently $z_1 < T$, the MLEs of μ and ϑ exist. Writing $z_0 = \mu$ and using the identity in (5.4), (14.2) can be rewritten as

$$\mathsf{L}(\mu,\vartheta;\mathbf{z}_d) = \frac{\prod_{j=1}^d \gamma_j}{\vartheta^d} \exp\Big\{-\frac{1}{\vartheta}\Big[\sum_{j=1}^d \gamma_j (z_j - z_{j-1}) + \gamma_{d+1}(T - z_d)\Big]\Big\},$$
$$\mu \le z_1 \le \cdots \le z_d \le T. \quad (14.7)$$

The likelihood function in (14.7) is an increasing function in μ for any $\vartheta > 0$. Therefore, $\widehat{\mu} = z_1$ is the MLE of μ, and we get the upper bound

$$\mathsf{L}(\widehat{\mu},\vartheta;\mathbf{z}_d) = \frac{\prod_{j=1}^d \gamma_j}{\vartheta^d} \exp\Big\{-\frac{1}{\vartheta}\Big[\sum_{j=2}^d \gamma_j (z_j - z_{j-1}) + \gamma_{d+1}(T - z_d)\Big]\Big\}.$$

Using standard arguments, the MLE of ϑ is given by

$$\widehat{\vartheta} = \frac{1}{d}\Big[\sum_{j=2}^d \gamma_j (Z_j^{(\mathsf{I})} - Z_{j-1}^{(\mathsf{I})}) + \gamma_{d+1}(T - Z_d^{(\mathsf{I})})\Big].$$

Notice that $\widehat{\vartheta}$ is bounded and that $\widehat{\vartheta} = \frac{1}{d} V_d$, with V_d as in (5.12). Thus, we get from $Z_j^{(\mathsf{I})} \le T$, $1 \le j \le d$, and (5.4) with μ replaced by $Z_1^{(\mathsf{I})}$ the upper bounds

$$\mu \le \widehat{\mu} \le T, \quad 0 \le \widehat{\vartheta} \le \frac{T-\mu}{d}\Big(\sum_{j=2}^d (R_j + 1) + \gamma_{d+1}\Big) = \gamma_2(T-\mu), \qquad d \ge 1,$$

which establishes that both $\widehat{\mu}$ and $\widehat{\vartheta}$ have finite supports provided $D \ge 1$. These results enable us to derive the distribution function of $\widehat{\mu}$, given $D \ge 1$.

Theorem 14.1.4. The conditional distribution of $\widehat{\mu}$, given $D \ge 1$, is a right truncated $\mathsf{Exp}(\mu, \vartheta/n)$ distribution with truncation point $T \ge \mu$. Specifically,

$$P\big(\widehat{\mu} > t \big| D \ge 1\big) = \frac{\mathrm{e}^{-n(t-\mu)/\vartheta} - \mathrm{e}^{-n(T-\mu)/\vartheta}}{1 - \mathrm{e}^{-n(T-\mu)/\vartheta}}, \quad \mu \le t \le T. \quad (14.8)$$

Proof. The result follows directly from the observation that the event $\{D \ge 1\}$ is equivalent to $\{Z_{1:m:n} < T\} = \{Z_{1:n} < T\}$. Then,

$$\begin{aligned} P\big(\widehat{\mu} > t \big| D \ge 1\big) &= P\big(\widehat{\mu} > t \big| Z_{1:m:n} < T\big) \\ &= P\big(Z_{1:m:n} > t \big| Z_{1:m:n} < T\big) \\ &= \frac{F_{1:n}(T) - F_{1:n}(t)}{F_{1:n}(T)} = \frac{\mathrm{e}^{-n(t-\mu)/\vartheta} - \mathrm{e}^{-n(T-\mu)/\vartheta}}{1 - \mathrm{e}^{-n(T-\mu)/\vartheta}}. \end{aligned}$$

Remark 14.1.5. Notice that, for $t \ge \mu$, the conditional cumulative distribution function in Theorem 14.1.4 converges to $\mathrm{e}^{-n(t-\mu)/\vartheta}$ when $T \to \infty$, which is the survival function of the minimum in an IID sample of size n from an $\mathsf{Exp}(\mu,\vartheta)$-distribution.

Moreover, it should be noted that, given $D \geq 1$, the distribution of $\widehat{\mu}$ is absolutely continuous w.r.t. the Lebesgue measure. Since $\widehat{\mu} = T$ for $D = 0$, the unconditional distribution function is given by

$$P\Big(\frac{\widehat{\mu}-\mu}{\vartheta} \leq t\Big) = (1 - \mathrm{e}^{-nt})\mathbb{1}_{[0,(T-\mu)/\vartheta)}(t) + \mathbb{1}_{[(T-\mu)/\vartheta,\infty)}(t), \quad t \in \mathbb{R},$$

which has a jump of height $\mathrm{e}^{-n(T-\mu)/\vartheta}$ at $(T-\mu)/\vartheta$.

The expression in (14.8) is very useful in proving the following monotonicity result.

Corollary 14.1.6. $P_\mu\big(\widehat{\mu} > t \big| D \geq 1\big)$ is a monotone increasing function of μ.

Proof. Let $\mu \leq t \leq T$. From (14.8), we have

$$P_\mu\big(\widehat{\mu} > t \big| D \geq 1\big) = P_\mu\Big(\frac{\widehat{\mu}-\mu}{\vartheta} > \frac{t-\mu}{\vartheta} \Big| D \geq 1\Big) = \frac{\mathrm{e}^{-nt/\vartheta} - \mathrm{e}^{-nT/\vartheta}}{\mathrm{e}^{-n\mu/\vartheta} - \mathrm{e}^{-nT/\vartheta}},$$

which is an increasing function of μ. □

Given $D \geq 1$, the MLE of ϑ is given by

$$\widehat{\vartheta} = \frac{1}{D}\Big[\sum_{j=2}^{D} \gamma_j (Z_j^{(\mathrm{I})} - Z_{j-1}^{(\mathrm{I})}) + \gamma_{D+1}(T - Z_d^{(\mathrm{I})})\Big] = \frac{1}{D} V_D,$$

where the sum is defined to be zero when $D = 1$. The conditional bivariate density function of $\widehat{\mu}$ and $\widehat{\vartheta}$ can be taken from (5.14) and is given by

$$\begin{aligned} f^{\widehat{\mu},\widehat{\vartheta}|D=d}(z,t) &= \frac{nd(T-z)^{d-1}\prod_{j=2}^{d+1}\gamma_j}{(d-1)!\vartheta^{d+1} f_{d+1:m:n}(T)} \\ &\quad \times B_{d-2}\Big(dt|\gamma_{d+1}\Big(T-z\Big),\ldots,\gamma_2\Big(T-z\Big)\Big)\exp\Big\{-\frac{n(z-\mu)+dt}{\vartheta}\Big\}, \\ &\qquad \mu \leq z \leq T, 0 \leq t \leq \frac{\gamma_2}{d}(T-\mu). \end{aligned}$$

Expressions for $f^{\widehat{\vartheta}|D\geq 1}$ can be obtained from this expression by integrating w.r.t. z and the following procedure as the one used in the case of a known location parameter.

14.1.2 Other Distributions

Inference for distributions other than exponential has also been addressed in the literature. For instance, Lin et al. [607] and Mokhtari et al. [654] considered Weibull lifetimes. They discussed maximum likelihood estimation as well as approximate maximum likelihood estimation (see Balakrishnan and Varadan [127]). Mokhtari et al. [654] addressed additionally Bayesian inference. For Maxwell distributions, we refer to Tomer and Panwar [848].

14.2 Likelihood Inference for Type-II Progressive Hybrid Censored Data

Given $X_1^{(\mathrm{II})}, \ldots, X_D^{(\mathrm{II})}$ and D as in (5.18) and (2.40) with realizations $x_1, \ldots, x_d$ and d, the likelihood function is given by

$$\mathsf{L}(\boldsymbol{\theta}; \mathbf{x}_d, d) = \begin{cases} f^{\mathscr{R}}_{1,\ldots,d:d:n;\boldsymbol{\theta}}(\mathbf{x}_d), & d \geq m + R_m \\ f^{\mathscr{R}}_{1,\ldots,d+1:d+1:n;\boldsymbol{\theta}}(\mathbf{x}_d, T), & m \leq d < m + R_m \\ f^{\mathscr{R}}_{1,\ldots,m:m:n;\boldsymbol{\theta}}(\mathbf{x}_m), & d < m \end{cases}.$$

14.2.1 Exponential Distribution

Assuming an $\mathsf{Exp}(\vartheta)$-distribution, Childs et al. [260] obtained the MLE

$$\widehat{\vartheta} = \begin{cases} \frac{1}{m}\sum_{j=1}^{m}(R_j+1)X_{j:m:n}, & D < m, \\ \frac{1}{D}\Big(\sum_{j=1}^{m}(R_j+1)X_{j:m:n} + \sum_{j=m+1}^{D} X_{j:m+R_m:n} + \gamma_{D+1}T\Big), & D \geq m. \end{cases}$$

Notice that $\gamma_{m+R_m+1} = 0$.

Then, as pointed out in Cramer et al. [315], the density function of the MLE can be obtained directly from Theorem 5.2.3 and is given in the following theorem. An alternative representation has been established by Childs et al. [260] using a moment generating function approach.

Theorem 14.2.1. The density function of the maximum likelihood estimator $\widehat{\vartheta} = \frac{1}{D}S_D$ is given by

$$f^{\widehat{\vartheta}}(s) = \frac{m^m s^{m-1}\mathrm{e}^{-ms/\vartheta}}{(m-1)!\vartheta^m} - \frac{T^m\prod_{j=1}^m \gamma_j}{(m-1)!\vartheta^m} B_{m-1}(ms|0, \gamma_m T, \ldots, \gamma_1 T)\mathrm{e}^{-ms/\vartheta}$$
$$+ \sum_{d=m}^{m+R_m} \frac{T^d\prod_{j=1}^d \gamma_j}{(d-1)!\vartheta^d} B_{d-1}(ds|\gamma_{d+1}T, \ldots, \gamma_1 T)\mathrm{e}^{-ds/\vartheta}, \quad s \geq 0.$$

Expressions for the cumulative distribution function and the moments of $\widehat{\vartheta}$ are presented in Cramer et al. [315]. Furthermore, the joint density function of the MLEs $\widehat{\mu}$ and $\widehat{\vartheta}$ is presented in the location–scale setting. Using the moment generating function approach, Ganguly et al. [392] have established alternative expressions for Type-II hybrid censored data.

Remark 14.2.2. Bayesian inference for one-parameter exponential distribution has been studied in Kundu and Joarder [561]. The two-parameter exponential distribution has been considered by Kundu et al. [565]. A competing risk model with progressive hybrid censored data has been investigated by Kundu and Joarder [560].

14.2.2 Other Distributions

Statistical inference for hybrid censored data has also been addressed for distributions other than exponential. Type-II progressively hybrid censored Weibull lifetimes have been investigated by Lin et al. [607] and Mokhtari et al. [654]. They obtained (approximate) maximum likelihood estimators for scale and shape parameters. (Approximate) maximum likelihood estimators for a location–scale family of extreme value distributions are discussed in Joarder et al. [479]. Log-normal distributions under Type-II progressive hybrid censoring are discussed by Hemmati and Khorram [439].

14.3 Inferential Results for Adaptive Progressive Type-II Censoring

Let $F_{\boldsymbol{\theta}}$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^q$, be an absolutely continuous cumulative distribution function with density function $f_{\boldsymbol{\theta}}$. The data are given by the sample $\mathbf{y}_m, \mathbf{R}_m$. Then, from (6.3), the likelihood function is given by

$$\begin{aligned} \mathsf{L}(\boldsymbol{\theta}|\mathbf{y}_m, \mathbf{R}_m) &= f_{\boldsymbol{\theta}}^{\mathbf{Y}_{(m)}, \mathbf{R}_m^\star}(\mathbf{y}_m, \mathbf{R}_m) \\ &= f_{\boldsymbol{\theta}}^*(\mathbf{y}_m|\mathbf{R}_{m-1}) g^*(\mathbf{R}_m|\mathbf{y}_m) \propto f_{\boldsymbol{\theta}}^*(\mathbf{y}_m|\mathbf{R}_{m-1}). \end{aligned} \tag{14.9}$$

This illustrates that the maximum likelihood estimators in both the adaptive and nonadaptive cases are identical. This yields directly the following theorem.

Theorem 14.3.1 (Cramer and Iliopoulos [294]). Let $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}(\mathbf{Y}_{(m)}, \mathbf{R}_m)$ be the maximum likelihood estimator of $\boldsymbol{\theta}$ when $\mathbf{R}_m$ is a prefixed censoring scheme. Then, $\widehat{\boldsymbol{\theta}}^* = \widehat{\boldsymbol{\theta}}^*(\mathbf{Y}_{(m)}, \mathbf{R}_m^\star)$ is the maximum likelihood estimator of $\boldsymbol{\theta}$ when $\mathbf{R}_m^\star$ is an adaptive censoring scheme.

Theorem 14.3.1 shows that, for the maximum likelihood estimator, it does not matter whether the censoring scheme has been prefixed in advance or adaptively adjusted in the censoring process. Therefore, the same estimators result as in the nonadaptive model. However, the distribution of the estimators may be different due to the adaptive process. An exception is given by those cases, where Theorem 6.1.1 can be applied.

Example 14.3.2. Suppose $X_j \sim \text{Exp}(\mu, \vartheta)$, $1 \le j \le n$. Then, for $\mu = 0$, the maximum likelihood estimator of ϑ is given by

$$\widehat{\vartheta} = \frac{1}{m} \sum_{j=1}^{m} (1 + R_j^\star) Y_{(j)}.$$

Then, Theorem 6.1.2 yields $2m\widehat{\vartheta}/\vartheta \sim \chi^2(2m)$ so that $E\widehat{\vartheta} = \vartheta$ and $\text{Var}\,\widehat{\vartheta} = \vartheta^2/m$. As in the nonadaptive case, $\widehat{\vartheta}$ is an unbiased estimator of ϑ. It can be shown that it is consistent and that $\sqrt{m}(\widehat{\vartheta} - \vartheta)/\vartheta$ and $\sqrt{m}(\widehat{\vartheta} - \vartheta)/\widehat{\vartheta}$ are asymptotically normal. A two-sided confidence interval for ϑ is given by

$$\Big[\frac{2m\widehat{\vartheta}}{\chi^2_{1-\alpha/2}(2m)}, \frac{2m\widehat{\vartheta}}{\chi^2_{\alpha/2}(2m)}\Big].$$

For an unknown location parameter, the maximum likelihood estimators are given by

$$\widehat{\mu} = Y_{(1)} \quad \text{and} \quad \widehat{\vartheta} = \frac{1}{m} \sum_{j=2}^{m} \Gamma_j (Y_{(j)} - Y_{(j-1)})$$

which, by Theorem 6.1.2, are independent with distributions $\text{Exp}(\mu, \vartheta/n)$ and $\Gamma(\vartheta/m, m-1)$, respectively. Hence, confidence intervals and statistical tests can be constructed as in the nonadaptive model (see, e.g., Corollary 17.1.1).

Remark 14.3.3. Bobotas and Kourouklis [210] addressed a two-sample model that covers adaptive progressive Type-II censoring in the exponential case. They considered estimation of the scale parameters and the hazard rate of the population distribution as well as estimation of the ratio of scales. They established improved estimators for these quantities which are of Stein-type, Brewster and Zidek-type, and Strawderman-type, respectively.

Remark 14.3.4. Adapting the idea of Ng et al. [690], Lin and Huang [600] proposed an adaptive censoring scheme for Type-I censoring called adaptive Type-I progressive hybrid censoring scheme. This procedure works as follows. Given progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{m:m:n}$ and a threshold T, the experiment terminates always at time T. But, the number of observed failure times depends on $X_{m:m:n}$:

(i) if $X_{m:m:n} > T$, the data are given by $X_{1:m:n}, \ldots, X_{D:m:n}$, or

(ii) if $X_{m:m:n} \le T$, the remaining R_m units are not withdrawn from the experiment and all failures are observed until termination time T. This procedure follows the idea of Type-II progressive hybrid censoring.

Lin and Huang [600] presented exact conditional distributions for the MLE of the scale parameter when the population distribution is exponential. They also provided confidence intervals and results on Bayesian inference.

Weibull distributions are discussed in Lin et al. [607] (MLEs and AMLEs) and Lin et al. [610] (likelihood and Bayesian inference). Log-normal distributions are investigated by Hemmati and Khorram [439].

14.3.1 Ng–Kundu–Chan Model

The Ng–Kundu–Chan model, introduced in Ng et al. [690], has been discussed in Sect. 6.2.2. Defining by $\mathbf{R}_m$ the effectively applied censoring scheme in the Ng–Kundu–Chan model given in (6.1), Ng et al. [690] calculated the maximum likelihood estimator of λ provided that the baseline distribution is an $\mathsf{Exp}(1/\lambda)$ distribution. It is given by

$$\begin{aligned}\widehat{\lambda} &= \frac{m}{\sum_{j=1}^{J}(S_j+1)Y_{(j)} + \sum_{j=J+1}^{m-1} Y_{(j)} + \big(n-m-\sum_{j=1}^{J} S_j\big)Y_{(m)}}\\ &= \frac{m}{\sum_{j=1}^{m}(R_j+1)Y_{(j)}},\end{aligned}$$

where the random variable J denotes the change point in the censoring procedure and $\mathbf{R}_m$ is the censoring scheme given in (6.1). Clearly, this result can be directly taken from Theorem 14.3.1 (see also Example 14.3.2). Moreover, Ng et al. [690] sketched several approaches to construct confidence intervals for λ including conditional inference, normal approximations, likelihood-ratio-based inference, bootstrapping, and Bayesian inference. They also presented a formula for the probability mass function of J. From the construction process, it follows that

$$P(J=j) = P(X_{j:m:n} < T \le X_{j+1:m:n}), \quad j=0,\ldots,m,$$

where $X_{0:m:n} = 0$ and $X_{m+1:m:n} = +\infty$. Hence, we find

$$\begin{aligned}P(J=0) &= 1 - F^{X_{1:m:n}}(T) = \mathrm{e}^{-n\lambda T},\\ P(J=j) &= F^{X_{j:m:n}}(T) - F^{X_{j+1:m:n}}(T), \quad j=1,\ldots,m-1,\\ P(J=m) &= F^{X_{m:m:n}}(T).\end{aligned}$$

Simple explicit expressions can be taken from Theorem 2.4.2 and Corollary 2.4.7 (see also Kamps and Cramer [503]). Moreover, an expression for the conditional distribution of $\widehat{\lambda}$, given $J=j$, $j=1,\ldots,m$, is obtained via the conditional moment generating function. However, noticing the results presented in Example 14.3.2, there is no need to consider conditional inference since the maximum likelihood estimator and its distribution can be obtained from Theorems 6.1.2 and 14.3.1.

Lin et al. [607] considered the Ng–Kundu–Chan model with Weibull lifetimes. Using a log-transformation of the data, they got an extreme value distribution as population distribution for which they derived the maximum likelihood estimates of the transformed parameters. However, writing the problem in terms of the effectively applied censoring scheme $\mathbf{R}_m$ given in (6.1), it follows that the likelihood function equals that given in Balakrishnan et al. [136]. Therefore, the maximum likelihood estimates can directly be taken from Balakrishnan et al. [136] which is also clear from Theorem 14.3.1. By analogy with the nonadaptive case, expressions for the approximate maximum likelihood estimates can be obtained because they are deduced from the likelihood equations which are the same in the adaptive and nonadaptive settings (see also Sect. 12.9.2).

14.3.2 Progressive Censoring with Random Removals

The notion of progressive censoring with random removals has been presented in Sect. 6.2.4. Inferential issues for this model have been addressed in numerous papers considering different lifetime distributions as well as various probability mass functions g^*. For further reading, we refer to, e.g., Yuen and Tse [936], Tse and Yuen [859], Tse et al. [860], Tse and Xiang [857], Tse and Yang [858], Wu [905], Wu et al. [916], Wu et al. [919], and Amin [45]. It should be noted that in some of these models (like binomial removals), the probability mass function g^* may depend on parameters which have to be estimated from the data. The same models are also denoted by *progressive Type-II censoring with random scheme* (see, e.g., Sarhan and Al-Ruzaizaa [774]).

In order to illustrate the close connection of inference in the standard model and in the model with random removals, we consider the following standard approach. Suppose the lifetime cumulative distribution function $F_{\boldsymbol{\theta}}$ depends on a parameter $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^q$ and that the probability mass function of $R_m^\star$ depends on a parameter p. Given the data $\mathbf{y}_m, \mathbf{R}_m$, the likelihood function is given by (see (14.9))

$$\mathsf{L}(\boldsymbol{\theta}, p|\mathbf{y}_m, \mathbf{R}_m) = f_{\boldsymbol{\theta}}^*(\mathbf{y}_m|\mathbf{R}_{m-1}) g_p^*(\mathbf{R}_m|\mathbf{y}_m). \tag{14.10}$$

Hence, the likelihood function factorizes into two likelihood functions

$$\mathsf{L}_1(\boldsymbol{\theta}|\mathbf{y}_m, \mathbf{R}_m) = f_{\boldsymbol{\theta}}^*(\mathbf{y}_m|\mathbf{R}_{m-1}) \quad \text{and} \quad \mathsf{L}_2(p|\mathbf{y}_m, \mathbf{R}_m) = g_p^*(\mathbf{R}_m),$$

respectively, where we have also used the fact that the distribution of $R_m^\star$ does not depend on the failure times $\mathbf{y}_m$ (see (6.6)). Thus, the maximization w.r.t. $\boldsymbol{\theta}$ and

p can be handled separately. In particular, the MLE for the parameter $\boldsymbol{\theta}$ is the same as in the standard model (which follows already from (14.9)). The MLE for p depends on the particular assumption on the distributions guiding the random removal of items. However, it is independent of the lifetime cumulative distribution function $F_{\boldsymbol{\theta}}$. Similar comments apply to Bayesian inference in this framework.

In order to present an example for the likelihood inference w.r.t. the parameter p, we present the estimation of p for binomial removals as can be found in, e.g., Tse et al. [860] or Wu and Chang [910]. Then, the (conditional) probability mass functions are given by

$$P(R_1^\star = R_1) = \binom{n-m}{R_1} p^{R_1}(1-p)^{n-m-R_1},$$

$$\begin{aligned} P(R_j^\star = R_j | \mathbf{R}_{j-1}^\star = \mathbf{R}_{j-1}) &= \binom{n-m-R_{\bullet j-1}}{R_j} p^{R_j}(1-p)^{n-m-R_{\bullet j}} \\ &= \binom{n-m-R_{\bullet j-1}}{R_j} p^{R_j}(1-p)^{\gamma_j - m + j - 1 - R_j}, \end{aligned}$$

$2 \le j \le m-1$, so that the part of the likelihood involving p is

$$\mathsf{L}_2(p|\mathbf{y}_m, \mathbf{R}_m) = \binom{n-m}{R_1, \ldots, R_m} p^{R_{\bullet m-1}}(1-p)^{(m-1)(n-m)-\sum_{j=1}^{m-1}(m-j)R_j}.$$

Obviously, this expression is proportional to the likelihood function for a binomial distribution $\mathsf{bin}\big((m-1)(n-m) - \sum_{j=1}^{m-2} R_{\bullet j}, p\big)$ with $R_{\bullet m-1}$ successes, which results in the MLE

$$\begin{aligned} \widehat{p} &= \frac{R_{\bullet m-1}}{(m-1)(n-m) - \sum_{j=1}^{m-1}(m-j-1)R_j} \\ &= \frac{R_{\bullet m-1}}{\sum_{j=1}^{m-1}(\gamma_j - (m-j+1))} = \frac{R_{\bullet m-1}}{\sum_{j=1}^{m-1}\sum_{i=j}^{m} R_i}. \end{aligned}$$

Chapter 15
Bayesian Inference for Progressively Type-II Censored Data

Bayesian inference with progressively Type-II censored data has been studied extensively by many authors in the last decade. Assuming a particular lifetime distribution and a suitable prior distribution for the parameters, Bayesian estimates for the parameters have been obtained w.r.t. several loss functions like squared-error loss or LINEX loss. We use the following notation in this chapter. Let $\mathbf{x} = (x_1, \ldots, x_m)$ be the observed progressively Type-II censored sample, $\mathsf{L}(\boldsymbol{\theta}; \mathbf{x})$ be the likelihood function, and $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^k$ be the parameter (vector) with $k \geq 1$. The prior distribution is defined by the density function $\pi_{\mathbf{a}}(\boldsymbol{\theta})$ with hyperparameters $\mathbf{a} \in \mathfrak{A}$. Then, the posterior distribution exhibits the density function

$$\pi_{\mathbf{a}}^*(\boldsymbol{\theta}|\mathbf{x}) = \frac{\pi_{\mathbf{a}}(\boldsymbol{\theta})\mathsf{L}(\boldsymbol{\theta}; \mathbf{x})}{\int \mathsf{L}(\boldsymbol{\xi}; \mathbf{x})\pi_{\mathbf{a}}(\boldsymbol{\xi})d\boldsymbol{\xi}}. \tag{15.1}$$

Several loss functions have been applied in Bayesian estimation with progressively Type-II censored data. For convenience, we provide the definitions of the subsequently used loss functions:

(i) The squared-error loss (SEL) function is defined by $\mathcal{L}_1(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}) = \|\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\|_2^2$, where $\|\cdot\|_2$ denotes the Euclidean norm. For a one-dimensional parameter, it is given by $\mathcal{L}_1(\theta, \widehat{\theta}) = (\theta - \widehat{\theta})^2$;

(ii) The linear loss function is defined by $\mathcal{L}_2(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}) = \|\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\|_1$, where $\|\cdot\|_1$ is defined by $\|\mathbf{x}\|_1 = \sum_{j=1}^k |x_j|$, $\mathbf{x} \in \mathbb{R}^k$;

(iii) The linear exponential (LINEX) loss function for a one-dimensional parameter is defined by

$$\mathcal{L}_3(\theta, \widehat{\theta}) = b\left[\mathrm{e}^{a\Delta} - a\Delta - 1\right], \quad a \neq 0, b > 0,$$

where $\Delta = \theta - \widehat{\theta}$. A k-parameter extension is defined by $\Delta_j = \theta_j - \widehat{\theta}_j$, $j = 1, \ldots, k$, and

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_15,

$$\mathcal{L}_3(\theta, \widehat{\theta}) = \sum_{j=1}^{k} b_j \big[\mathrm{e}^{a_j \Delta_j} - a_j \Delta_j - 1\big], \quad b_j > 0, j = 1, \dots, k.$$

This loss function was introduced by Varian [870] (see also Varian [871, Chap. 9]) as a measure of asymmetric loss. It was introduced in Bayesian inference by Zellner [938].

The risk function is defined as the expected loss, i.e.,

$$R(\pi_{\mathbf{a}}) = \int E_{\boldsymbol{\theta}}(\mathcal{L}(\boldsymbol{\theta}, \widehat{\boldsymbol{\theta}}))\pi_{\mathbf{a}}(\boldsymbol{\theta}) d\boldsymbol{\theta}.$$

Bayes estimators are defined as estimators leading to minimum risks. For the quadratic loss function, the posterior mean is a Bayes estimator, whereas for the linear loss function, the posterior median provides a solution. Another concept is highest posterior density estimation which leads to the posterior mode as an estimator.

15.1 Exponential and Weibull Distributions

Due to their importance in lifetime analysis, exponential and Weibull distributions have received great attention in Bayesian inference for progressively Type-II censored data. The exponential distribution has been addressed as particular case of the $\mathsf{Weibull}(\alpha, \beta)$-distribution with known shape parameter $\beta = 1$. Schenk [782] and Schenk et al. [783] have considered s independent possibly multiply censored samples of sequential order statistics which includes progressively Type-II censored order statistics as a particular case. It should also be mentioned that the scale parametrization differs in the sense that some works use ϑ or $1/\vartheta$ as a scale parameter. This affects the choice of the prior distribution which is commonly chosen as a gamma or inverse gamma distribution, respectively.

In order to present the results, we assume first the shape parameter $\beta > 0$ to be known which includes the one-parameter exponential distribution with $\beta = 1$. Then, a conjugate prior is given by the gamma density function

$$\pi_{a,b}(\alpha) = \frac{b^a}{\Gamma(a)} \alpha^{a-1} \mathrm{e}^{-b\alpha}, \quad \alpha > 0, \tag{15.2}$$

with hyperparameters $a, b > 0$. Given observations $x_1, \dots, x_m$, this gamma prior yields a $\Gamma(1/[b + \sum_{j=1}^m (R_j + 1)x_j^\beta], a + m)$-distribution as posterior distribution. Under squared-error loss function, the Bayes estimate of the scale parameter α is given by the posterior mean

$$\widehat{\alpha}_{\mathsf{B}} = \frac{a + m}{b + \sum_{j=1}^m (R_j + 1) X_{j:m:n}^\beta}$$

(see Kundu [557]). For $\beta = 1$, the respective estimate for the exponential distribution results. For $\beta = 2$, we get the Rayleigh distribution which has been considered, for instance, in Ali Mousa and Al-Sagheer [36], Wu et al. [915], and Kim and Han [529]. These authors use a different parametrization of the Rayleigh distribution (details are provided Sect. 15.2).

Remark 15.1.1. Schenk et al. [783] considered $\mathsf{Exp}(\vartheta)$-distribution which can be seen as $\mathsf{Weibull}(1/\vartheta, 1)$-distribution. They assumed an inverse gamma prior with a similar parametrization of the gamma distribution given in (15.2)

$$\pi_{a,b}(\vartheta) = \frac{b^a}{\Gamma(a)}\vartheta^{-(a+1)}\exp\Big\{-\frac{b}{\vartheta}\Big\}, \quad \vartheta > 0, \tag{15.3}$$

with hyperparameters $a, b > 0$. The corresponding Bayes estimator of ϑ has the representation

$$\begin{aligned}\widehat{\vartheta}_{\mathsf{B}} &= \frac{1}{a+m-1}\Big(\sum_{j=1}^{m}(R_j+1)X_{j:m:n} + b\Big)\\ &= \frac{1}{a+m-1}\Big(\sum_{j=1}^{m}\gamma_j(X_{j:m:n} - X_{j-1:m:n}) + b\Big),\end{aligned} \tag{15.4}$$

where $X_{0:m:n} = 0$.

In addition to the case of a known shape parameter, Kundu [557] also addressed the problem of two unknown parameters $\alpha, \beta > 0$. Since a continuous bivariate conjugate prior does not exist in this case, the preferable prior distribution is not clear. Soland [810] has shown that a continuous–discrete conjugate prior exists, where the shape parameter may only take on values in a finite set $\{\beta_1, \ldots, \beta_k\}$. Denoting by $p_j = P(\beta = \beta_j)$, $j \in \{1, \ldots, k\}$, the probability mass function of β, the joint prior of α and β is assumed to be

$$\pi_{a_j,b_j}(\alpha, \beta) = p_j\frac{b_j^{a_j}}{\Gamma(a_j)}\alpha^{a_j-1}\mathsf{e}^{-b_j\alpha}, \quad \alpha > 0, \tag{15.5}$$

where $a_j, b_j > 0$. Notice that the conditional prior of α given $\beta = \beta_j$ is a gamma distribution with hyperparameters a_j and b_j (see (15.2)). Proceeding as above, the posterior distribution is specified by

$$\pi^*_{a_j,b_j}(\alpha, \beta_j) = \frac{p_j b_j^{a_j}\beta_j^m\Big[\prod_{i=1}^m x_i\Big]^{\beta_j-1}}{\Gamma(a_j)c(\mathbf{x})}\alpha^{m+a_j-1}\mathsf{e}^{-\alpha w_j},$$

where $c(\mathbf{x}) = \sum_{\nu=1}^{k} p_\nu b_\nu^{a_\nu} \beta_\nu^m \Big[\prod_{i=1}^m x_i\Big]^{\beta_\nu - 1} \frac{\Gamma(m+a_\nu)}{\Gamma(a_\nu)} w_\nu^{-(m+a_\nu)}$ and $w_j = b_j + \sum_{i=1}^m (R_i + 1) x_i^{\beta_j}$. As in the one-parameter scale model, the posterior distribution is of the same type as the prior distribution. Hence, statistical inference can be carried out as above. Although this approach is quite tempting for computational reasons, it has been criticized in the literature. Kaminskiy and Krivtsov [494] pointed out that the discrete part of the prior exhibits some difficulties in practice. They criticized the problem of evaluating the prior information. To overcome this, Kundu [557] adapted an approach proposed by Berger and Sun [191] (see also Kundu and Gupta [559]). The random parameters α and β are supposed to be independent. The marginal priors are gamma distributions $\pi_\mathbf{a}$ and $\pi_\mathbf{c}$ as in (15.2). Using Lindley's approximation (see Lindley [612]) under squared-error loss, Kundu [557] established the following estimates ($\widehat{\alpha}$ and $\widehat{\beta}$ denote the MLEs of α and β, respectively):

$$\begin{aligned}\widehat{\alpha}_\mathsf{B} = \widehat{\alpha} + \frac{1}{2}\bigg[&\bigg(\frac{2m}{\widehat{\beta}^3} - \widehat{\alpha}\sum_{i=1}^m (R_i+1) x_i^{\widehat{\beta}} (\log x_i)^3\bigg)\tau_{11}\tau_{12} \\ &+ \frac{2m}{\widehat{\alpha}^3}\tau_{22}^2 - (\tau_{12}\tau_{22} + 2\tau_{12}^2)\sum_{i=1}^m (R_i+1) x_i^{\widehat{\beta}}(\log x_i)^2\bigg] \\ &+ \tau_{21}\Big(\frac{c-1}{\widehat{\beta}} - d\Big) + \tau_{22}\Big(\frac{a-1}{\widehat{\alpha}} - b\Big), \\ \widehat{\beta}_\mathsf{B} = \widehat{\beta} + \frac{1}{2}\bigg[&\bigg(\frac{2m}{\widehat{\beta}^3} - \widehat{\alpha}\sum_{i=1}^m (R_i+1) x_i^{\widehat{\beta}} (\log x_i)^3\bigg)\tau_{11}^2 \\ &+ \frac{2m}{\widehat{\alpha}^3}\tau_{21}\tau_{22} - 3\tau_{11}\tau_{12}\sum_{i=1}^m (R_i+1) x_i^{\widehat{\beta}}(\log x_i)^2\bigg] \\ &+ \tau_{11}\Big(\frac{c-1}{\widehat{\beta}} - d\Big) + \tau_{12}\Big(\frac{a-1}{\widehat{\alpha}} - b\Big),\end{aligned}$$

where $\mathbf{a} = (a, b)$ and $\mathbf{c} = (c, d)$ are the hyperparameters, and

$$U = \frac{m}{\widehat{\beta}^2} + \widehat{\alpha}\sum_{i=1}^m (R_i+1) x_i^{\widehat{\beta}}(\log x_i)^2,\ V = \sum_{i=1}^m (R_i+1) x_i^{\widehat{\beta}} \log x_i,\ W = \frac{m}{\widehat{\alpha}^2},$$

$$\tau_{11} = \frac{W}{UW - V^2},\ \tau_{22} = \frac{U}{UW - V^2},\ \tau_{12} = \tau_{21} = -\frac{V}{UW - V^2}.$$

Although Lindley's approach yields explicit estimates of the distribution parameters, it does not provide credible estimates. Therefore, Kundu [557] proposed MCMC technique to compute the Bayesian estimates and to construct credible

11	35	49	170	329
958	1925	2223	2400	2568

Table 15.1 Progressively Type-II censored data from the electrical alliance test data of Lawless [575] as given in Kundu [557]

intervals (see Sect. 17.6). Alternatively, Kundu [557] utilized a Gibbs sampling procedure and an idea of Geman and Geman [394]. A detailed description of the simulational algorithm as well as an extensive simulation comparing the estimates can be found in Kundu [557].

Example 15.1.2. From failure data for electrical alliance test reported in Lawless [575, p. 8], Kundu [557] generated a progressively Type-II censored sample with censoring plan $\mathscr{R} = (2^{*9}, 8)$ $(n = 36, m = 10)$. It is given in Table 15.1. The applied censoring scheme is given by $\mathscr{R} = (0^{*2}, 3, 0^{*2}, 2, 0^{*2}, 2, 0, 2, 1, 0)$. The maximum likelihood estimates were computed as $\widehat{\alpha} = 0.0627$ and $\widehat{\beta} = 0.6298$. Assuming noninformative gamma priors for both parameters, Lindley's approximation yields the estimates 0.0699 and 0.6283, respectively. Using the MCMC technique, Kundu [557] reported the estimates 0.0679 and 0.6223. Notice that these results were obtained for the above data divided by 100.

Remark 15.1.3. Bayesian inference for Weibull distributions has also been considered by Li et al. [588]. Exponentiated Weibull distributions with cumulative distribution function

$$F(t) = \left[1 - \mathrm{e}^{-t^{\beta}}\right]^{\alpha}, \quad t \geq 0,$$

have been investigated by Kim et al. [531]. Estimates were obtained by Lindley's approximation method. Exponentiated exponential distributions

$$F(t) = \left[1 - \mathrm{e}^{-\alpha t}\right]^{\beta}, \quad t \geq 0,$$

are considered in Madi and Raqab [627] and Kundu and Pradhan [562] for a quadratic loss function (the latter is along the lines of Kundu [557]). Progressively Type-II censored samples generated from data given in Lieblein and Zelen [595] are applied to illustrate the method in both papers. The paper of Elkahlout [349] deals with the same topic. In addition, LINEX loss functions are considered.

Remark 15.1.4. Fernández [364] established results for general progressively Type-II censored data from exponential distributions using an inverse gamma prior (15.3). According to Sect. 1.1.1, general progressive Type-censored data can be seen as a left censored sample $X_{r+1:m:n}, \dots, X_{m:m:n}$ with censoring scheme $\mathscr{R} = (0^{*r}, R_{r+1}, \dots, R_m)$, where $1 \le r < m$. Introducing the quantities $(p, q > 0)$

$$C_r[p,q] = p^q \sum_{j=0}^{r} \frac{(-1)^j \binom{r}{j}}{(p + jx_{r+1})^q},$$

$$D_r[p,q;\vartheta] = p^q \sum_{j=0}^{r} \frac{(-1)^j \binom{r}{j}}{(p + jx_{r+1})^q} \mathsf{IG}((p + jx_{r+1})/\vartheta, q),$$

the posterior density function is given by

$$\pi^*_{a,b}(\vartheta|\mathbf{x}) = \frac{(b+w)^{a+m-r}\left[1 - \exp\{-x_{r+1}/\vartheta\}\right]^r \exp\{-(b+w)/\vartheta\}}{\Gamma(a+m-r)C_r[b+w, a+m-r]\vartheta^{a+m-r+1}}, \quad \vartheta > 0.$$

Here, $w = \sum_{j=r+1}^{m}(R_j+1)x_j$, where $x_{r+1}, \dots, x_m$ denote the observations. For squared-error loss, Fernández [364] calculated the Bayes estimate as

$$\widehat{\vartheta}_{\mathsf{B}} = \frac{C_r[b+W, a+m-r-1]}{C_r[b+W, a+m-r]} \cdot \frac{b+W}{a+m-r-1}.$$

For a complete sample of progressively Type-II censored order statistics, i.e., $r = 1$, this yields the Bayes estimator

$$\widehat{\vartheta}_{\mathsf{B}} = \frac{b+T_m}{a+m-1}, \tag{15.6}$$

where $T_m = \sum_{j=1}^{m} \gamma_j (X_{j:m:n} - X_{j-1:m:n})$, $X_{0:m:n} = 0$, as given in (15.4).

Fernández [364] also discussed HPD estimation. He showed that the HPD estimator of ϑ is the solution of the equation

$$(a+m-r+1)\vartheta + \frac{r x_{r+1}}{\exp\{x_{r+1}/\vartheta\} - 1} = b + w,$$

which is quite similar to the likelihood equation (12.5) to be solved for the MLE of ϑ (choose $a = -1$ and $b = 0$). HPD estimators of the reliability $R(t)$ as well as of the reciprocal ϑ^{-1} are also established. Finally, the results are applied to Nelson's insulating fluid data 1.1.5, where the first failure time is additionally censored.

Remark 15.1.5. Multiply Type-II censored samples from an exponential-type family of distributions were addressed by Abdel-Aty et al. [2]. They assumed a family of distributions defined by the cumulative distribution function

$$F_\theta(t) = 1 - \mathsf{e}^{-\lambda_\theta(t)}, \quad t \ge 0, \tag{15.7}$$

where λ_θ is a nonnegative continuous differentiable function with limits $\lim_{t\to 0+}\lambda_\theta(t) = 0$ and $\lim_{t\to\infty}\lambda_\theta(t) = \infty$. θ is supposed to be the parameter with $\theta \in \Theta \subseteq \mathbb{R}^k$, where $k \geq 1$. Priors are chosen as in AL-Hussaini and Ahmad [28] who considered Bayesian inference in the framework of m-generalized order statistics.

Remark 15.1.6. Two-parameter exponential distributions $\mathsf{Exp}(\mu, \vartheta)$ have been addressed in a multi-sample setting of sequential order statistics with multiply censoring by Schenk [782] and Schenk et al. [783] (see also Shafay et al. [794]). They used a bivariate prior $\pi(\vartheta, \mu)$ previously discussed in Evans and Nigm [357] and Varde [869]:

$$\pi(\vartheta, \mu) \propto \vartheta^{-(a+1)} \exp\left\{-\frac{b - c\mu}{\vartheta}\right\}\mathbb{1}_{[N,M]}(\mu), \quad \vartheta > 0, \mu \in \mathbb{R},$$

with $b > cM$, $a > 1$, $c > 0$, and $M > N$ with $N, M \in \mathbb{R}$.

Let $m + a - 2 > 0$, $c > -n$, $N < x_1$, and $H(y) = \sum_{j=1}^{m}(R_j + 1)X_{j:m:n} + b - y(c + n)$. In the case of one sample of progressively Type-II right censored order statistics, the corresponding Bayes estimators of μ and ϑ are given by

$$\widehat{\vartheta}_{\mathsf{B}} = \frac{1}{m + a - 2}\frac{H^{-(m+a-2)}(M_0) - H^{-(m+a-2)}(N)}{H^{-(m+a-1)}(M_0) - H^{-(m+a-1)}(N)},$$

$$\widehat{\mu}_{\mathsf{B}} = \frac{M_0 H^{-(m+a-1)}(M_0) - N H^{-(m+a-1)}(N)}{H^{-(m+a-1)}(X_{1:m:n}) - H^{-(m+a-1)}(N)} - \widehat{\vartheta}_{\mathsf{B}}(c + n)^{-1},$$

where $M_0 = \min\{X_{1:m:n}, M\}$. Further details like posterior density function and posterior survival function can be found in Schenk et al. [783].

15.2 Rayleigh Distribution

As mentioned above, different parametrizations of the Rayleigh distribution are used in Bayesian inference. Moreover, in addition to the scale model, location–scale families of Rayleigh distributions have been considered. In the following, we present the results of Ali Mousa and Al-Sagheer [36] and Wu et al. [915] who use different parametrizations. Wu et al. [915] presumed a scale model with $F_\vartheta(t) = 1 - \exp\{-t^2/(2\vartheta^2)\}$. As a prior, they used a square-root inverted gamma distribution defined by the density function

$$\pi_{a,b}(\vartheta) = \frac{b^a}{\Gamma(a)2^{a-1}}\vartheta^{-(2a+1)} \exp\left\{-\frac{b}{2\vartheta^2}\right\}, \quad \vartheta > 0, \tag{15.8}$$

with hyperparameters $a, b > 0$. Notice that this is the natural conjugate prior in this setting (see Fernández [363]). This yields a posterior distribution of the same kind so that, for squared-error loss, the expression

17.88	28.92	33.00	42.12	45.60	48.48	51.84
51.96	67.80	68.64	84.12	93.12	127.92	

Table 15.2 Progressively Type-II censored data from Lieblein's groove ball bearings data as given in Wu et al. [915]

$$\widehat{\vartheta}_{\mathrm{B}} = \sqrt{\frac{1}{2}(b+W)} \cdot \frac{\Gamma(a+m-1/2)}{\Gamma(a+m)} \tag{15.9}$$

of the Bayes estimator of ϑ results, where $W = \sum_{j=1}^{m}(R_j + 1)X_{j:m:n}^2$. The corresponding Bayes estimate of the reliability $R_\vartheta(t)$, $t > 0$, is given by

$$\widehat{R}_{\mathrm{B}}(t) = \left[\frac{W+b}{W+b+t^2}\right]^{m+a+1}.$$

Wu et al. [915] also considered highest posterior (HPD) estimation which means that the mode of the posterior density function is used as an estimate. The HPD estimator of ϑ and the reliability $R_\vartheta(t)$ are given by

$$\widehat{\vartheta}_{\mathrm{B}}^* = \sqrt{\frac{b+W}{2(a+m)+1}}, \quad \widehat{R}_{\mathrm{B}}^*(t) = \exp\left\{-\frac{(a+m-1)t^2}{W+b-t^2}\right\},$$

respectively.

Example 15.2.1. Wu et al. [915] generated a progressively Type-II censored sample from data reported originally by Lieblein and Zelen [595]. The data are measurements for the fatigue life of groove ball bearings. The measurements are the number of million revolutions to failure for each of $n = 23$ ball bearings in a fatigue test. Probability plots suggest a good fit by Weibull and log-normal distributions. Raqab and Madi [744] suggested a one-parameter Rayleigh distribution to analyze the complete sample. Following the Bayesian analysis of Raqab and Madi [744], Wu et al. [915] applied a square-root inverted gamma prior with $a = b = 2$ as above to illustrate their results. The progressively Type-II censored data with $m = 13$ observations are given in Table 15.2. The applied censoring scheme is given by $\mathscr{R} = (0^{*2}, 3, 0^{*2}, 2, 0^{*2}, 2, 0, 2, 1, 0)$. The resulting Bayes estimates of ϑ and $R(1)$ are given by $\widehat{\vartheta}_{\mathrm{B}} = 0.6052$ and $\widehat{R}_{\mathrm{B}}(1) = 0.2554$, respectively. Notice that these results were obtained for the above data divided by 100.

Ali Mousa and Al-Sagheer [36] assumed a location–scale model with cumulative distribution function

$$F_\vartheta(t) = 1 - \exp\{-(t-\mu)^2/(2\vartheta)\}, \quad t \geq 0, \quad \mu \in \mathbb{R}, \vartheta < 0.$$

Choosing an inverse gamma distribution as a prior in the scale case (i.e., $\mu = 0$; see (15.3)) and squared-error loss, Ali Mousa and Al-Sagheer [36] obtained the Bayes estimator

$$\widehat{\vartheta}_{\mathsf{B}} = \frac{b + W/2}{a + m - 1}$$

of α. The corresponding Bayes estimate of the reliability is given by

$$\widehat{R}_{\mathsf{B}}(t) = \left[\frac{W + 2b}{W + 2b + t^2}\right]^{m+a+1}.$$

Notice that the expression in Ali Mousa and Al-Sagheer [36] is in error.

In the location–scale model, Ali Mousa and Al-Sagheer [36] proposed the joint prior

$$\pi_{a,b,s}(\vartheta, \mu) = \frac{b^a}{s\Gamma(a)}\vartheta^{-(a+1)} \exp\left\{-\frac{b}{\vartheta}\right\}, \quad \mu \in [0, s], \vartheta > 0,$$

with hyperparameters $a, b, s > 0$. Notice that the parameters μ and ϑ are supposed to be independent. μ has a uniform distribution, whereas ϑ is supposed to have an inverse gamma distribution. For a quadratic loss function, the resulting estimators are only available as integral expressions which have to be evaluated by numerical integration. Further details as well as some simulation results can be found in Ali Mousa and Al-Sagheer [36].

Remark 15.2.2. Bayesian inference for Rayleigh distributions with general progressively Type-II censored data has been considered in Kim and Han [529]. Bayesian inference for finite mixtures of Rayleigh distributions has been addressed by Soliman [812].

15.3 Pareto Distribution

Ali Mousa [34] considered progressively Type-II censored data from two-parameter Pareto distribution with cumulative distribution function as in (12.24). He distinguished three cases: λ known, α known, and both parameters unknown. For a known scale parameter λ, he assumed a gamma prior as given in (15.2). The corresponding posterior distribution is a $\Gamma(1/[\beta + \sum_{j=1}^{m}(R_j + 1)\log(\lambda x_j)], m + a)$-distribution. For squared-error loss function, this yields the Bayes estimator

$$\widehat{\alpha}_{\mathsf{B}} = \left[\frac{1}{m + a}\Big(b + \sum_{j=1}^{m}(R_j + 1)\log(\lambda X_{j:m:n})\Big)\right]^{-1}.$$

The Bayes estimator of the reliability is given by

$$\widehat{R}_{\mathrm{B}}(t) = \left[1 + \frac{\log(\lambda t)}{b + \sum_{j=1}^{m}(R_j+1)\log(\lambda X_{j:m:n})}\right]^{-(m+a)}.$$

When the shape parameter is supposed to be known, a Pareto prior with density function

$$\pi_{c,d}(\lambda) = cd(d\lambda)^{-(c+1)}, \quad d\lambda > 1,$$

is known to be conjugate to the Pareto distribution (12.24) with scale parameter λ. The posterior distribution is a Pareto distribution with density function

$$\pi^*_{c,d}(\lambda) = \frac{n\alpha + c}{\widehat{\delta}^{n\alpha+c}}\lambda^{-(n\alpha+c+1)}, \quad \widehat{\delta}\lambda > 1,$$

where $\widehat{\delta} = \min(d, x_1)$. This defines a two-parameter Pareto distribution as in (12.24) with scale parameter $n\alpha+c$ and $\widehat{\delta}$. For squared-error loss, the corresponding Bayes estimators are the posterior means given by

$$\widehat{\lambda}_{\mathrm{B}} = \frac{1}{\widehat{\delta}}\Big(1 + \frac{1}{n\alpha + c - 1}\Big) \quad \text{and} \quad \widehat{R}_{\mathrm{B}}(t) = \frac{n\alpha + c}{(n+1)\alpha + c}\Big(\frac{t}{\widehat{\delta}}\Big)^{-\alpha}.$$

In the two-parameter setup, Ali Mousa [34] assumed a prior with density function

$$\pi_{a,b,c,d}(\lambda, \alpha) = \frac{cdb^a}{\Gamma(a)}\alpha^a(d\lambda)^{-(c\alpha+1)}\mathrm{e}^{-b\alpha}, \quad \alpha > 0, d\lambda > 1.$$

This prior was proposed first by Lwin [623] (see also Arnold and Press [53, 54]). The posterior density function is given by

$$\pi^*_{a,b,c,d}(\lambda, \alpha) = C\alpha^{m+a}\lambda^{-((c+n)\alpha+1)}\mathrm{e}^{-\alpha T}, \quad \alpha > 0, \widehat{\delta}\lambda > 1,$$

where $\widehat{\delta} = \min(d, x_1)$ and

$$T = b + c\log d + \sum_{j=1}^{m}(R_j+1)\log x_j, \quad C = \frac{(n+c)(m+a)}{\Gamma(m+a)}[T - (n+c)\log\widehat{\delta}].$$

Explicit expressions for the resulting Bayes estimators are generally not available. Therefore, Ali Mousa [34] used the approximation method of Tierney and Kadane [845] to establish expressions for approximate Bayesian estimates.

i	1	2	3	4	5	6	7	8	9	10	11
$x_{i:11:17}$	23	27	39	127	136	280	624	730	836	1,024	1,349

Table 15.3 Progressively Type-II censored sample as reported in Soliman [813] with censoring scheme $\mathscr{R} = (1, 0^{*3}, 1, 0^{*2}, 2, 0, 2, 0)$

Soliman [813] studied Bayesian estimation for Lomax distributions from general progressively Type-II censored samples under squared-error loss and LINEX loss. He used the parametrization

$$F(t) = 1 - \left(1 + \frac{t}{\vartheta}\right)^{-\alpha}, \quad t \geq 0, \alpha, \vartheta > 0.$$

As a joint prior, he assumed a discrete–continuous prior similar to that proposed by Soland [810] for the Weibull distribution (see (15.5)). The author established explicit expressions for the Bayes estimators of the distribution parameters as well as the reliability. Further results in this direction can be found in Fu et al. [384] (see also Amin [45]). They analyzed progressively Type-II censored sample generated from data reported by Crowley and Hu [317]. The data set is given in Table 15.3.

15.4 Burr Distributions

For the two-parameter Burr-XII distribution with cumulative distribution function

$$F(t) = 1 - (1 + t^{\beta})^{-\alpha}, \quad t \geq 0, \alpha, \beta > 0,$$

Ali Mousa and Jaheen [38] considered Bayesian inference in one- and two-parameter settings. Supposing β to be known, they suggested a gamma prior as in (15.2) resulting in a $\Gamma(1/[b + T], m + a)$-posterior distribution with $T = \sum_{j=1}^{m}(R_j + 1)\log(1 + x_i^{\beta})$. For squared-error loss, the Bayesian estimates are given by

$$\widehat{\alpha}_{\mathsf{B}} = \frac{b + T}{m + a} \quad \text{and} \quad \widehat{R}_{\mathsf{B}}(t) = \left[1 + \frac{\log(1 + t^{\beta})}{b + T}\right]^{-(m+a)}.$$

Asgharzadeh and Valiollahi [63] used an exponential prior $\mathsf{Exp}(1/b)$, $b > 0$, for α. Assuming a linear loss function, they calculated the median of the gamma posterior distribution as

$$\widehat{\alpha}_{\mathsf{B}} = \frac{\chi^2_{1/2}(2m + 2)}{2(b + T)}.$$

They also applied a logarithmic loss function introduced by Brown [221] resulting in the estimate $\widehat{\alpha}_B = e^{\Psi(m+1)}/(b+T)$, where Ψ denotes the Digamma function. Li et al. [587] studied the same problem under LINEX loss. Furthermore, Asgharzadeh and Valiollahi [63] and Li et al. [587] considered empirical Bayesian estimation.

In the two-parameter model, they assumed a joint prior suggested by AL-Hussaini and Jaheen [29]:

$$\pi_{a,b,c,d}(\alpha,\beta) = \frac{1}{\Gamma(d)\Gamma(a)c^d b^a}\beta^{a+d}\alpha^a \exp\Big\{-\beta[1/c+\alpha/b]\Big\}, \quad \alpha,\beta > 0.$$

Using the approximation of Tierney and Kadane [845], they obtained approximate Bayesian estimates following the lines of AL-Hussaini and Jaheen [30] who considered Type-II right censored data. The same prior has been suggested by Jaheen [475] for m-generalized order statistics who used the approach of Lindley [612] to compute approximate Bayesian estimates. Soliman [811] utilized a discrete–continuous prior similar to that proposed by Soland [810] for the Weibull distribution (see (15.5)). Results for squared-error loss, LINEX loss, and general entropy loss were obtained. Due to the structure of the prior distribution, explicit representations of the Bayes estimates result. The results are used to analyze Nelson's insulation fluid data 1.1.5.

15.5 Other Distributions

Bayesian inference for linear hazard rate distributions with

$$F(t) = 1 - e^{-(\lambda t + \nu t^2/2)}, \quad t \geq 0, \lambda \geq 0, \nu > 0, \tag{15.10}$$

has been investigated by Lin et al. [604] for a general progressively Type-II censored sample $X_{r+1:m:n}, \dots, X_{m:m:n}$. The applied censoring scheme is given by $\mathscr{R} = (0^{*r}, R_{r+1}, \dots, R_m)$, where $1 \leq r < m$. As Ashour et al. [65], they suggested a bivariate gamma prior with independent marginals

$$\pi_{a_1,b_1,a_2,b_2}(\lambda,\nu) \propto \lambda^{a_1-1}\nu^{a_2-1}e^{-(b_1\lambda+b_2\nu)}, \quad \lambda,\nu > 0.$$

The corresponding posterior distribution has a density function

$$\pi^*_{a_1,b_1,a_2,b_2}(\lambda,\nu) \propto \lambda^{a_1-1}\nu^{a_2-1}\Big\{1 - e^{-(\lambda x_{r+1}+\nu x_{r+1}^2/2)}\Big\}^r e^{-(\lambda T_1+\nu T_2)} \times \prod_{j=1}^{m-r}(\lambda+\nu x_{r+j}) \quad \lambda,\nu > 0,$$

where $T_1 = \sum_{j=1}^{m-r}(R_{r+j}+1)x_{r+j} + b_1$ and $T_2 = \frac{1}{2}\sum_{j=1}^{m-r}(R_{r+j}+1)x_{r+j}^2 + b_2$. Explicit expressions for the Bayes estimators of λ and ν have been established

1.96	1.96	3.60	3.80	4.79	5.66	5.78	6.27	6.30	6.76	7.65	7.99
8.51	9.18	10.13	10.24	10.25	10.43	11.45	11.75	11.81	12.34	12.78	

Table 15.4 Progressively Type-II censored data set generated from 33 observations of the floods of the Fox River data by Lin et al. [604]. The applied censoring scheme was given by $\mathscr{R} = (0^{*5}, 1, 0^{*4}, 1, 0^{*2}, 1, 0^{*4}, 2, 0^{*3}, 5)$. $m = 23$ observations are available; ten observations have been progressively censored

for squared-error loss. Additionally, an efficient MCMC method to generate the posterior distributions of interest has been implemented. Moreover, the estimators are applied to a progressively Type-II censored data set generated from the floods of the Fox River data earlier analyzed by Gumbel and Mustafi [422] and Bain and Engelhardt [73]. The progressively Type-II censored sample is given in Table 15.4.

Example 15.5.1. Lin et al. [604] applied their results to the data given in Table 15.4. Using the MCMC approach, the estimates $\widehat{\lambda} = 0.000086$ and $\widehat{\nu} = 0.016201$ result. Details on the prediction results can be found in Table 3 of Lin et al. [604].

Remark 15.5.2. Bayesian inference for linear hazard rate distributions has also recently been discussed by Sen et al. [792] using a unified treatment of both progressive Type-I and Type-II censoring. Using gamma priors as above, mixtures of gamma distributions result as posterior distributions.

A location–scale family of extreme value distributions with standard member $F(t) = 1 - \mathrm{e}^{-\mathrm{e}^t}$, $t \in \mathbb{R}$, has been considered by Al-Aboud [26] under squared-error loss, LINEX loss, and general entropy loss as proposed by Calabria and Pulcini [238]. The prior distribution is suggested as

$$\pi_{a,b}(\mu, \vartheta) \propto \frac{b^a}{\Gamma(a)} \vartheta^{-(a+2)} \mathrm{e}^{-b/\vartheta}, \quad \vartheta > 0, \mu \in \mathbb{R}.$$

Notice that Jeffrey's noninformative prior on $\mathbb{R}$ is used for the location parameter μ. The results are applied to log-times of Nelson's insulating fluid data given in Table 17.5. Exponentiated modified Weibull distributions with cumulative distribution function as in (12.35) are studied by Klakattawi et al. [534] using independent gamma priors. Prediction problems are considered in Klakattawi et al. [535].

Rastogi et al. [749] and Sarhan et al. [778] discussed Bayesian inference for the bathtub-shaped distribution given in (12.43). They considered squared-error, LINEX, and entropy loss functions with gamma priors. The results are illustrated by the data given in Table 12.4. Bayesian inference for logistic-type distributions has been addressed by some authors. Half-logistic distributions are considered by Kim and Han [530]. Rastogi et al. [749] discussed exponentiated half-logistic distributions under progressive Type-II censoring. Results for generalized half-normal distributions with cumulative distribution function in (12.44) have been established in Ahmadi et al. [19] assuming independent gamma priors.

Chapter 16
Point Prediction from Progressively Type-II Censored Samples

Prediction problems are discussed extensively in lifetime analysis. Based on a sample $X_1, \ldots, X_m$, we wish to predict the outcome of a random variable Y or random vector $\mathbf{Y}$. This random variable may be a (censored) observation from the same experiment (one-sample prediction) or may be part of an independent future sample (two-sample prediction). In the first setup, Y and the X sample will be correlated. In the following, we consider two setups. Given a progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$, we want to predict the lifetime of an item censored in the progressive censoring procedure. Moreover, we address the problem of predicting random variables in a future sample $Y_1, \ldots, Y_r$. Here, it is important which distributional assumption is put on that sample. Point and interval predictions are both discussed.

16.1 Prediction Concepts

Before turning to specific prediction problems, some general concepts of point prediction are briefly introduced. We assume an informative sample $\mathbf{Y} = (Y_1, \ldots, Y_m)$ used to predict a single observation Y. The best unbiased predictor (BUP) $\widehat{\pi}_{\mathsf{BU}}(\mathbf{Y})$ is given by the conditional expectation $\widehat{\pi}_{\mathsf{BU}}(\mathbf{Y}) = E_{\boldsymbol{\theta}}(Y|\mathbf{Y})$. If $\boldsymbol{\theta}$ is known, the expectation can be calculated directly from the conditional distribution of Y, given $\mathbf{Y}$. Unknown parameters can be replaced by appropriate estimates. If $\boldsymbol{\theta}$ is unknown, we can apply a result due to Ishii [471] which is reported in Takada [832]: $\widehat{\pi}_{\mathsf{BU}}(\mathbf{Y})$ is BUP iff

$$E_{\boldsymbol{\theta}}\big([Y - \widehat{\pi}_{\mathsf{BU}}(\mathbf{Y})] \cdot \varepsilon_0(\mathbf{Y})\big) = 0 \quad \text{for all } \boldsymbol{\theta} \tag{16.1}$$

for any squared integrable estimator $\varepsilon_0(\mathbf{Y})$ which is an unbiased estimator of zero. This characterization result is similar to that for unbiased estimators.

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_16,

An important concept in point prediction is linear prediction. A linear predictor is a linear function of the observations $Y_1, \ldots, Y_m$: $\widehat{\pi}_{\mathsf{L}}(\mathbf{Y}) = \sum_{j=1}^{m} c_j Y_j$. It is said to be unbiased if its expectation equals EY for every choice of the parameter. Best linear unbiased prediction is widely used since it can be applied in various contexts. A best linear unbiased predictor (BLUP) is defined as the linear unbiased predictor of Y which minimizes (standardized) mean squared error. The mathematical treatment is presented in, e.g., Goldberger [405] or Christensen [264]. In location–scale families [see (11.1)] with unknown parameter $\boldsymbol{\theta} = (\mu, \vartheta)'$, the BLUP is given by

$$\widehat{\pi}_{\mathsf{LU}}(\mathbf{Y}) = \widehat{\mu}_{\mathsf{LU}} + \alpha_Y \widehat{\vartheta}_{\mathsf{LU}} + \boldsymbol{\varpi}' V^{-1}(\mathbf{Y} - \widehat{\mu}_{\mathsf{LU}} \mathbb{1} - \widehat{\vartheta}_{\mathsf{LU}} \boldsymbol{\alpha}), \tag{16.2}$$

where $\alpha_Y = EY$, $\boldsymbol{\alpha} = E\mathbf{Y}$, $V = \mathsf{Cov}(\mathbf{Y})$, $\boldsymbol{\varpi} = \mathsf{Cov}(Y, \mathbf{Y})$, and $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$ are the BLUEs of μ and ϑ, respectively. The corresponding best linear equivariant predictor (BLEP) has a similar representation, in which the BLUEs of the parameters are replaced by the BLEEs (see Balakrishnan et al. [139]):

$$\widehat{\pi}_{\mathsf{LE}}(\mathbf{Y}) = \widehat{\mu}_{\mathsf{LE}} + \alpha_Y \widehat{\vartheta}_{\mathsf{LE}} + \boldsymbol{\varpi}' V^{-1}(\mathbf{Y} - \widehat{\mu}_{\mathsf{LE}} \mathbb{1} - \widehat{\vartheta}_{\mathsf{LE}} \boldsymbol{\alpha}).$$

An alternative concept proposed by Kaminsky and Rhodin [496] is maximum likelihood prediction. Given the informative sample $\mathbf{y}$, the so-called predictive likelihood function (PLF) $\mathsf{L}(\boldsymbol{\theta}, y; \mathbf{y})$ is considered and maximized simultaneously with regard to the observation and the parameter $\boldsymbol{\theta}$. A solution $\widehat{\pi}_{\mathsf{ML}}(\mathbf{y})$ of the maximization problem

$$\sup_{(y,\boldsymbol{\theta})} \mathsf{L}(\theta, y; \mathbf{y})$$

defines the maximum likelihood predictor (MLP) $\widehat{\pi}_{\mathsf{ML}}(\mathbf{Y})$. The solution for $\boldsymbol{\theta}$ is called the predictive maximum likelihood estimator of $\boldsymbol{\theta}$. Kaminsky and Rhodin [496] applied this approach to predict an order statistic $X_{s:n}$ given a sample $X_{1:n}, \ldots, X_{r:n}$ with $1 \le r < s \le n$. Since explicit expressions of the MLP $\widehat{\pi}_{\mathsf{ML}}(\mathbf{y})$ will only be available in exceptional cases, Basak and Balakrishnan [175] used a Taylor approximation of order one to linearize the likelihood equations generated by the PLF. The resulting solution $\widehat{\pi}_{\mathsf{AM}}(\mathbf{y})$ is called approximate maximum likelihood predictor (AMLP).

The median unbiased predictor (MUP) $\widehat{\pi}_{\mathsf{MU}}(\mathbf{Y})$ is defined via the generalized median condition

$$P_{\boldsymbol{\theta}}(\widehat{\pi}_{\mathsf{MU}}(\mathbf{Y}) \le Y) = P_{\boldsymbol{\theta}}(\widehat{\pi}_{\mathsf{MU}}(\mathbf{Y}) \ge Y). \tag{16.3}$$

This idea was employed by Takada [834] for Type-II censored samples. He showed that the MUP leads to a smaller value of Pitman's measure of closeness than the BLUP. The median $\widehat{\pi}_{\mathsf{CM}}(\mathbf{Y})$ of the conditional distribution $f^{Y|\mathbf{Y}}$ is called conditional median predictor (CMP). Since it obviously satisfies (16.3), it is also a MUP.

16.2 Prediction of Failure Times of Censored Units

Balakrishnan and Rao [115] considered the problem of predicting the lifetime of an item censored in the last step of the progressive censoring procedure. Here, it has to be assumed that, at the termination time $X_{m:m:n}$ of the experiment, $R_m \geq 1$ observations are left (see also Balakrishnan and Aggarwala [86, Chap. 8]). This approach has been extended by Basak et al. [178] to predict lifetimes of items progressively censored in the lifetime experiment at some stage of the censoring procedure. This approach is based on a distributional result previously presented in Sect. 9.1.2. Interpreting the observed progressively Type-II censored order statistics $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ as observed information and the random vector $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_m)$, where $\mathbf{W}_j = (W_{j1}, \ldots, W_{jR_j})$ denotes those random variables corresponding to items withdrawn in the jth step of the progressive censoring procedure, as missing data, we get the following representation of the conditional density function [see (9.7)]:

$$f^{\mathbf{W}|\mathbf{X}^{\mathscr{R}}}(\mathbf{w}|\mathbf{x}) = \prod_{r=1}^{m}\prod_{k=1}^{R_r} \frac{f(w_{rk})}{1-F(x_r)} = \prod_{r=1}^{m}\prod_{k=1}^{R_r} f^{W_{rk}|X_{r:m:n}}(w_{rk}|x_r),$$
$$\min \mathbf{w}_r > x_r,\ 1 \leq r \leq m. \quad (16.4)$$

This shows that $\mathbf{W}_j$, $1 \leq j \leq m$, are conditionally independent, given $\mathbf{X}^{\mathscr{R}}$. If we are interested in predicting the ordered lifetimes of the items censored in the jth step of the progressive censoring procedure, the order statistics $W_{j,1:R_j}, \ldots, W_{j,R_j:R_j}$ have to be predicted. Integrating (16.4), we get

$$f^{\mathbf{W}_j|\mathbf{X}^{\mathscr{R}}}(\mathbf{w}_j|\mathbf{x}) = \prod_{k=1}^{R_j} f^{W_{jk}|X_{j:m:n}}(w_{jk}|x_j),$$

showing that the distribution of $\mathbf{W}_j$, given $\mathbf{X}^{\mathscr{R}}$, depends only on the progressively Type-II censored order statistic $X_{j:m:n}$. Thus, this applies also to the distribution of the order statistics $\mathbf{W}_j^{\leq} = (W_{j,1:R_j}, \ldots, W_{j,R_j:R_j})$ yielding the joint density function

$$f^{\mathbf{W}_j^{\leq}|\mathbf{X}^{\mathscr{R}}}(\mathbf{w}_j|\mathbf{x}) = R_j! \prod_{k=1}^{R_j} f^{W_{jk}|X_{j:m:n}}(w_{jk}|x_j), \quad w_{j1} \leq \cdots \leq w_{jR_j}.$$

If we are interested in the rth order statistic of the lifetimes withdrawn in the jth censoring step only, the respective density function is

$$f^{W_{j,r:R_j}|\mathbf{X}^{\mathscr{R}}}(w|\mathbf{x}) = f^{W_{j,r:R_j}|X_{j:m:n}}(w|x_j)$$
$$= \frac{R_j!}{(r-1)!(R_j-r)!} \frac{[F(w)-F(x_j)]^{r-1}[1-F(w)]^{R_j-r}}{[1-F(x_j)]^{R_j}} f(w), \quad (16.5)$$

where $w > x_r$.

For maximum likelihood prediction, we need the PLF $\mathsf{L}(\boldsymbol{\theta}, w; \mathbf{x})$, where w is the (future) realization of $W_{j,r:R_j}$. From (16.5), we get

$$\begin{aligned}\mathsf{L}(\boldsymbol{\theta}, w; \mathbf{x}) &= f_{\boldsymbol{\theta}}^{W_{j,r:R_j}|X_{j:m:n}}(w|x_j) \cdot f_{\boldsymbol{\theta}}^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}) \\ &= \frac{R_j!}{(r-1)!(R_j-r)!} \frac{[F_{\boldsymbol{\theta}}(w) - F_{\boldsymbol{\theta}}(x_j)]^{r-1}[1 - F_{\boldsymbol{\theta}}(w)]^{R_j-r}}{[1 - F_{\boldsymbol{\theta}}(x_j)]^{R_j}} f_{\boldsymbol{\theta}}(w) \\ &\quad \times \prod_{i=1}^{m} \gamma_i! f_{\boldsymbol{\theta}}(x_i)(1 - F_{\boldsymbol{\theta}}(x_i))^{R_i}. \end{aligned} \tag{16.6}$$

Assuming that $\boldsymbol{\theta}$ is known, Basak et al. [178] have established weak conditions guaranteeing the existence of a unique predictor by studying the corresponding likelihood equation. But, in this case, (16.6) can be seen as

$$\mathsf{L}(\boldsymbol{\theta}, w; \mathbf{x}) \propto \frac{R_j!}{(r-1)!(R_j-r)!} \frac{[F_{\boldsymbol{\theta}}(w) - F_{\boldsymbol{\theta}}(x_j)]^{r-1}[1 - F_{\boldsymbol{\theta}}(w)]^{R_j-r}}{[1 - F_{\boldsymbol{\theta}}(x_j)]^{R_j}} f_{\boldsymbol{\theta}}(w),$$

which is the density function of the rth order statistic in a sample of size R_j. Therefore, uniqueness of the predictor is directly connected to unimodality of $F_{\boldsymbol{\theta}}$ (see Sect. 2.7). Applying results of Alam [31] (see also Barlow and Proschan [167], Huang and Ghosh [460, 461], Dharmadhikari and Joag-dev [339], and Cramer [286]), we get simple conditions ensuring unimodality of $F_{\boldsymbol{\theta}}$. Namely, convexity of $1/f_{\boldsymbol{\theta}}$ yields the desired result. Moreover, it follows that log-concavity of the density function is sufficient, too. This includes, e.g., uniform, exponential, normal, Weibull, gamma, and Cauchy distributions.

Throughout, we assume a location or a scale model as introduced in (11.1a) and (11.1b). The location–scale model (11.1c) can be treated similarly with some additional mathematical difficulties. Most of the results are taken from Basak et al. [178] and Basak and Balakrishnan [175].

16.2.1 Exponential Distribution

In this section, a scale family of exponential distributions $\mathsf{Exp}(\vartheta)$ is considered for a sample $\mathbf{Z}^{\mathscr{R}} = (Z_{1:m:n}, \ldots, Z_{m:m:n})$. We present the predictors for a censored observation $W_{j,r:R_j}$ provided that $R_j > 0$. A crucial tool in the derivations is the conditional density function given in (16.5). It is given by

$$\begin{aligned} f^{W_{j,r:R_j}|Z_{j:m:n}}(w|z_j) \\ &= \frac{R_j!}{(r-1)!(R_j-r)!}[1 - \mathrm{e}^{-(w-z_j)/\vartheta}]^{r-1}\mathrm{e}^{-(R_j-r+1))(w-z_j)/\vartheta}. \end{aligned}$$

Then, $W_{j,r:R_j} - Z_{j:m:n}$ has the distribution of $\vartheta Z^*_{r:R_j}$, where $Z^*_{r:R_j}$ denotes the rth order statistic in a sample of size R_j from a standard exponential distribution. This shows that the distribution of $W_{j,r:R_j} - Z_{j:m:n}$ is independent of $Z_{j:m:n}$.

Scale Parameter $\vartheta > 0$ Known

First, we assume that the scale parameter is known. The BUP of $W_{j,r:R_j}$ is given by

$$\begin{aligned}\widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}}) &= E_\vartheta(W_{j,r:R_j}|Z_{j:m:n}) = Z_{j:m:n} + E_\vartheta(W_{j,r:R_j} - Z_{j:m:n}|Z_{j:m:n}) \\ &= Z_{j:m:n} + \vartheta E Z^*_{r:R_j} = Z_{j:m:n} + \vartheta \sum_{k=1}^{r} \frac{1}{R_j - k + 1}. \end{aligned} \tag{16.7}$$

Since this predictor is linear, it equals the BLUP of $W_{j,r:R_j}$. The mean squared predictive error (MSPE) is directly calculated from the above property as

$$\begin{aligned}\mathsf{MSPE}_\vartheta(\widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})) &= E_\vartheta\Big(W_{j,r:R_j} - \widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})\Big)^2 \\ &= E_\vartheta\Big(W_{j,r:R_j} - Z_{j:m:n} - \vartheta\sum_{k=1}^{r}\frac{1}{R_j - k + 1}\Big)^2 \\ &= \mathsf{Var}_\vartheta\left(W_{j,r:R_j} - Z_{j:m:n}\right) = \vartheta^2\,\mathsf{Var}(Z^*_{r:R_j}) \\ &= \vartheta^2\sum_{k=1}^{r}\frac{1}{(R_j - k + 1)^2}.\end{aligned}$$

For maximum likelihood prediction, we consider the PLF given in (16.6). For a one-parameter exponential distribution, this objective can be interpreted as a likelihood function from two-parameter exponential distribution with known scale parameter ϑ and unknown location parameter w. The sample is given by a single order statistic. The resulting maximum results as in Kambo [493] and the corresponding MLP are given by

$$\widehat{\pi}_{\mathsf{ML}}(\mathbf{Y}) = Z_{j:m:n} + \vartheta\log\Big(\frac{R_j}{R_j - r + 1}\Big). \tag{16.8}$$

Notice that the PLF is decreasing for $r = 1$ leading to the predictor $\widehat{\pi}_{\mathsf{ML}}(\mathbf{Y}) = Z_{j:m:n}$ of the first failure time of the removed items. This predictor is obviously biased with bias

$$E_\vartheta(\widehat{\pi}_{\mathsf{ML}}(\mathbf{Y}) - W_{j,r:R_j}) = \vartheta\Big[\log\Big(\frac{R_j}{R_j - r + 1}\Big) - \sum_{k=1}^{r}\frac{1}{R_j - k + 1}\Big].$$

The MSPE is given by

$$\mathsf{MSPE}_\vartheta(\widehat{\pi}_{\mathsf{ML}}(\mathbf{Z}^{\mathscr{R}})) = \vartheta^2\Big[\sum_{k=1}^{r}\frac{1}{(R_j-k+1)^2} + \Big(\log\Big(\frac{R_j}{R_j-r+1}\Big) - \sum_{k=1}^{r}\frac{1}{R_j-k+1}\Big)^2\Big].$$

Basak et al. [178] showed that the bias is negative and that, for $R_j - r \to \infty$, the MLP is an asymptotically unbiased and consistent predictor.

Using the results of Takada [834], Basak et al. [178] obtained that the CMP $\widehat{\pi}_{\mathsf{CM}}(\mathbf{Z}^{\mathscr{R}})$ and the MUP $\widehat{\pi}_{\mathsf{MU}}(\mathbf{Z}^{\mathscr{R}})$ are identical in the exponential case with known scale parameter. They are given by

$$\widehat{\pi}_{\mathsf{CM}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + \vartheta \operatorname{med}\big(F^{Z^*_{r:R_j}}\big),$$

where $Z^*_{r:R_j}$ denotes the rth order statistic in a sample of size R_j from the standard exponential distribution.

Scale Parameter $\vartheta > 0$ Unknown

For an unknown scale parameter, we consider first best unbiased prediction. Replacing ϑ in (16.7) by its maximum likelihood estimator $\widehat{\vartheta}^*_{\mathsf{MLE}}$ [see (12.4)], we get the predictor

$$\widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + \widehat{\vartheta}^*_{\mathsf{MLE}}\sum_{k=1}^{r}\frac{1}{R_j-k+1}. \tag{16.9}$$

We show that $\widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})$ satisfies condition (16.1) for every unbiased estimator $\varepsilon(\mathbf{Z}^{\mathscr{R}})$ of zero and, thus, is the BUP of $W_{j,r:R_j}$. Writing $c = \sum_{k=1}^{r}\frac{1}{R_j-k+1}$, we get

$$\begin{aligned}
&E_\vartheta\big([W_{j,r:R_j} - \widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})]\varepsilon(\mathbf{Z}^{\mathscr{R}})\big)\\
&= E_\vartheta\big([W_{j,r:R_j} - Z_{j:m:n} - c\widehat{\vartheta}^*_{\mathsf{MLE}}]\varepsilon(\mathbf{Z}^{\mathscr{R}})\big)\\
&= E_\vartheta\big([W_{j,r:R_j} - Z_{j:m:n}]\varepsilon(\mathbf{Z}^{\mathscr{R}})\big) - cE_\vartheta\big(\widehat{\vartheta}^*_{\mathsf{MLE}}\varepsilon(\mathbf{Z}^{\mathscr{R}})\big)\\
&= E_\vartheta\big([W_{j,r:R_j} - Z_{j:m:n} - \vartheta c]\varepsilon(\mathbf{Z}^{\mathscr{R}})\big) + c\vartheta E_\vartheta\varepsilon(\mathbf{Z}^{\mathscr{R}}) - cE_\vartheta\big(\widehat{\vartheta}^*_{\mathsf{MLE}}\varepsilon(\mathbf{Z}^{\mathscr{R}})\big)\\
&= -cE_\vartheta\big(\widehat{\vartheta}^*_{\mathsf{MLE}}\varepsilon(\mathbf{Z}^{\mathscr{R}})\big).
\end{aligned}$$

Here, we have used the fact that, from (16.1), $E_\vartheta\big([W_{j,r:R_j} - Z_{j:m:n} - \vartheta c]\varepsilon(\mathbf{Z}^{\mathscr{R}})\big) = 0$ because $EZ_{j:m:n} + \vartheta c$ is the BUP of $W_{j,r:R_j}$ for a known scale parameter ϑ.

From Theorem 12.1.1, we know that the MLE $\widehat{\vartheta}^*_{\mathsf{MLE}}$ forms a complete sufficient statistic and that it is the UMVUE of ϑ (see also Cramer and Kamps [299]). Hence, we conclude from Zacks [937, Theorem 3.3.1] that

$$E_\vartheta\big(\widehat{\vartheta}^*_{\mathsf{MLE}}\varepsilon(\mathbf{Z}^{\mathscr{R}})\big) = 0 \text{ for every unbiased estimator } \varepsilon(\mathbf{Z}^{\mathscr{R}}) \text{ of zero.} \tag{16.10}$$

This proves the desired result. Notice that $\widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})$ is a linear predictor so that it equals the BLUP, too. Since $\widehat{\vartheta}^*_{\mathsf{MLE}}$ is also the BLUE of ϑ, the result for best unbiased linear prediction can be obtained also using the general result from Doganaksoy and Balakrishnan [342]: The BLUP can be constructed by replacing the parameter by its BLUE. The MSPE is given by ($c = \sum_{k=1}^r \frac{1}{R_j-k+1}$)

$$\begin{aligned}
\mathsf{MSPE}_\vartheta(\widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})) &= E_\vartheta\Big(W_{j,r:R_j} - \widehat{\pi}_{\mathsf{BU}}(\mathbf{Z}^{\mathscr{R}})\Big)^2 \\
&= E_\vartheta\Big(W_{j,r:R_j} - Z_{j:m:n} - c\widehat{\vartheta}^*_{\mathsf{MLE}}\Big)^2 \\
&= \mathsf{Var}_\vartheta\left(W_{j,r:R_j} - Z_{j:m:n}\right) + c^2\,\mathsf{Var}_\vartheta(\vartheta^*_{\mathsf{MLE}}) \\
&\quad + 2cE_\vartheta\Big((W_{j,r:R_j} - Z_{j:m:n} - c\vartheta)(\widehat{\vartheta}^*_{\mathsf{MLE}} - \vartheta)\Big) \\
&= \vartheta^2\,\mathsf{Var}(Z^*_{r:R_j}) + c^2\,\mathsf{Var}_\vartheta(\widehat{\vartheta}^*_{\mathsf{MLE}}) \\
&= \vartheta^2\Big[\sum_{k=1}^r \frac{1}{(R_j-k+1)^2} + \frac{1}{m}\Big(\sum_{k=1}^r \frac{1}{R_j-k+1}\Big)^2\Big].
\end{aligned}$$

We have used the fact that (16.10) implies

$$E_\vartheta\Big((W_{j,r:R_j} - Z_{j:m:n} - c\vartheta)(\widehat{\vartheta}^*_{\mathsf{MLE}} - \vartheta)\Big) = 0$$

and that $\mathsf{Var}_\vartheta(\widehat{\vartheta}^*_{\mathsf{MLE}}) = \vartheta^2/m$ (see Theorem 12.1.1).

Following Takada [832, Theorem 3], the best equivariant predictor of ϑ has a similar form as in (16.9):

$$\widehat{\pi}_{\mathsf{BE}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + c\widehat{\vartheta}^*_{\mathsf{MLE}}$$

with an appropriately chosen constant c. Obviously, this predictor is linear so that it must equal the BLEP. Therefore, we find the representation of the BEP from a general result of Balakrishnan et al. [139] which parallels that of Doganaksoy and Balakrishnan [342] for unbiased estimation. We can replace the unknown parameter by its BLEE. Since the BLEE of ϑ is given by

$$\widehat{\vartheta}_{\mathsf{LE}} = \frac{1}{m+1}\sum_{j=1}^m (R_j+1)Z_{j:m:n} = \Big(1 - \frac{1}{m+1}\Big)\widehat{\vartheta}^*_{\mathsf{MLE}}$$

[see (11.5)], the BLEP and BEP of $W_{j,r:R_j}$ have the representation

$$\widehat{\pi}_{\mathsf{BE}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + \Big(1 - \frac{1}{m+1}\Big)\widehat{\vartheta}^*_{\mathsf{MLE}} \sum_{k=1}^{r} \frac{1}{R_j - k + 1}.$$

The MLP is derived by considering the logarithm of the PLF. Up to an additive term, it is given by

$$-(m+1)\log\vartheta - \frac{T(w)}{\vartheta} + (r-1)\log\Big(1 - \mathrm{e}^{-(w-z_j)/\vartheta}\Big), \quad w \ge x_j,$$

where $T(w) = \sum_{i=1}^m (R_i+1)z_i + (R_j - r + 1)(w - z_j)$. Solving the maximization problem yields the predictor

$$\widehat{\pi}_{\mathsf{ML}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + \widehat{\vartheta}^{**} \log\Big(\frac{R_j}{R_j - r + 1}\Big)$$

which is (16.8), where the parameter ϑ is replaced by the biased estimator $\widehat{\vartheta}^{**} = \frac{1}{m+1}\sum_{i=1}^m (R_i+1)Z_{i:m:n}$. The MSPE is given by

$$\begin{aligned}
&\mathsf{MSPE}_\vartheta(\widehat{\pi}_{\mathsf{ML}}(\mathbf{Z}^{\mathscr{R}})) \\
&\quad = \vartheta^2\Big[\sum_{k=1}^{r}\frac{1}{(R_j-k+1)^2} + \Big\{\log\Big(\frac{R_j}{R_j-r+1}\Big) - \sum_{k=1}^{r}\frac{1}{R_j-k+1}\Big\}^2 \\
&\qquad\qquad \frac{\log\big(\frac{R_j}{R_j-r+1}\big)}{m+1}\Big\{2\sum_{k=1}^{r}\frac{1}{R_j-k+1} - \log\Big(\frac{R_j}{R_j-r+1}\Big)\Big\}\Big].
\end{aligned}$$

For median unbiased prediction, Basak et al. [178] obtained the predictor

$$\widehat{\pi}_{\mathsf{MU}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + \widehat{\vartheta}^*_{\mathsf{MLE}}\,\mathrm{med}\big(F^{T_{r:R_j}}\big),$$

where $\widehat{\vartheta}^*_{\mathsf{MLE}}$ denotes the MLE (and UMVUE) of ϑ and $T_{r:R_j} = (W_{j,r:R_j} - Z_{j:m:n})/\widehat{\vartheta}_{\mathsf{MLE}}$. Following arguments as in Lawless [569], Basak et al. [178] found the survival function

$$\begin{aligned}
&P(T_{r:R_j} \ge t) \\
&\quad = \frac{1}{\mathsf{B}(r, R_j - r + 1)}\sum_{k=0}^{r-1}\frac{(-1)^k\binom{r-1}{k}}{R_j - r + k + 1}\Big[1 + \frac{(R_j - r + k + 1)t}{m}\Big]^{-m}.
\end{aligned}$$

The CMP has the representation

$$\widehat{\pi}_{\mathsf{CM}}(\mathbf{Z}^{\mathscr{R}}) = Z_{j:m:n} + \widehat{\vartheta}^*_{\mathsf{MLE}}\,\mathrm{med}\big(F^{Z^*_{r:R_j}}\big).$$

Hence, it is the CMP for a known scale parameter with ϑ replaced by the MLE $\widehat{\vartheta}^*_{\mathsf{MLE}}$.

	BUP/BLUP	BLEP	MLP	MUP	CMP
$W_{3,1:3}$	3.98875	3.65222	0.96000	3.15300	3.05937
$W_{3,2:3}$	8.53187	7.69056	4.23481	7.53806	7.25811
$W_{3,3:3}$	17.61813	15.7672	9.83313	15.98540	15.30198
$W_{5,1:3}$	5.80875	5.47222	2.78000	4.97300	4.87937
$W_{5,2:3}$	10.35188	9.51056	6.05481	9.35806	9.07811
$W_{5,3:3}$	19.43813	17.5872	11.65313	17.80540	17.12198
$W_{8,1:5}$	9.16725	8.96533	7.35000	8.66580	8.60962
$W_{8,2:5}$	11.43881	10.98450	9.15226	10.92206	10.77189
$W_{8,3:5}$	14.46756	13.67672	11.47577	13.92729	13.64811
$W_{8,4:5}$	19.01069	17.71506	14.75057	18.36360	17.88066
$W_{8,5:5}$	28.09694	25.79172	20.34889	26.86161	25.92652

Table 16.1 Predicted values for censored observations $W_{j,r:R_j}$ in Data 1.1.5 taken partly from Basak et al. [178]

Example 16.2.1. Basak et al. [178] applied the results for the exponential distribution to Nelson's insulating fluid Data 1.1.5. The results for this approach are summarized in Table 16.1. The results for the BLUP are also included in Balakrishnan and Aggarwala [86, pp. 160–162]. Notice that the censoring scheme is given by $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$ so that progressive censoring takes place only at the third step (3 items), fifth step (3 items), and at the last step (5 items).

Remark 16.2.2. Prediction of progressively censored failure times with exponentiated exponential distributions has been considered by Madi and Raqab [627]. They illustrated their results by the insulating fluid Data 1.1.4. They also analyzed progressively Type-II censored data discussed by Pradhan and Kundu [727] which is given in Table 12.5. Using the data belonging to the censoring scheme $\mathscr{R} = (15, 5, 4, 0^{*9})$, Madi and Raqab [627] obtained Bayesian predictive values as well as predictive intervals for the censored failure times.

16.2.2 Extreme Value Distribution

For the extreme value distribution with density function

$$f_\vartheta(t) = \frac{1}{\vartheta} \mathrm{e}^{t/\vartheta} \mathrm{e}^{-\mathrm{e}^{t/\vartheta}}, \quad t \in \mathbb{R}, \vartheta > 0,$$

Basak et al. [178] discussed the preceding prediction problem. As for the exponential distribution, the cases of a known and unknown scale parameter have to be treated separately. $X_{1:m:n}, \ldots, X_{m:m:n}$ denotes the informative sample of progressively Type-II censored order statistics from this extreme value distribution.

Scale Parameter $\vartheta > 0$ Known

The BUP is calculated as expectation of the conditional density function (16.5) and has the representation

$$\widehat{\pi}_{\mathsf{BU}}(\mathbf{X}^{\mathscr{R}}) = X_{j:m:n} + \vartheta E\Big[\log\Big(1 + \mathrm{e}^{-X_{j:m:n}/\vartheta} Z^*_{r:R_j}\Big)\Big| X_{j:m:n}\Big],$$

where $Z^*_{r:R_j}$ denotes the rth order statistic in a sample of size R_j from the $\mathsf{Exp}(1)$-distribution. The BLUP has the representation

$$\widehat{\pi}_{\mathsf{LU}}(\mathbf{X}^{\mathscr{R}}) = \alpha_W \vartheta + \boldsymbol{\varpi} \Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - \vartheta\boldsymbol{\alpha}), \tag{16.11}$$

where

$$\alpha_W = EW_{j,r:R_j}, \boldsymbol{\alpha} = E\mathbf{X}^{\mathscr{R}}, \Sigma = \mathsf{Cov}(\mathbf{X}^{\mathscr{R}}), \text{ and } \boldsymbol{\varpi} = \mathsf{Cov}(W_{j,r:R_j}, \mathbf{X}^{\mathscr{R}}). \tag{16.12}$$

The moments have to be calculated numerically. A discussion using representation (2.25), due to Kamps and Cramer [503], is included in Basak et al. [178, Appendix A].

The MLP also has an explicit representation. Maximization of the PLF yields the predictor

$$\widehat{\pi}_{\mathsf{ML}}(\mathbf{X}^{\mathscr{R}}) = \vartheta \log\Big[\mathrm{e}^{X_{j:m:n}/\vartheta} + \log\Big(\frac{R_j}{R_j - r + 1}\Big)\Big]. \tag{16.13}$$

The CMP is given by

$$\widehat{\pi}_{\mathsf{CM}}(\mathbf{X}^{\mathscr{R}}) = \vartheta \log\Big[\mathrm{e}^{X_{j:m:n}/\vartheta} + \mathrm{med}\big(F^{Z^*_{r:R_j}}\big)\Big]. \tag{16.14}$$

Scale Parameter $\vartheta > 0$ Unknown

Using the result from Doganaksoy and Balakrishnan [342], the BLUP can be directly obtained from (16.11) by replacing the parameter ϑ by its BLUE as

$$\widehat{\vartheta}_{\mathsf{LU}} = \frac{\boldsymbol{\alpha}'\Sigma^{-1}}{\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{\alpha}}\mathbf{X}^{\mathscr{R}},$$

where $\boldsymbol{\alpha}$ and Σ are given in (16.12). The MSPE can be directly computed from the variance of the BLUE, i.e., from $\vartheta^2/(\boldsymbol{\alpha}'\Sigma^{-1}\boldsymbol{\alpha})$. The MLP can be deduced from (16.13) by replacing the parameter ϑ by the predictive MLE $\widehat{\vartheta}^{**}$ obtained from the equation

$$\widehat{\vartheta}^{**} = \frac{1}{m+1}\sum_{k=1}^{m}(R_k + 1)X_{k:m:n}\mathrm{e}^{X_{k:m:n}/\widehat{\vartheta}^{**}}.$$

	BLUP	MMLP	CMP
$W_{3,1:3}$	3.56217	0.96000	3.31520
$W_{3,2:3}$	8.68718	5.18736	8.76462
$W_{3,3:3}$	20.13858	15.27098	20.57151
$W_{5,1:3}$	7.88749	2.78000	5.36904
$W_{5,2:3}$	13.43723	7.73934	11.06764
$W_{5,3:3}$	25.46373	18.56257	23.13913
$W_{8,1:5}$	8.91500	7.35000	9.06435
$W_{8,2:5}$	9.70274	10.52387	12.11114
$W_{8,3:5}$	13.63622	15.01828	16.33133
$W_{8,4:5}$	19.94702	21.99980	22.81920
$W_{8,5:5}$	33.13133	35.37716	35.84896

Table 16.2 Predicted values for censored observations $W_{j,r:R_j}$ in Data 1.1.5 taken from Basak et al. [178]

Basak et al. [178] pointed out that this may cause some computational difficulties. Alternatively, they proposed the linearized predictive MLE as discussed by Thomas and Wilson [843]:

$$\widetilde{\vartheta} = \frac{1}{m+1}\sum_{k=1}^{m}(R_k+1)\kappa_k X_{k:m:n},$$

where $\kappa_k = \sum_{i=1}^{k}\frac{1}{\gamma_i}$, $1 \le k \le m$. The resulting modified MLP of $W_{j,r:R_j}$ is given by

$$\widehat{\pi}_{\mathsf{MM}}(\mathbf{X}^{\mathscr{R}}) = \widetilde{\vartheta}\log\Big[\mathrm{e}^{X_{j:m:n}/\widetilde{\vartheta}} + \log\Big(\frac{R_j}{R_j - r + 1}\Big)\Big].$$

Notice that, for $r = 1$, this predictor simplifies to $\widehat{\pi}_{\mathsf{MM}}(\mathbf{X}^{\mathscr{R}}) = X_{j:m:n}$.

Since the CMP is also difficult to obtain in this setting, Basak et al. [178] proposed a modified CMP using (16.14) and replacing the parameter ϑ by the corresponding BLUE $\widehat{\vartheta}_{\mathsf{LU}}$. The resulting predictor is given by

$$\widehat{\pi}_{\mathsf{MC}}(\mathbf{X}^{\mathscr{R}}) = \widehat{\vartheta}_{\mathsf{LU}}\log\Big[\mathrm{e}^{X_{j:m:n}/\widehat{\vartheta}_{\mathsf{LU}}} + \operatorname{med}\big(F^{Z^*_{r:R_j}}\big)\Big],$$

where $Z^*_{r:R_j}$ denotes the rth order statistic in a sample of size R_j from the standard exponential distribution.

Example 16.2.3. The preceding predictors are applied by Basak et al. [178] to the log-times of Nelson's insulating fluid Data 17.5. The resulting predicted values are transformed back to the original time scale. The results from this approach are given in Table 16.2 (see also Table 16.1).

16.2.3 Normal Distribution

The present prediction problem has been discussed by Basak and Balakrishnan [175] for a normal distribution $\mathsf{N}(\mu, 1)$ with location parameter $\mu \in \mathbb{R}$. As above, the cases of a known and unknown location parameter have to be handled separately.

Location Parameter μ Known

Given the sample $X_{1:m:n}, \ldots, X_{m:m:n}$, the BUP is calculated from the expectation of the conditional density function in (16.5) and has the integral representation

$$\widehat{\pi}_{\mathsf{BU}}(\mathbf{X}^{\mathscr{R}}) = X_{j:m:n} + \frac{R_j!}{(r-1)!(R_j-r)!}\frac{1}{(1-\Phi(Y_{j:m:n}))}$$
$$\times \int_0^\infty t\varphi(t+Y_{j:m:n})\Big[\Phi(t+Y_{j:m:n})-\Phi(Y_{j:m:n})\Big]^{r-1}\Big[1-\Phi(t+Y_{j:m:n})\Big]^{R_j-r}\,dt,$$

where $Y_{j:m:n} = X_{j:m:n} - \mu$, $1 \le j \le m$, and φ and Φ denote the density function and cumulative distribution function of a standard normal distribution, respectively. The integral has been evaluated numerically but, obviously, is not linear. The BLUP is given by the expression

$$\widehat{\pi}_{\mathsf{LU}}(\mathbf{X}^{\mathscr{R}}) = \mu + \alpha_W + \boldsymbol{\varpi}'\Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - \mu\mathbb{1} - \boldsymbol{\alpha}), \tag{16.15}$$

where α_W, $\boldsymbol{\alpha}$, $\boldsymbol{\varpi}$, and Σ are given in (16.12). The moments can be computed using (2.25). A detailed discussion is provided by Basak and Balakrishnan [175] who showed that this works quite efficiently. They have pointed out that only single and product moments of order statistics from a standard normal distribution are necessary to compute the above quantities. These moments can be taken from Tietjen et al. [846] up to sample size 50.

The MLP can be obtained by maximizing the PLF w.r.t. the variable w. In this case, the PLF is proportional to the density function (16.5), or to

$$h(w) = \varphi(w)(\Phi(w) - \Phi(y_j))^{r-1}[1 - \Phi(w)]^{R_j - r}, \quad w \ge y_j.$$

Notice that, according to Basak et al. [178], a unique MLP exists.

Location Parameter μ Unknown

For an unknown location parameter, the BLUP can be directly obtained from (16.15) by replacing μ by its BLUE $\widehat{\mu}_{\mathsf{LU}} = \mathbb{1}'\Sigma^{-1}(\mathbf{X}^{\mathscr{R}} - \boldsymbol{\alpha})/(\mathbb{1}'\Sigma^{-1}\mathbb{1})$.

Maximum likelihood prediction has also been discussed in Basak and Balakrishnan [175]. They proposed two predictors referred to as one- and two-stage MLPs. The one-stage MLP is obtained as solution of the predictive likelihood equations resulting from the derivatives of the log-PLF:

$$\begin{aligned} \sum_{k=1}^{m} y_k + \sum_{k=1,k\neq j}^{m} R_k h(y_k) + (r-1)h^*(y_j, w-\mu) = 0, \\ -w + \mu + (r-1)[1 - h^*(y_j, w-\mu)] - (R_j - r)h(w-\mu) = 0, \end{aligned} \tag{16.16}$$

where $h(t) = \Phi(t)/[1-\Phi(t)]$, $h^*(s,t) = \Phi(s)/[\Phi(t)-\Phi(s)]$, $t > s$, $y_k = x_k - \mu$, $1 \le k \le m$. The corresponding predictor of $W_{j,r:R_j}$ is defined as $\widehat{\pi}_{\mathsf{ML}_1}(\mathbf{X}^{\mathscr{R}}) = \max\{X_{j:m:n}, \breve{X}_1\}$, where $\breve{X}_1$ is defined via the solution of (16.16) in w.

For the two-stage approach, Basak and Balakrishnan [175] used the fact that the PLF can be written as a product of the functions (up to multiplicative constants)

$$\begin{aligned} \mathsf{L}_1(\mu; \mathbf{y}) &= \prod_{k=1}^{m} \{\phi(y_k)(1-\Phi(y_k)\}^{R_k}, \\ \mathsf{L}_2(w; \mu, \mathbf{y}) &= \phi(w)\frac{(\Phi(w)-\Phi(y_j))^{r-1}}{[1-\Phi(y_j)]^{R_j}}[1-\Phi(w)]^{R_j - r}, w \ge y_j. \end{aligned}$$

Notice that these correspond to the joint density function of $\mathbf{Y}^{\mathscr{R}}$ and the conditional density function of $W_{j,r:R_j}$, given $\mathbf{Y}^{\mathscr{R}} = \mathbf{y}$ [see (16.5) and (16.6)]. Here, L_1 depends only on μ, whereas L_2 depends on both parameters. Using this decomposition, Basak and Balakrishnan [175] defined a two-stage MLP based on the solution of the equations resulting from differentiating the logarithm of L_1 w.r.t. μ and the logarithm of L_2 w.r.t. w:

$$\begin{aligned} \sum_{k=1}^{m} y_k + \sum_{k=1}^{m} R_k h(y_k) = 0, \\ -w + \mu + (r-1)[1 - h^*(y_j, w-\mu)] - (R_j - r)h(w-\mu) = 0. \end{aligned} \tag{16.17}$$

The corresponding predictor of $W_{j,r:R_j}$ is defined as

$$\widehat{\pi}_{\mathsf{ML}_2}(\mathbf{X}^{\mathscr{R}}) = \max\{X_{j:m:n}, \breve{X}_2\},$$

where $\breve{X}_2$ is defined via the solution of (16.17) in w.

Since explicit solutions of (16.16) generally do not exist, Basak and Balakrishnan [175] proposed a modified MLP which is defined via the solutions of (16.16) and (16.17), where the quantities $h(X_{k:m:n} - \mu)$, $h(W_{j,r:R_j} - \mu)$, and $h^*(X_{j:m:n} - \mu, W_{j,r:R_j} - \mu)$ are replaced by its expected values. Since the population distribution is a $\mathsf{N}(\mu, 1)$-distribution, it follows that $Y_{k:m:n} = X_{k:m:n} - \mu$ and $W_{j,r:R_j} - \mu$ can be seen as the corresponding random variables based on standard normal distribution. Therefore, the expectations are free of the location parameter μ. Consequently, the preceding equations are linear in μ and w and so can be easily solved provided that these expected values are available. Basak and Balakrishnan [175] established explicit but rather complicated expressions of these expected values in terms of moments of order statistics from a standard normal distribution. For instance, the expectation of $h(Y_{k:m:n})$ is given by

$$Eh(Y_{k:m:n}) = \prod_{i=1}^{k} \gamma_i \sum_{i=1}^{k} \frac{a_{i,k}}{(\gamma_i - 1)\gamma_i} EY_{1:\gamma_i}, \quad 1 \le k \le m,$$

provided that $\gamma_m > 1$ which means that at least one unit is Type-II censored in the experiment. If $\gamma_m = 1$, the last term in the representation of $Eh(Y_{m:m:n})$ has to be replaced by

$$\Big(\prod_{i=1}^{m-1} \gamma_i\Big) a_{m,m} \int_{-\infty}^{\infty} \frac{\varphi^2(t)}{1 - \Phi(t)}\, dt = 0.9031972856 \cdot \prod_{i=1}^{m-1} \frac{\gamma_i}{\gamma_i - 1}.$$

Expressions for $Eh(W_{j,r:R_j} - \mu)$ and $Eh^*(X_{j:m:n} - \mu, W_{j,r:R_j} - \mu)$ are presented in Basak and Balakrishnan [175].

Introducing the notation $e_k = Eh(X_{k:m:n} - \mu), a_j = Eh(W_{j,r:R_j} - \mu)$, and $b_j = Eh^*(X_{j:m:n} - \mu, W_{j,r:R_j} - \mu)$, the solutions of the modified equations (16.16) are given by

$$\widehat{\mu}_1 = \frac{1}{m}\Big(\sum_{k=1}^{m} x_k + \sum_{k=1,k\neq j}^{m} R_k e_k + (r-1)b_j\Big),$$

$$\widehat{w}_1 = \widehat{\mu}_1 + (r-1)[1 - b_1] - (R_j - r)a_j.$$

For the modified equations (16.17), the resulting solutions are

$$\widehat{\mu}_2 = \frac{1}{m}\Big(\sum_{k=1}^{m} x_k + \sum_{k=1}^{m} R_k e_k\Big),$$

$$\widehat{w}_2 = \widehat{\mu}_2 + (r-1)[1 - b_1] - (R_j - r)a_j.$$

The corresponding predictors of $W_{j,r:R_j}$ are defined as the predictors $\widehat{\pi}_{\mathsf{MM}_i}(\mathbf{X}^{\mathscr{R}}) = \max\{X_{j:m:n}, \check{X}_i\}$, where $\check{X}_i$ is defined via $\widehat{w}_i$, $i = 1, 2$.

Finally, Basak and Balakrishnan [175] discussed approximate maximum likelihood prediction. Here, the functions $h(X_{k:m:n} - \mu)$, $h(W_{j,r:R_j} - \mu)$, and $h^*(X_{j:m:n} - \mu, W_{j,r:R_j} - \mu)$ are replaced by a Taylor series expansion of order 1 in (16.16) and (16.17), respectively. For details, we refer to Basak and Balakrishnan [175]. Further, these authors also addressed conditional mean predictors which need to be computed numerically.

5.4258	5.4324	5.4410	5.4510	5.4515	5.4634
5.4984	5.5045	5.5146	5.5338	5.5502	5.5771

Table 16.3 Progressively Type-II censored data generated by Basak and Balakrishnan [175] from data simulated from a normal distribution. The censoring scheme is given by $\mathscr{R} = (0^{*3}, 2, 0^{*2}, 2, 0^{*4}, 4)$, $m = 12, n = 20$

	BLUP	MLP_1	MLP_2	MMLP_1	MMLP_2	CMP
$W_{4,1:2}$	5.4992	5.4510	5.4510	5.4510	5.4510	5.4802
$W_{4,2:2}$	5.6230	5.4921	5.6121	5.4992	5.6513	5.5784
$W_{7,1:2}$	5.5107	5.4984	5.4984	5.4984	5.4984	5.1023
$W_{7,2:2}$	7.1781	5.5423	6.9256	5.5621	7.0156	6.2384
$W_{12,1:4}$	6.2814	5.5771	5.5771	5.5771	5.5771	5.9123
$W_{12,2:4}$	7.8310	5.5771	7.0562	5.7108	7.1121	6.0256
$W_{12,3:4}$	9.1251	7.2341	9.2681	7.2108	9.3785	9.0587
$W_{12,4:4}$	11.2542	9.1081	12.0581	9.0256	11.3515	11.0568

Table 16.4 Predicted values for censored observations $W_{j,r:R_j}$ for the data given in Table 16.3 taken from Basak and Balakrishnan [175]

Example 16.2.4. Basak and Balakrishnan [175] generated a progressively Type-II data set (see Table 16.3) from simulated log-normal data presented in Cohen and Whitten [277] (see also Cohen [266]). The data were selected from normal random samples given in Mahalanobis et al. [628]. The resulting predicted values are given in Table 16.4.

16.2.4 Pareto Distributions

In this section, a location–scale family of Pareto distributions with cumulative distribution function

$$F(t) = 1 - \left(1 + \frac{x-\mu}{\vartheta}\right)^{-\alpha}, \quad x \geq \mu,$$

is considered. The shape parameter α is supposed to be known. The standard member of this family is a Lomax(α)-distribution. The conditional density function $f^{W_{j,r:R_j}|\mathbf{X}^{\mathscr{R}}}$ can be taken from (16.5). It can be written as

$$f^{W_{j,r:R_j}|X_{j:m:n}}(w|x_j) = \frac{R_j!}{(r-1)!(R_j-r)!} H^{r-1}(w|x_j)(1-H(w|x_j))^{R_j-r} h(w|x_j), \quad w \geq x_j,$$

where

$$H(w|x_j) = 1 - \Big[1 + \frac{w - x_j}{\vartheta + x_j - \mu}\Big]^{-\alpha} \quad \text{and}$$

$$h(w|x_j) = \frac{\alpha}{\vartheta + x_j - \mu}\Big[1 + \frac{w - x_j}{\vartheta + x_j - \mu}\Big]^{-(\alpha+1)}.$$

Therefore, we can interpret $W_{j,r:R_j}|X_{j:m:n} = x_j$ as the rth order statistic in a sample of size R_j with population cumulative distribution function $H(\cdot|x_j)$. Notice that this is the cumulative distribution function of a Pareto distribution with location parameter x_j and scale parameter $\vartheta + x_j - \mu$.

Suppose $V_{r:R_j}$ denotes the rth order statistic from a Lomax distribution in a sample of size R_j. Then, according to Corollary 7.2.6, $EV_{r:R_j} = \prod_{i=1}^{r} \frac{\alpha(R_j-i+1)}{\alpha(R_j-i+1)-1} - 1$ provided that $\alpha(R_j - r + 1) > 1$. Given that μ and ϑ are known, the BUP of $W_{j,r:R_j}$ is given by

$$\begin{aligned}\widehat{\pi}_{\mathsf{BU}}(\mathbf{X}^{\mathscr{R}}) &= E(W_{j,r:R_j}|X_{j:m:n}) = X_{j:m:n} + (\vartheta + X_{j:m:n} - \mu)EV_{r:R_j}\\ &= \mu - \vartheta + (\vartheta + X_{j:m:n} - \mu)\prod_{i=1}^{r}\frac{\alpha(R_j - i + 1)}{\alpha(R_j - i + 1) - 1}.\end{aligned} \tag{16.18}$$

Since this predictor is linear, it is also the BLUP for $W_{j,r:R_j}$ provided that μ and ϑ are known. Representation (16.18) shows that $\widehat{\pi}_{\mathsf{BU}}(\mathbf{X}^{\mathscr{R}}) \geq X_{j:m:n}$ almost surely and that $\widehat{\pi}_{\mathsf{BU}}(\mathbf{X}^{\mathscr{R}})$ is increasing in $r \in \{1, \ldots, R_j\}$.

Raqab et al. [746] considered best linear unbiased and (approximate) maximum likelihood prediction for a location–scale family of Pareto distributions with known shape parameter $\alpha > 0$ and standard member $\mathsf{Lomax}(\alpha)$. They obtained explicit expressions for the BLUPs using the representations of the moments given in Theorem 7.2.5. They can also be obtained from the BLUEs given in Theorem 11.2.4 and using the result of Doganaksoy and Balakrishnan [342] [see (16.2)]:

$$\widehat{\pi}_{\mathsf{LU}}(\mathbf{X}^{\mathscr{R}}) = \widehat{\mu}_{\mathsf{LU}} + \alpha_W\widehat{\vartheta}_{\mathsf{LU}} + \boldsymbol{\varpi}'V^{-1}(\mathbf{X}^{\mathscr{R}} - \widehat{\mu}_{\mathsf{LU}}\mathbb{1} - \widehat{\vartheta}_{\mathsf{LU}}\boldsymbol{\alpha}),$$

where the expected values $\alpha_W = EW_{j,r:R_j}, \boldsymbol{\alpha} = E\mathbf{X}^{\mathscr{R}}, V = \mathsf{Cov}(\mathbf{X}^{\mathscr{R}})$, and $\boldsymbol{\varpi} = \mathsf{Cov}(W_{j,r:R_j}, \mathbf{X}^{\mathscr{R}})$ are calculated for the standard member of the location–scale family, i.e., the $\mathsf{Lomax}(\alpha)$-distribution. The representation of the BUP in (16.18) can be used to derive expressions for α_W and $\boldsymbol{\varpi}$. Suppose $\mu = 0$ and $\vartheta = 1$. Then, we get using (16.5)

$$\begin{aligned}\alpha_W = EW_{j,r:R_j} &= E\Big[E\big(W_{j,r:R_j}|X_{j:m:n}\big)\Big]\\ &= -1 + E(X_{j:m:n} + 1)\cdot\prod_{i=1}^{r}\frac{\alpha(R_j - i + 1)}{\alpha(R_j - i + 1) - 1}\\ &= \prod_{k=1}^{j}\frac{\alpha\gamma_k}{\alpha\gamma_k - 1}\prod_{i=1}^{r}\frac{\alpha(R_j - i + 1)}{\alpha(R_j - i + 1) - 1} - 1,\end{aligned}$$

$$\begin{aligned}\beta_{rk} &= E[X_{k:m:n}W_{j,r:R_j}] = E\Big[X_{k:m:n}E\big(W_{j,r:R_j}\big|\mathbf{X}^{\mathscr{R}}\big)\Big]\\ &= E\Big[X_{k:m:n}E\big(W_{j,r:R_j}\big|X_{j:m:n}\big)\Big]\\ &= E\big[X_{k:m:n}X_{j:m:n}\big]\prod_{i=1}^{r}\frac{\alpha(R_j-i+1)}{\alpha(R_j-i+1)-1}\\ &\qquad + EX_{k:m:n}\Big[\prod_{i=1}^{r}\frac{\alpha(R_j-i+1)}{\alpha(R_j-i+1)-1}-1\Big],\quad 1\le k\le m,\\ \varpi_k &= \mathsf{Cov}(X_{j:m:n},X_{k:m:n})\prod_{i=1}^{r}\frac{\alpha(R_j-i+1)}{\alpha(R_j-i+1)-1}.\end{aligned}$$

Then, we find by the definition of V that $(\boldsymbol{\varpi}'V^{-1})_k = \prod_{i=1}^{r}\frac{\alpha(R_j-i+1)}{\alpha(R_j-i+1)-1}$ for $k=j$ and zero otherwise. This leads to the following representation of the BLUP:

$$\widehat{\pi}_{\mathsf{LU}}(\mathbf{X}^{\mathscr{R}}) = X_{j:m:n} + \Big[\prod_{i=1}^{r}\frac{\alpha(R_j-i+1)}{\alpha(R_j-i+1)-1}-1\Big](\widehat{\vartheta}_{\mathsf{LU}}+X_{j:m:n}-\widehat{\mu}_{\mathsf{LU}}).$$

Notice that this is the BLUP for known μ and ϑ with the parameters replaced by the corresponding BLUEs [see (16.18)].

The BLEPs can be obtained in a similar fashion using the expressions for the BLEEs given in Burkschat [231] and the devolution result due to Balakrishnan et al. [139]. Notice that moments of progressively Type-II censored order statistics from a $\mathsf{Lomax}(\alpha)$-distribution exist only under certain conditions on the parameters and on α (see Theorem 7.2.5). This has to be taken into account. Raqab et al. [746] also established the MLP

$$\widehat{\pi}_{\mathsf{ML}}(\mathbf{X}^{\mathscr{R}}) = \widehat{\mu} + \widehat{\vartheta}\Big[\Big(\frac{R_j+1/\alpha}{R_j-r+1+1/\alpha}\Big)^{1/\alpha}\Big(1+\frac{X_{j:m:n}-\widehat{\mu}}{\widehat{\vartheta}}\Big)-1\Big] \quad (16.19)$$

of $W_{j,r:R_j}$, where $\widehat{\mu} = X_{1:m:n}$ and $\widehat{\vartheta}$ are the predictive MLEs. $\widehat{\vartheta}$ is obtained as the solution of the equation

$$\alpha\gamma_2 - 1 = \sum_{i=2}^{m}[\alpha(R_i+1)+1]\frac{\vartheta}{\vartheta+x_i-x_1} + \frac{\vartheta}{\vartheta+x_j-x_1}.$$

This equation is quite similar to the corresponding likelihood equation given in (12.22).

Since the right-hand side of this equation is strictly increasing in ϑ, the equation has at most one solution. Since the limit for $\vartheta\to 0$ is given by 0 and for $\vartheta\to\infty$ is given by $\alpha\gamma_2+m>0$, the equation has a unique solution iff $\gamma_2 > 1/\alpha$ (and $m\ge 2$)

	BLUP	MLP_1	AMLP
$W_{1,1:3}$	0.0280	0.0141	0.0141
$W_{1,2:3}$	0.0651	0.0872	0.0869
$W_{1,3:3}$	0.2113	0.3552	0.3918
$W_{3,1:1}$	1.9146	1.2200	1.2200
$W_{4,1:2}$	3.1855	2.7200	2.7200
$W_{4,2:2}$	4.6215	4.7770	4.5104

Table 16.5 Predicted values for the observations $W_{j,r:R_j}$ censored in the first three censoring steps for the data given in Table 11.2 taken from Raqab et al. [746]

[see also the comments following (12.22)]. Notice that this means that $X_{2:m:n}$ has a finite first moment (see Theorem 7.2.5). In order to get an explicit expression for the estimate of ϑ, Raqab et al. [746] used a linearization of the above equation and proposed an approximate MLP similar to (16.19), where the predictive MLE $\widehat{\vartheta}$ is replaced by the predictive AMLE.

Example 16.2.5. Raqab et al. [746] presented the predicted values for the censored data as given in Table 16.5

16.3 Prediction of Future Observations

Suppose we observe progressively Type-II censored order statistics $X_{1:m:n} \leq \cdots \leq X_{r:m:n}$ with $r < m$ and censoring scheme $\mathscr{R}$. We are interested in predicting the future observations $X_{r+1:m:n} \leq \cdots \leq X_{m:m:n}$.

16.3.1 Linear Prediction

Burkschat [231] derived BLUPs and BLEPs in the setting of generalized order statistics for generalized Pareto distributions as given in Definition A.1.11. The results are established in a multi-sample scenario. For brevity, we consider only the one-sample setting as introduced above. Let $\ell \in \{r+1, \ldots, m\}$. Then, Burkschat [231] obtained the following result.

Theorem 16.3.1. Let c_j, d_j be as defined in (11.9) and $F \in \mathscr{GPD}$ with $q \neq 0$. Suppose that, for the sample $X_{1:m:n}^{\mathscr{R}}, \ldots, X_{r:m:n}^{\mathscr{R}}$, the censoring scheme $\mathscr{R}$ satisfies $\gamma_r + 2q > 0$. Then, the BLUP of $X_{\ell:m:n}$ is given by

$$\widehat{\pi}_{\mathsf{LU}}(X_{\ell:m:n}) = \frac{c_\ell d_r}{c_r d_\ell} X_{r:m:n} + \Big(1 - \frac{c_\ell d_r}{c_r d_\ell}\Big)\Big(\widehat{\mu}_{\mathsf{LU}} + \frac{1}{q}\widehat{\vartheta}_{\mathsf{LU}}\Big),$$

where the BLUEs $\widehat{\mu}_{\mathsf{LU}}, \widehat{\vartheta}_{\mathsf{LU}}$ are given in Theorem 11.2.4. In the exponential case, the BLUP is given by

$$\widehat{\pi}_{\mathsf{LU}}(X_{\ell:m:n}) = X_{r:m:n} + \widehat{\vartheta}_{\mathsf{LU}} \sum_{j=r+1}^{\ell} \frac{1}{\gamma_j}.$$

As pointed out in Burkschat [231], the BLEPs are obtained from Theorem 16.3.1 by replacing the BLUEs with the BLEEs given in Burkschat [231]. Furthermore, MSPEs are provided. Burkschat [231] observed that, for reflected power distributions, the BLUP is a convex combination of the largest observation $X_{r:m:n}$ and the BLUE of the right endpoint of the support given by $\widehat{\mu}_{\mathsf{LU}} + \frac{1}{q}\widehat{\vartheta}_{\mathsf{LU}}$. A similar property has been reported by Kaminsky and Nelson [495] for order statistics.

16.3.2 Bayesian Prediction

Suppose the population cumulative distribution function is given by $F_{\boldsymbol{\theta}}$ with density function $f_{\boldsymbol{\theta}}$ and that the prior distribution is defined by the prior density function $\pi_{\mathbf{a}}$. $\pi_{\mathbf{a}}^*(\cdot|\mathbf{x})$ denotes the posterior density function, given $\mathbf{X}^{\mathscr{R}} = \mathbf{x}$ [see (15.1)]. In this framework, two different setups are considered:

- One-sample setup
  Given a sample $X_{1:m:n}, \ldots, X_{j:m:n}$ of progressively Type-II censored order statistics, we are interested in predicting a future outcome $X_{k:m:n}$, $k > j$, from the same sample. The procedure is based on the predictive distribution of $X_{k:m:n}$, given the observations $\mathbf{X}_j^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{j:m:n})$. Then, the density function of $X_{k:m:n}$, given $\mathbf{X}_j^{\mathscr{R}} = \mathbf{x}_j$, is given by

  $$f_{\boldsymbol{\theta}}^{X_{k:m:n}|\mathbf{X}_j^{\mathscr{R}}=\mathbf{x}_j}(x_k) = f_{\boldsymbol{\theta}}^{X_{k:m:n}|X_{j:m:n}=x_j}(x_k), \quad x_1, \ldots, x_j, x_k \in \mathbb{R}.$$

  Due to the Markov property of progressively Type-II censored order statistics, this conditional density function depends only on the largest observation x_j. The predictive density function of $X_{k:m:n}$, given $\mathbf{X}_j^{\mathscr{R}} = \mathbf{x}_j$, is defined via

  $$f_k(t|\mathbf{x}_j) = \int f_{\boldsymbol{\theta}}^{X_{k:m:n}|X_{j:m:n}=x_j}(t)\pi_{\mathbf{a}}^*(\boldsymbol{\theta}|\mathbf{x}_j)d\boldsymbol{\theta}, \quad \mathbf{x}_j \in \mathbb{R}^j, t \in \mathbb{R},$$

  where $\pi_{\mathbf{a}}^*(\boldsymbol{\theta}|\mathbf{x}_j)$ denotes the posterior density function. In fact, from Theorem 2.5.2, the conditional density function $f_{\boldsymbol{\theta}}^{X_{k:m:n}|X_{j:m:n}=x_j}$ can be written as

  $$f_{\boldsymbol{\theta}}^{X_{k:m:n}|X_{j:m:n}=x_j}(t) = \Big(\prod_{i=j+1}^{k} \gamma_i\Big) \sum_{i=j+1}^{k} a_{i,k}^{(j)}(1 - G_{\boldsymbol{\theta},x_j}(t))^{\gamma_i-1} g_{\boldsymbol{\theta},x_j}(t),$$

where $a_{i,k}^{(j)} = \prod_{\substack{\nu=j+1\\ \nu\neq i}}^{k} \frac{1}{\gamma_\nu - \gamma_i}$, $G_{\boldsymbol\theta,x_j}(t) = 1 - [1 - F_{\boldsymbol\theta}(t)]/[1 - F_{\boldsymbol\theta}(x_j)]$ is a left-truncated cumulative distribution function, and $g_{\boldsymbol\theta,x_j}$ denotes the corresponding density function.

- Two-sample setup
  Given a sample $X_{1:m:n}, \ldots, X_{m:m:n}$ of progressively Type-II censored order statistics (the so-called informative sample), the goal of this method is to predict outcomes of an independent future sample $Y_{1:N} \le \cdots \le Y_{N:N}$ from the same population. The procedure is based on the predictive distribution of $Y_{k:N}$, given the informative sample $\mathbf{X}^{\mathscr{R}}$. Then, the density function of $Y_{k:N}$ is given by

$$f_{\boldsymbol\theta}^{Y_{k:N}}(y) = k\binom{N}{k} f_{\boldsymbol\theta}(y)(1 - F_{\boldsymbol\theta}(y))^{k-1} F_{\boldsymbol\theta}^{N-k}(y), \quad y \in \mathbb{R}.$$

For a posterior density function $\pi_{\mathbf{a}}^*(\cdot|\mathbf{x})$, the predictive density function of $Y_{k:N}$ given $\mathbf{X}^{\mathscr{R}} = \mathbf{x}$ is defined via

$$f_k(y|\mathbf{x}) = \int f_{\boldsymbol\theta}^{Y_{k:N}}(y)\pi_{\mathbf{a}}^*(\boldsymbol\theta)d\boldsymbol\theta, \quad y \in \mathbb{R}.$$

Under squared-error loss, the Bayes predictive estimator of $X_{k:m:n}$ (or $Y_{k:N}$) is defined as the expected value of the predictive distribution, i.e.,

$$\widehat{\pi}_{\mathsf{BA}}(\mathbf{X}^{\mathscr{R}}) = \int y f_k(y|\mathbf{X}^{\mathscr{R}})dy.$$

Under absolute error loss, the median $\widehat{\pi}_{\mathsf{BAM}}(\mathbf{X}^{\mathscr{R}})$ of the predictive distribution is the Bayes predictor of $X_{k:m:n}$ and $Y_{k:N}$, respectively.

Bayesian Prediction: One-Sample Case

Schenk et al. [783] have considered a multiply censored sample from an exponential population in terms of sequential order statistics with different model parameters. For illustration, we present here only the case of a progressively Type-II right censored sample. In the one-parameter setting, i.e., an exponential distribution $\mathsf{Exp}(\vartheta)$ and an inverse gamma prior $\pi_{a,b}(\vartheta)$ as in (15.3) is assumed, they found the posterior density function

$$\pi_{a,b}^*(\vartheta|\mathbf{x}) = \frac{(t_j + b)^{j+a}}{\Gamma(j+a)}\vartheta^{-(j+a+1)}\mathrm{e}^{-(t_j+b)/\vartheta}, \quad \vartheta > 0, \tag{16.20}$$

where $t_j = \sum_{\ell=1}^{j} \gamma_\ell(x_\ell - x_{\ell-1})$, $x_0 = 0$. Then, the predictive density function results as

$$f_k(t|\mathbf{x}_j) = \Big(\prod_{i=j+1}^{k}\gamma_i\Big)\sum_{i=j+1}^{k} a_{i,k}^{(j)} \frac{(j+a)(t_j+b)^{j+a}}{(t_j + b + \gamma_i(t_j - x_j))^{j+a+1}}.$$

Suppose $j+a-1>0$. Then, for squared-error loss, integration w.r.t. $t \geq x_j$ yields the Bayes predictor

$$\widehat{\pi}_{\mathsf{BA}}(\mathbf{X}^{\mathscr{R}}) = X_{j:m:n} + \frac{T_j+b}{j+a-1}\sum_{i=j+1}^{k}\frac{1}{\gamma_i} = X_{j:m:n} + \widehat{\vartheta}_{\mathsf{B}}\sum_{i=j+1}^{k}\frac{1}{\gamma_i},$$

where $\widehat{\vartheta}_{\mathsf{B}}$ denotes the Bayes estimator of ϑ in this setting [see (15.4)]. For $j+a-2>0$, the posterior variance is given by

$$\mathsf{Var}(X_{k:m:n}|\mathbf{X}_j^{\mathscr{R}} = \mathbf{x}_j) = \widehat{\vartheta}_{\mathsf{B}}^2\Big[\frac{1}{j+a-2}\Big(\sum_{i=j+1}^{k}\frac{1}{\gamma_i}\Big)^2 + \sum_{i=j+1}^{k}\frac{1}{\gamma_i^2}\Big]$$

(see Schenk et al. [783]). The predictive survival function is given by

$$\overline{F}_k(t|\mathbf{x}_j) = \Big(\prod_{i=j+1}^{k}\gamma_i\Big)\sum_{i=j+1}^{k}\frac{a_{i,k}^{(j)}}{\gamma_i}\frac{(T_j+b)^{j+a}}{(T_j+b+\gamma_i(t-x_j))^{j+a}}.$$

It can be used to compute a $(1-\alpha)$-HPD prediction interval $[\widehat{\ell}_k, \widehat{u}_k]$ for $X_{k:m:n}$ solving the equations

$$\overline{F}_k(\widehat{\ell}_k|\mathbf{x}_j) - \overline{F}_k(\widehat{u}_k|\mathbf{x}_j) = 1-\alpha, \qquad f_k(\widehat{\ell}_k|\mathbf{x}_j) = f_k(\widehat{u}_k|\mathbf{x}_j).$$

The corresponding formulas for conventionally Type-II censored data have been established by Fernández [365].

A similar approach can be applied in the setting of a Weibull model with known shape parameter β. In this setting, we have to replace the observations x_j by the transformed observations x_j^β resulting in quite similar expressions. This comment applies to, e.g., Rayleigh distributions, where $\beta = 2$. Prediction intervals for a general class of distributions as given in (15.7), including exponential, Weibull, and Pareto distributions, are presented in Abdel-Aty et al. [2]. Their results are based on priors previously used by AL-Hussaini and Ahmad [28]. The scenario is further discussed in Mohie El-Din and Shafay [652].

Bayesian Prediction: Two-Sample Case

Given the informative sample $X_{1:m:n}, \ldots, X_{m:m:n}$ of progressively Type-II censored order statistics, the goal of this method is to predict outcomes of an independent future sample $Y_{1:N} \leq \cdots \leq Y_{N:N}$ from the same population. The procedure is based on the predictive distribution of $Y_{k:N}$, given the informative sample $\mathbf{X}^{\mathscr{R}}$. Suppose the population cumulative distribution function is given by $F_{\boldsymbol{\theta}}$ with density function $f_{\boldsymbol{\theta}}$. Then, the density function of $Y_{k:N}$ is given by

$$f_{\boldsymbol{\theta}}^{Y_{k:N}}(y) = \frac{N!}{(k-1)!(N-k)!} f_{\boldsymbol{\theta}}(y) F_{\boldsymbol{\theta}}^{k-1}(y)(1-F_{\boldsymbol{\theta}}(y))^{N-k}, \quad y \in \mathbb{R}.$$

Given a prior density function $\pi_{\mathbf{a}}$, the predictive density function of $Y_{k:N}$, given $\mathbf{X}^{\mathscr{R}} = \mathbf{x}$, is defined via

$$f_k(y|\mathbf{x}) = \int f_{\boldsymbol{\theta}}^{Y_{k:N}}(y)\pi_{\mathbf{a}}^*(\boldsymbol{\theta})d\boldsymbol{\theta}, \quad y \in \mathbb{R},$$

where $\pi_{\mathbf{a}}^*$ denotes the posterior density function.

Apparently, Bayesian prediction of future order statistics from an exponential distribution has not been considered. For completeness, we present the results which are easy to derive. Using the gamma prior given in (15.3), the posterior density function (16.20) with $j = m$ results as in the one-sample prediction problem with $t_m = \sum_{i=1}^{m} \gamma_i(x_i - x_{i-1})$, $x_0 = 0$. Then, the density function of $Y_{k:N}$ is given by

$$f_{\vartheta}^{Y_{k:N}}(y) = k\binom{N}{k}\frac{1}{\vartheta}\mathrm{e}^{-(N-k+1)y/\vartheta}\left(1-\mathrm{e}^{-y/\vartheta}\right)^{k-1}, \quad y > 0,$$

leading to the predictive density function

$$f_k(y|\mathbf{x}) = k\binom{N}{k}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}\frac{(m+a)(t_m+b)^{m+a}}{(t_m+b+(N-k+j+1)y)^{m+a+1}}.$$

Supposing $m + a > 1$, we get for squared-error loss the Bayes predictor of $Y_{k:N}$ as

$$\widehat{\pi}_{\mathsf{BA}}(\mathbf{X}^{\mathscr{R}}) = \left[k\binom{N}{k}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}(N-k+j+1)^{-2}\right]\widehat{\vartheta}_{\mathsf{B}},$$

where $\widehat{\vartheta}_{\mathsf{B}}$ is the Bayes estimator of ϑ [see (15.6)]. Prediction of k-records based on a progressively Type-II censored sample has been discussed by Ahmadi et al. [17].

Bayesian prediction issues of future order statistics for Rayleigh distributions have been addressed by Ali Mousa and Al-Sagheer [35], Wu et al. [915], and Kim and Han [529]. Assuming the prior in (15.8) (see Sect. 15.2), Wu et al. [915] obtained the Bayesian point predictor of $Y_{k:N}$ as

$$\widehat{\pi}_{\mathsf{BA}}(\mathbf{X}^{\mathscr{R}}) = \left[k\binom{N}{k}\sqrt{\frac{\pi}{2}}\sum_{j=0}^{k-1}(-1)^j\binom{k-1}{j}(N-k+j+1)^{-3/2}\right]\widehat{\vartheta}_{\mathsf{B}}, \qquad (16.21)$$

where $\widehat{\vartheta}_{\mathsf{B}} = \frac{T_m+b}{m+a-1}$ denotes the Bayes estimator of ϑ given in (15.9). The derivations are similar to those for the exponential distribution and therefore omitted. Kim and Han [529] obtained the same expression of the Bayes predictor as in (16.21) when the sample is generally progressively Type-II censored. But, the Bayes estimator $\widehat{\vartheta}_{\mathsf{B}}$ exhibits a much more complicated representation. Ali Mousa and Al-Sagheer [35] considered Bayesian interval prediction for a future progressively Type-II censored order statistic using a different prior. Fernández [367] discussed Bayes prediction based on a multiply censored sample of order statistics.

Remark 16.3.2. Bayesian prediction issues have also been addressed for other distributions. Ali Mousa [34] and Ali Mousa and Jaheen [37] discussed prediction intervals for Pareto distributions and Burr distributions, respectively [see also Mohie El-Din and Shafay [652] in terms of the family (15.7)].

Chapter 17
Statistical Intervals for Progressively Type-II Censored Data

Though considerable attention is usually paid to point estimation, in many practical situations an experimenter may also require uncertainty information such as those conveyed by confidence intervals, tolerance intervals, or prediction intervals for the life characteristics of interest. The primary motivation in constructing confidence intervals is that they provide a range of plausible values for the life parameter of interest based on the observed progressively censored data at a required level of confidence. The practical relevance of confidence intervals and their construction has been well demonstrated by Hahn and Meeker [426].

In Sects. 17.1 and 17.2, we discuss the derivation of exact (conditional) confidence intervals for distribution parameters, quantiles, and reliability. We present parametric and nonparametric approaches. Additionally, asymptotic confidence intervals are sketched in Sect. 17.3. Finally, this section is supplemented by results on prediction intervals (Sect. 17.4) and tolerance intervals (Sect. 17.5).

17.1 Exact Confidence Intervals

In this section, we present exact confidence intervals for various parametric families of distributions as well as a nonparametric approach. The results include confidence intervals for distribution parameters, quantiles, and other quantities like reliability.

17.1.1 Exponential Distribution

Suppose the progressively Type-II censored order statistics are based on one- or two-parameter exponential distribution. The construction of confidence intervals for the distribution parameters is based on the results presented in Theorems 12.1.1 and 12.1.4.

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_17,

Corollary 17.1.1. Let $\alpha \in (0,1)$ and let $\chi^2_\beta(k)$ and $\mathsf{F}_\beta(2,k)$ denote the β-quantiles of a χ^2- and F-distribution with the corresponding degrees of freedom, respectively. Moreover, let $\widehat{\vartheta}^*_{\mathsf{LU}} = \frac{1}{m}\sum_{j=1}^m (R_j+1)(Z_{j:m:n}-\mu)$ be the BLUE of ϑ in the scale model and

$$\widehat{\mu}_{\mathsf{LU}} = Z_{1:m:n}, \quad \widehat{\vartheta}_{\mathsf{LU}} = \frac{1}{m-1}\sum_{j=2}^m (R_j+1)(Z_{j:m:n}-Z_{1:m:n})$$

be the BLUEs of μ and ϑ in the location–scale model.

Then, two-sided $(1-\alpha)$ confidence intervals for ϑ are given by

(i) $\left[\frac{2m\widehat{\vartheta}^*_{\mathsf{LU}}}{\chi^2_{1-\alpha/2}(2m)}, \frac{2m\widehat{\vartheta}^*_{\mathsf{LU}}}{\chi^2_{\alpha/2}(2m)}\right]$,

(ii) $\left[\frac{(2m-2)\widehat{\vartheta}_{\mathsf{LU}}}{\chi^2_{1-\alpha/2}(2m-2)}, \frac{(2m-2)\widehat{\vartheta}_{\mathsf{LU}}}{\chi^2_{\alpha/2}(2m-2)}\right]$.

A two-sided $(1-\alpha)$ confidence interval for μ is given by

(iii) $\left[\widehat{\mu}_{\mathsf{LU}} - 2\mathsf{F}_{1-\alpha}(2,2m-2)\frac{\widehat{\vartheta}_{\mathsf{LU}}}{n}, \widehat{\mu}_{\mathsf{LU}}\right]$.

One-sided confidence intervals may be constructed similarly.

Proof. (i) The result follows directly from Theorem 12.1.1 because $\frac{2m}{\vartheta}\widehat{\vartheta}^*_{\mathsf{LU}} \sim \chi^2(2m)$.

(ii) For a $\Gamma(\vartheta, m-1)$-distributed random variable U, $2\frac{U}{\vartheta} \sim \chi^2(2(m-1))$. Thus, $\frac{2(m-1)\widehat{\vartheta}_{\mathsf{LU}}}{\vartheta} \sim \chi^2(2(m-1))$. Hence, we find for $\vartheta > 0$,

$$P_\vartheta\left(\vartheta \in \left[\frac{2(m-1)\widehat{\vartheta}_{\mathsf{LU}}}{\chi^2_{1-\alpha/2}(2(m-1))}, \frac{2(m-1)\widehat{\vartheta}_{\mathsf{LU}}}{\chi^2_{\alpha/2}(2(m-1))}\right]\right)$$
$$= P_\vartheta\left(\chi^2_{\alpha/2}(2(m-1)) \le \frac{2(m-1)\widehat{\vartheta}_{\mathsf{LU}}}{\vartheta} \le \chi^2_{1-\alpha/2}(2(m-1))\right) = 1-\alpha.$$

(iii) From $\frac{\widehat{\mu}_{\mathsf{LU}}-\mu}{\vartheta/n} \sim \mathsf{Exp}(1) = \chi^2(2)$, we obtain for

$$U = \frac{\widehat{\mu}_{\mathsf{LU}}-\mu}{\vartheta/n} \quad \text{and} \quad V = \frac{2(m-1)\widehat{\vartheta}_{\mathsf{LU}}}{\vartheta}$$

that $U \sim \chi^2(2)$, $V \sim \chi^2(2(m-1))$, and U, V are independent by Theorem 12.1.4. Therefore,

$$\frac{n(\widehat{\mu}_{\mathsf{LU}}-\mu)}{2\widehat{\vartheta}_{\mathsf{LU}}} = \frac{\frac{1}{2}U}{\frac{1}{2(m-1)}V} \sim \mathsf{F}(2, 2(m-1)).$$

This implies, for $\mu \in \mathbb{R}$, that

$$P_\mu\left(\mu \in \left[\widehat{\mu}_{\mathsf{LU}} - 2\mathsf{F}_{1-\alpha}(2, 2(m-1))\frac{\widehat{\vartheta}_{\mathsf{LU}}}{n}, \widehat{\mu}_{\mathsf{LU}}\right]\right) = P_\mu\left(0 \le \frac{n(\widehat{\mu}_{\mathsf{LU}} - \mu)}{2\widehat{\vartheta}_{\mathsf{LU}}} \le \mathsf{F}_{1-\alpha}(2, 2(m-1))\right) = 1-\alpha,$$

which proves the assertion (see also Remark 12.1.5). □

Remark 17.1.2. Under general progressive censoring, Balakrishnan and Lin [110] utilized an algorithm of Huffer and Lin [466] to compute exact confidence intervals for the scale parameter ϑ. The approach is based on the representation of the BLUEs as a linear function of the spacings. From (11.8) and (2.13), it is obvious that the BLUEs can be written in such a way.

Simultaneous confidence regions for μ and ϑ in the location–scale model have been obtained by Wu [909] following the ideas of Wu [904]. Wu [909] proposed two approaches to construct joint confidence regions. The constructions are based on the independence of the quantities

$$2n\frac{\widehat{\mu}_{\mathsf{LU}} - \mu}{\vartheta} \sim \chi^2(2), \quad \frac{2(m-1)}{\vartheta}\widehat{\vartheta}_{\mathsf{LU}} \sim \chi^2(2(m-1)),$$

and

$$n\frac{\widehat{\mu}_{\mathsf{LU}} - \mu}{\widehat{\vartheta}_{\mathsf{LU}}} \sim \mathsf{F}(2, 2m-2), \quad \frac{2}{\vartheta}\left(m\widehat{\vartheta}^*_{\mathsf{LU}} - n\mu\right) \sim \chi^2(2m).$$

Theorem 17.1.3 (Wu [909]). Let $\alpha \in (0,1)$ and $\chi^2_\beta(k)$ and $\mathsf{F}_\beta(2, 2k)$ denote the β-quantiles of a χ^2- and F-distribution with the corresponding degrees of freedom, respectively. Then, the following statistical intervals are joint confidence regions for μ and ϑ. Moreover, let $\alpha_- = (1-\sqrt{1-\alpha})/2$ and $\alpha_+ = (1+\sqrt{1-\alpha})/2$:

(i) $\mathscr{K}_1 = \Big\{(\mu,\vartheta) \,\Big|\, \widehat{\mu}_{\mathsf{LU}} - \frac{\chi^2_{\alpha_+}(2)}{2n}\vartheta < \mu < \widehat{\mu}_{\mathsf{LU}} - \frac{\chi^2_{\alpha_-}(2)}{2n}\vartheta,$
$\frac{2(m-1)}{\chi^2_{\alpha_+}(2m-2)}\widehat{\vartheta}_{\mathsf{LU}} < \vartheta < \frac{2(m-1)}{\chi^2_{\alpha_-}(2m-2)}\widehat{\vartheta}_{\mathsf{LU}}\Big\}$,

(ii) $\mathscr{K}_2 = \Big\{(\mu,\vartheta) \,\Big|\, \widehat{\mu}_{\mathsf{LU}} - \frac{\mathsf{F}_{\alpha_+}(2,2m-2)}{n}\widehat{\vartheta}_{\mathsf{LU}} < \mu < \widehat{\mu}_{\mathsf{LU}} - \frac{\mathsf{F}_{\alpha_-}(2,2m-2)}{n}\widehat{\vartheta}_{\mathsf{LU}},$
$2\frac{m\widehat{\vartheta}^*_{\mathsf{LU}} - n\mu}{\chi^2_{\alpha_+}(2m)} < \vartheta < 2\frac{m\widehat{\vartheta}^*_{\mathsf{LU}} - n\mu}{\chi^2_{\alpha_-}(2m)}\Big\}$,

where $\widehat{\vartheta}^*_{\mathsf{LU}} = \frac{1}{m}\sum_{j=1}^m (R_j + 1)Z_{j:m:n}$.

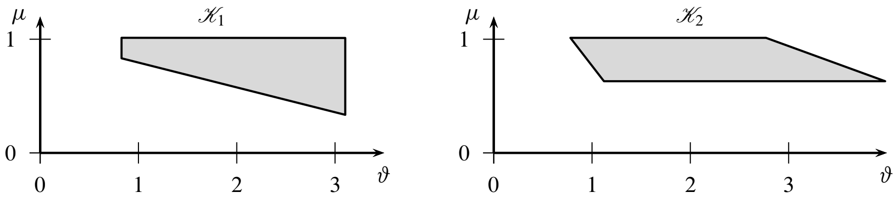

Fig. 17.1 Joint confidence regions $\mathscr{K}_1$ and $\mathscr{K}_2$ from Example 17.1.4

1.013	1.034	1.109	1.266	1.509	1.533	1.563
1.929	1.965	2.061	2.344	2.546	2.626	

Table 17.1 Progressively Type-II censored data generated by Wu [909] from a sample of duration of remission of $n = 20$ leukemia patients

Example 17.1.4. Wu [909] applied the above joint confidence regions to data taken from Lawless [574]. The data set of duration of remission of $n = 20$ leukemia patients treated by one drug has been progressively Type-II censored by the censoring scheme $\mathscr{R} = (1, 1, 0^{*10}, 5)$ leading to the censored data (measurements are in years) given in Table 17.1 of $m = 13$ observations.

The BLUEs in the two-parameter exponential model are given by $\widehat{\mu}_{\mathsf{LU}} = 1.013$ and $\widehat{\vartheta}_{\mathsf{LU}} = 1.45125$. Using the above methods, Wu [909] computed the following simultaneous 95 % confidence regions:

(i) $\mathscr{K}_1 = \Big\{(\mu, \vartheta) \,\Big|\, 1.013 - 0.2185\vartheta \le \mu \le 1.013 - 0.0006\vartheta, 0.8278 \le \vartheta \le 3.1038\Big\}$,
(ii) $\mathscr{K}_2 = \Big\{(\mu, \vartheta) \,\Big|\, 0.6305 \le \mu \le 1.0121, 1.5946 - 0.8946\mu \le \vartheta \le 5.6638 - 3.1776\mu\Big\}$.

Both confidence regions are depicted in Fig. 17.1. The regions $\mathscr{K}_1$ and $\mathscr{K}_2$ have an area of 0.9746 and 0.9254, respectively. Wu [909] conducted a simulation study revealing that the second confidence region covers a smaller area than the first one.

Remark 17.1.5. The multi-sample case has been discussed in Balakrishnan et al. [130] [see also (12.8)]. In this case, similar confidence intervals have been established using the fact that the MLEs of μ and ϑ are independent (see Cramer and Kamps [299]). Moreover, the estimates have an exponential and χ^2-distribution, respectively. Details can be found in Balakrishnan et al. [130].

From the above results, we can construct confidence intervals for quantiles of the exponential distribution. The quantile function of a two-parameter exponential distribution is given by $F^{\leftarrow}(t) = \mu - \vartheta \log(1 - t)$, $t \in (0, 1)$. Hence, we can estimate the quantile ξ_p, $p \in (0, 1)$, by the estimates

$$\widehat{\xi}_p^* = -\widehat{\vartheta}_{\mathsf{LU}}^* \log(1 - p), \quad \widehat{\xi}_p = \widehat{\mu}_{\mathsf{LU}} - \widehat{\vartheta}_{\mathsf{LU}} \log(1 - p)$$

if either $\mu = 0$ or μ is unknown, respectively.

From Theorem 12.1.1, we know that the BLUE $\widehat{\vartheta}^*_{\mathsf{LU}}$ has a $\Gamma(\vartheta/m, m)$-distribution. Therefore, $\widehat{\xi}^*_p$ is also gamma distributed but with parameters $-\vartheta \log(1-p)/m$ and m. From Corollary 17.1.1, we get the $(1-\alpha)$ confidence interval $\left[\frac{2m\widehat{\xi}^*_p}{\chi^2_{1-\alpha/2}(2m)}, \frac{2m\widehat{\xi}^*_p}{\chi^2_{\alpha/2}(2m)}\right]$ for ξ_p. In the location–scale model, a two-sided confidence interval can be obtained via the construction

$$\left[\widehat{\mu}_{\mathsf{LU}} - c_- \widehat{\vartheta}_{\mathsf{LU}} \log(1-p) \leq \xi_p \leq \widehat{\mu}_{\mathsf{LU}} - c_+ \widehat{\vartheta}_{\mathsf{LU}} \log(1-p)\right]. \tag{17.1}$$

Straightforward calculations lead to the equivalent relation

$$\frac{n}{2} c_+ \leq \frac{\frac{1}{2}U + \frac{n}{2}\log(1-p)}{\frac{1}{2m-2}V} \leq \frac{n}{2} c_-,$$

where $U = n(\widehat{\mu}_{\mathsf{LU}} - \mu)/\vartheta \sim \mathsf{Exp}(1) = \chi^2(2)$ and $V = 2(m-1)\widehat{\vartheta}_{\mathsf{LU}}/\vartheta \sim \chi^2(2m-2)$ are independent random variables. To compute the bounds, we have to calculate the distribution of this ratio.

Remark 17.1.6. The calculation of the distribution of the ratio $Q = \frac{\frac{1}{2}U + \frac{n}{2}\log(1-p)}{\frac{1}{2m-2}V}$ is closely related to calculating the OC curve given in (22.3) (see also Remark 22.5.2). This can be seen from its survival function which can be written as

$$P(Q > k) = P\Big(U - \frac{k}{m-1}V > -n\log(1-p)\Big).$$

In the same spirit, we can calculate exact confidence intervals for the reliability at a mission time t_0. In the scale model, it is easy to see from Corollary 17.1.1 that

$$c = \exp\left\{-\frac{t_0 \chi^2_{1-\alpha}(2m)}{2m\widehat{\vartheta}^*_{\mathsf{LU}}}\right\} \tag{17.2}$$

is a lower confidence limit for $R(t_0)$ at a confidence level $1-\alpha$. In the location–scale case, the situation is more involved. The MLE for the reliability $R(t_0) = \exp\{-(t_0-\mu)/\vartheta\}$ is given by $\widehat{R}(t_0) = \exp\{-(t_0-\widehat{\mu})/\widehat{\vartheta}\}$ provided that $t_0 \geq \widehat{\mu}$. This problem can be traced back to confidence intervals for a certain quantile. Suppose we are interested in an upper confidence interval for $R(t_0)$. Then, for a statistic $\tau = \tau(\mathbf{X}^{\mathscr{R}})$,

$$R(t_0) \geq \tau \iff \mu - \vartheta \log(\tau) \geq t_0 \iff \xi_{1-\tau} \geq t_0.$$

Recalling the construction in (17.1), this tells us that $[\tau, 1]$ is an upper confidence interval for $R(t_0)$ if t_0 is a lower confidence limit for $\xi_{1-\tau}$. Hence, τ is determined as the solution of the equation $\widehat{\xi}_{1-\tau(\mathbf{x})} = t_0$ for a given sample $\mathbf{x}$.

Example 17.1.7. Using the results given in Example 12.1.6, the estimate of the reliability at $t_0 = 2$ is given by $\widehat{R}(2) = 0.8024$ in the scale model. Then, (17.2) provides an upper 95 % confidence interval for $R(2)$, i.e., $[0.6965, 1]$.

17.1.2 Weibull Distribution

Wu [904] has obtained confidence intervals for the scale and shape parameters of a Weibull(ϑ, β)-distribution based on progressively Type-II censored data. The construction is based on the idea that progressively Type-II censored order statistics $X_{1:m:n}, \ldots, X_{m:m:n}$ from such a Weibull distribution can be expressed in terms of exponential progressively Type-II censored order statistics $Z_{1:m:n}, \ldots, Z_{m:m:n}$. Namely, we have

$$(X_{j:m:n})_{1\le j\le m} \stackrel{\mathsf{d}}{=} \left(Z_{j:m:n}^{1/\beta}\right)_{1\le j\le m}. \tag{17.3}$$

From the independence of spacings in the exponential case [see (2.9)], we conclude that

$$Z_1 = \frac{n}{\vartheta} X_{1:m:n}^{\beta}, \quad Z_j = \frac{\gamma_j}{\vartheta}(X_{j:m:n}^{\beta} - X_{j-1:m:n}^{\beta}), 2 \le j \le m, \tag{17.4}$$

are IID standard exponential random variables. Moreover, this shows that $V = 2\frac{n}{\vartheta} X_{1:m:n}^{\beta} \sim \chi^2(2)$ and

$$U = 2\sum_{j=2}^{m} Z_j = \frac{2}{\vartheta}\sum_{j=2}^{m}(R_j + 1)(X_{j:m:n}^{\beta} - X_{1:m:n}^{\beta}) \sim \chi^2(2(m-1)) \tag{17.5}$$

are independent random variables. Hence, $T_1 = U/[(m-1)V] \sim \mathsf{F}(2(m-1), 2)$. Wu [904] established the confidence interval for β using that, for $1 < t_2 \le \cdots \le t_m$, the function

$$\psi : (0, \infty) \longrightarrow (0, \infty), \quad \beta \mapsto \psi(\beta) = \frac{1}{n(m-1)}\sum_{j=2}^{m}(R_j + 1)(t_i^{\beta} - 1)$$

is strictly increasing with $\lim_{\beta\to 0}\psi(\beta) = 0$ and $\lim_{\beta\to\infty}\psi(\beta) = \infty$. This shows that the equation $\psi(\beta) = x$ has a unique solution in β for any $x > 0$.

Theorem 17.1.8 (Wu [904]). Let $\alpha \in (0, 1)$ and $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ be a sample of progressively Type-II censored order statistics from a two-parameter Weibull distribution Weibull(ϑ, β). Then, a $(1-\alpha)$ confidence interval for β is given by

$$\mathscr{K} = \Big[\psi^*(\mathbf{X}^{\mathscr{R}}, \mathsf{F}_{\alpha/2}(2(m-1),2)), \psi^*(\mathbf{X}^{\mathscr{R}}, \mathsf{F}_{1-\alpha/2}(2(m-1),2))\Big],$$

where $\psi^*(\mathbf{X}^{\mathscr{R}}, \omega)$ is the unique solution for β of the equation

$$\sum_{j=2}^{m}(R_j+1)\Big(\frac{X_{j:m:n}}{X_{1:m:n}}\Big)^{\beta} = \gamma_2 + n(m-1)\omega. \tag{17.6}$$

Proof. First, notice that (17.6) is equivalent to the equation $\psi(\beta) = \omega$ with t_j replaced by $X_{j:m:n}/X_{1:m:n}$. Hence, we have a unique solution in β for any $\omega > 0$. Therefore, the condition $\beta \in \mathscr{K}$ is equivalent to

$$\mathsf{F}_{1-\alpha/2}(2(m-1),2) \le \psi(\beta) \le \mathsf{F}_{\alpha/2}(2(m-1),2).$$

Since $\psi(\beta) = T_1 \sim \mathsf{F}(2(m-1),2)$, the assertion is proved. □

Recently, Wang et al. [890] have established a confidence interval for β via a different approach. Using (17.4), they considered the partial sums $S_i = \sum_{j=1}^{i} Z_j$, $1 \le i \le m$, and the ratios S_i/S_m, $1 \le i \le m-1$. These ratios are distributed as uniform order statistics from a sample of size $m-1$ so that the pivotal quantity

$$\begin{aligned}\tau(\mathbf{X}^{\mathscr{R}}, \beta) &= -2\sum_{j=1}^{m-1}\log(S_j/S_m)\\ &= 2\sum_{j=1}^{m-1}\log\left(\frac{\sum_{i=1}^{m}(R_i+1)X_{i:m:n}^{\beta}}{\sum_{i=1}^{j-1}(R_i+1)X_{i:m:n}^{\beta} + \gamma_j X_{j:m:n}^{\beta}}\right)\end{aligned}$$

has a χ^2-distribution with $2(m-1)$ degrees of freedom. Moreover, it is free of the scale parameter ϑ. Wang et al. [890] showed that $\tau(\mathbf{x}, \beta)$ is strictly increasing in β for any $\mathbf{x} > 0$. Additionally, $\lim_{\beta\to 0}\tau(\mathbf{X}^{\mathscr{R}}, \beta) = 0$ and $\lim_{\beta\to\infty}\tau(\mathbf{X}^{\mathscr{R}}, \beta) = \infty$ for $0 < x_1 < \cdots < x_m$. Hence, the equation $\tau(\mathbf{x}, \beta) = t$ has a unique solution for any $t > 0$. This yields the following result.

Theorem 17.1.9 (Wang et al. [890]). Let $\alpha \in (0,1)$ and $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ be a sample of progressively Type-II censored order statistics from a two-parameter Weibull(ϑ, β)-distribution. Then, a $(1-\alpha)$ confidence interval for β is given by

$$\mathscr{K} = \Big[\tau^{-1}(\mathbf{X}^{\mathscr{R}}, \chi^2_{\alpha/2}(2(m-1))), \tau^{-1}(\mathbf{X}^{\mathscr{R}}, \chi^2_{1-\alpha/2}(2(m-1)))\Big],$$

where $\tau^{-1}(\mathbf{X}^{\mathscr{R}}, \omega)$ is the unique solution for β of the equation $\tau(\mathbf{X}^{\mathscr{R}}, \beta) = \omega$ with $\omega > 0$.

Example 17.1.10. The above confidence intervals are computed for Nelson's progressively Type-II censored data 1.1.5. Using the approach of Wu [904], the 95 % confidence interval $(0.3242, 1.7692)$ for β results. Wang et al. [890] obtained the 95 % confidence interval $(0.3909, 1.4872)$. Notice that it is included in the interval of Wu [904] so that it provides a shorter interval in the given situation.

Remark 17.1.11. In general, the confidence intervals proposed by Wu [904] and Wang et al. [890] need not be subsets. For instance, replacing the largest observation 7.35 in Nelson's progressively Type-II censored data used in Example 17.1.10 by 73.5, the estimates $[0.2364, 1.1304]$ (Wu) and $[0.2025, 0.6844]$ (Wang et al.) result.

Using the substitution method of Weerahandi [894, 895], Wang et al. [890] proposed generalized confidence intervals for, e.g., the scale parameter ϑ and quantiles ξ_p. Writing $T(\beta) = \sum_{j=1}^{m}(R_j + 1)X_{i:m:n}^{\beta}$, it follows that $V = 2S_m = \frac{2}{\vartheta}T(\beta)$ has a $\chi^2(2m)$-distribution. On the other hand, $\vartheta = 2T(\beta)/V$. Following the construction of generalized confidence intervals, Wang et al. [890] make use of the generalized pivotal quantity

$$Y_t = \frac{2S_m}{2\sum_{j=1}^{m}(R_j + 1)x_j^{\tau^{-1}(\mathbf{x},t)}},$$

where $\mathbf{x} = (x_1, \ldots, x_m)$ is the given data and $t \in (0, 1)$. $\tau^{-1}(\mathbf{x}, t)$ is the solution of the equation $\tau(\mathbf{x}, \beta) = t$ for the observed data. Conditionally on $\mathbf{X}^{\mathscr{R}} = \mathbf{x}$, the distribution of $Y_t - \vartheta$ and the distribution of Y_t are free of any unknown parameter. For $\alpha \in (0, 1)$, the bounds Y_α and $Y_{1-\alpha}$ form a generalized confidence interval. These bounds can be determined by Monte Carlo simulations. A similar approach can be used to estimate the mean, a quantile ξ_p, and the reliability function at a given point.

The following theorem provides a simultaneous confidence region for the distribution parameters.

Theorem 17.1.12 (Wu [904]). Let $\alpha \in (0, 1)$ and $\chi^2_\beta(k)$ and $\mathsf{F}_\beta(k, 2)$ denote the β-quantiles of a χ^2- and F-distribution with the corresponding degrees of freedom, respectively. Then, the following statistical interval is a joint confidence region for β and ϑ. Moreover, let $\alpha_- = (1 - \sqrt{1-\alpha})/2$ and $\alpha_+ = (1 + \sqrt{1-\alpha})/2$:

$$\mathscr{K} = \Big\{(\beta, \vartheta) \,\Big|\, \psi^*(\mathbf{X}^{\mathscr{R}}, \mathsf{F}_{\alpha_-}(2(m-1), 2)) \le \beta \le \psi^*(\mathbf{X}^{\mathscr{R}}, \mathsf{F}_{\alpha_+}(2(m-1), 2)),$$
$$\frac{2\sum_{j=1}^{m}(R_j + 1)X_{j:m:n}^{\beta}}{\chi^2_{\alpha_+}(2m)} \le \vartheta \le \frac{2\sum_{j=1}^{m}(R_j + 1)X_{j:m:n}^{\beta}}{\chi^2_{\alpha_-}(2m)}\Big\}, \quad (17.7)$$

where $\psi^*(\mathbf{X}^{\mathscr{R}}, \omega)$ is the unique solution for β of the Eq. (17.6).

Proof. From the representation in (17.5), we get that $T_2 = U + V = \frac{2}{\vartheta}\sum_{j=1}^{m}(R_j + 1)X_{j:m:n}^{\beta}$ has a χ^2-distribution with $2m$ degrees of freedom. Moreover, from Johnson et al. [483, pp. 350], we conclude that T_2 is independent of the ratio $T_1 = \frac{U}{(m-1)V} \sim \mathsf{F}(2(m-1), 2)$. This implies that

$$P\Big(\mathsf{F}_{\alpha_-}(2(m-1),2)) \le T_1 \le \mathsf{F}_{\alpha_+}(2(m-1),2)\Big) = \sqrt{1-\alpha},$$

$$P\Big(\chi^2_{\alpha_-}(2m) \le T_2 \le \chi^2_{\alpha_+}(2m)\Big) = \sqrt{1-\alpha}.$$

Using the independence of T_1 and T_2, we can multiply these probabilities to arrive at

$$\begin{aligned}1-\alpha &= P\Big(\mathsf{F}_{\alpha_-}(2(m-1),2)) \le T_1 \le \mathsf{F}_{\alpha_+}(2(m-1),2),\\ &\qquad\qquad \chi^2_{\alpha_-}(2m) \le T_2 \le \chi^2_{\alpha_+}(2m)\Big)\\ &= P\Big(\psi^*(\mathbf{X}^{\mathscr{R}}, \mathsf{F}_{\alpha_-}(2(m-1),2)) \le \beta \le \psi^*(\mathbf{X}^{\mathscr{R}}, \mathsf{F}_{\alpha_+}(2(m-1),2)),\\ &\qquad \frac{2\sum_{j=1}^{m}(R_j+1)X_{j:m:n}^{\beta}}{\chi^2_{\alpha_+}(2m)} \le \vartheta \le \frac{2\sum_{j=1}^{m}(R_j+1)X_{j:m:n}^{\beta}}{\chi^2_{\alpha_-}(2m)}\Big).\end{aligned}$$

This completes the proof. □

Example 17.1.13. For Nelson's progressively Type-II censored data 1.1.5, we get the joint confidence region

$$\begin{aligned}\mathscr{K} = \Big\{(\beta,\vartheta)\,\Big|\, & 0.2807 \le \beta \le 1.9648,\\ & \frac{2\sum_{j=1}^{m}(R_j+1)x_j^{\beta}}{31.2070} < \vartheta < \frac{2\sum_{j=1}^{m}(R_j+1)x_j^{\beta}}{6.0684}\Big\}.\end{aligned}$$

The resulting set is depicted in Fig. 17.2. Notice that the shape of the area is different from that given in Wu [904]. This is due to a different parametrization of the Weibull distribution. Wu [904] has used the parametrization ϑ^{β}. This leads to very huge bounds for ϑ^{β} when β is small. In order to represent the region in graphic form, Wu [904] suggested a logarithmic scaling of the ϑ-axis. This is not necessary in our setting.

Remark 17.1.14. Mann [634] has addressed exact interval estimation for a distribution quantile t_R, where R denotes a specified survival proportion. In particular, she was interested in determining a lower confidence bound for the reliability $R(t_0)$ for a given time t_0. Assuming a Weibull$(\vartheta^{1/\beta}, 1/\beta)$-distribution, she proposed a lower bound which depends on three progressively Type-II censored order statistics $Y_{\nu:m:n} = \log X_{\nu:m:n}$, $Y_{p:m:n} = \log X_{p:m:n}, Y_{q:m:n} =$

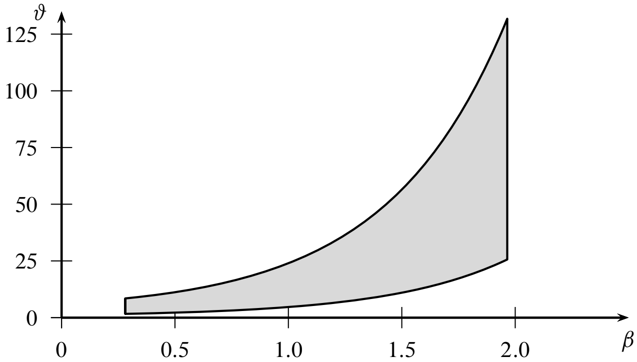

Fig. 17.2 Joint confidence region $\mathscr{K}$ for (β, ϑ) from Example 17.1.13.

$\log X_{q:m:n}$, $1 \le \nu \le m$, $1 \le p < q \le m$. The result is based on the derivation of the exact distribution of the ratio

$$V_{\nu,p,q} = \frac{\log(-\log R) - Y^*_{\nu:m:n}}{Y^*_{q:m:n} - Y^*_{p:m:n}},$$

where $Y^*_{k:m:n} = (Y_{k:m:n} - \log\vartheta)/\beta$.

17.1.3 Pareto Distribution

Exact joint confidence intervals for the parameters of two-parameter Pareto distributions with cumulative distribution function

$$F(t;\lambda,\beta) = 1 - \left(\frac{\beta}{t}\right)^{1/\vartheta}, \quad t \ge \beta, \tag{17.8}$$

are discussed in Kuş and Kaya [566], Parsi et al. [712], Wu [908], and Fernández [368]. Let $T = \frac{1}{n}\sum_{j=1}^{m}(R_j + 1)\log(X_{j:m:n}/X_{1:m:n})$. Using arguments similar to the Weibull case [see (17.3)], i.e., that

$$Z_{j:m:n} = \frac{1}{\vartheta}\log(X_{j:m:n}/\beta), \quad 1 \le j \le m,$$

are exponential progressively Type-II censored order statistics, Kuş and Kaya [566] obtained the joint $(1-\alpha)$ confidence set for β and ϑ as

$$\mathscr{K} = \Big\{(\beta,\vartheta) \,\Big|$$

$$X_{1:m:n}\exp\left\{-\frac{T}{(m-1)\mathsf{F}_{\sqrt{1-\alpha}}(2(m-1),2)}\right\} \le \beta \le X_{1:m:n},$$

$$2n\frac{T-\log(\beta/X_{1:m:n})}{\chi^2_{\alpha+}(2m)} \leq \vartheta \leq 2n\frac{T-\log(\beta/X_{1:m:n})}{\chi^2_{\alpha-}(2m)}\Big\},$$

where $\alpha_- = (1-\sqrt{1-\alpha})/2$ and $\alpha_+ = (1+\sqrt{1-\alpha})/2$. Parsi et al. [712] established a similar confidence region where the β-part is identical, but the ϑ-part is not constructed symmetrically. They used levels α_1^* and α_2^* with

$$\alpha_1^* + \alpha_2^* = 1 - \sqrt{1-\alpha} \tag{17.9}$$

such that $\chi^2_{1-\alpha_1^*}(2m) - \chi^2_{\alpha_2^*}(2m)$ has minimum width. They found a confidence region for $(\beta, 1/\vartheta)$ that always covers a smaller area than that proposed by Kuş and Kaya [566]. Here, for fixed β, the width of the confidence region in direction of ϑ are proportional to the difference

$$\ell_\beta = 2n\left[\frac{1}{\chi^2_{\alpha^*}(2m)} - \frac{1}{\chi^2_{\alpha^*+\sqrt{1-\alpha}}(2m)}\right]\left[T - \log(\beta/X_{1:m:n})\right],$$

where $\alpha^* = 1-\alpha_1^*$ and $\alpha_2^* = \alpha^* + \sqrt{1-\alpha}$ from (17.9). The same comment applies to the coverage area which can be written in our parametrization as

$$\begin{aligned} A &= 2n\left[\frac{1}{\chi^2_{\alpha^*}(2m)} - \frac{1}{\chi^2_{\alpha^*+\sqrt{1-\alpha}}(2m)}\right]\int_{x_1 \mathrm{e}^{-q}}^{x_1} [T - \log(\beta/x_1)]d\beta \\ &= 2n\left[\frac{1}{\chi^2_{\alpha^*}(2m)} - \frac{1}{\chi^2_{\alpha^*+\sqrt{1-\alpha}}(2m)}\right]x_1\left[T - (T+q)\mathrm{e}^{-q}\right], \end{aligned}$$

where $q = T/[(m-1)\mathsf{F}_{\sqrt{1-\alpha}}(2(m-1),2)]$ and $\alpha^* \in [0, 1-\sqrt{1-\alpha}]$. Therefore, we look for α^* such that

$$\Delta(\alpha^*; m, \alpha) = \frac{1}{\chi^2_{\alpha^*}(2m)} - \frac{1}{\chi^2_{\alpha^*+\sqrt{1-\alpha}}(2m)} \tag{17.10}$$

is minimal for $\alpha^* \in [0, 1-\sqrt{1-\alpha}]$. A plot of $\Delta(\cdot; 5, 0.05)$ is depicted in Fig. 17.3. The minimum is attained for $\alpha^* = 0.02445$ leading to $1-\alpha_1^* = 0.0245$ and $\alpha_2^* = 0.9991$ and $\Delta(\alpha^*; 5, 0.05) = 0.2764$. Using α_+ and α_-, the corresponding values are $\alpha_- = 0.0127$ and $\alpha_+ = 0.9873$ and $\Delta(\alpha_-; 5, 0.05) = 0.3238$. Therefore, the minimal choice of the α^* yields a 15 % smaller coverage area than the proposal by Kuş and Kaya [566]. A similar argument applies to the confidence region (17.7) obtained for the Weibull distribution.

The cumulative distribution function (17.8) is covered by the approach of Wang et al. [890] [see (17.11)] with $h(t;\beta) = \log(t/\beta)$. Therefore, a $(1-\alpha)$ confidence interval for β can be directly obtained as in Theorem 17.1.9 using (17.12) with

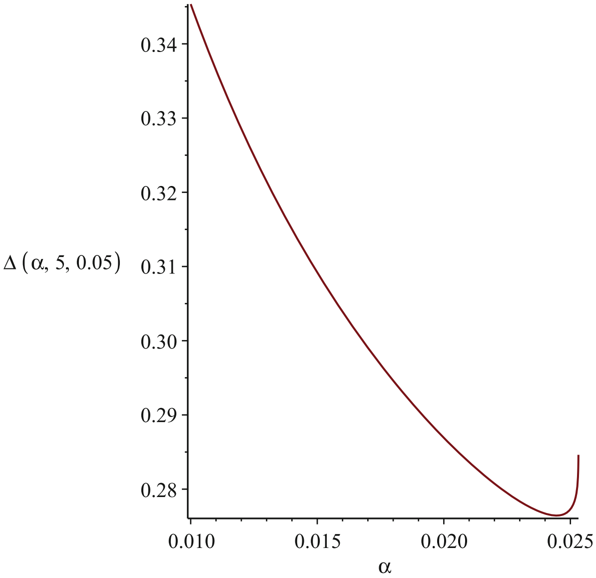

Fig. 17.3 Plot of $\Delta(\cdot; 5, 0.05)$ as in (17.10) in the interval $[0, 1-\sqrt{0.95}] = [0, 0.0253]$

$$\tau(\mathbf{X}^{\mathscr{R}}, \beta) = 2\sum_{j=1}^{m-1} \log\left(\frac{\sum_{i=1}^{m}(R_i+1)\log(X_{i:m:n}/\beta)}{\sum_{i=1}^{j-1}(R_i+1)\log(X_{i:m:n}/\beta) + \gamma_j \log(X_{j:m:n}/\beta)}\right).$$

Notice that the ratio $\log(a/\beta)/\log(b/\beta)$, $a > b$, is decreasing in β so that the bounds of the confidence interval have to be reversed [see Sect. 17.1.4 and comments after (17.12)].

Remark 17.1.15 (Doubly). Type-II right censored samples have been discussed in Chen [247], Wu [907], and, more recently, Zhang [939] who has provided simplified versions of the confidence intervals.

Recently, Fernández [368] established a generalization of those confidence regions proposed by Chen [247] using Pareto order statistics. He showed that the versions of Kuş and Kaya [566] and Parsi et al. [712] for progressively censored data are contained as special cases in his construction. A similar generalization has been obtained for the confidence region obtained in Wu [908]. Furthermore, smallest area confidence regions for Pareto parameters are presented in Fernández [368]. Using the fiducial argument due to Fisher [372, 374], the optimal joint confidence region w.r.t. smallest area has been derived and a simple regula falsi procedure is used to compute the optimal confidence region. The results have been illustrated by a progressively Type-II censored sample given in Table 17.2 which were generated from data reported in Wu et al. [918]. Fernández [368] pointed out that the covered area is reduced significantly by using the smallest area confidence regions.

i	1	2	3	4	5	6	7	8
$x_{i:8:20}$	0.0098	0.0376	0.0661	0.0849	0.1112	0.1447	0.1904	0.2463

Table 17.2 Progressively Type-II censored sample as reported in Fernández [368] with censoring scheme $\mathscr{R} = (1, 0, 2, 0, 3, 2, 0, 4)$

17.1.4 Other Parametric Distributions

Wang et al. [890] obtained confidence intervals for a distribution parameter by a method similar to that used in the Weibull case (see Theorem 17.1.9). They considered exponentiated distributions with cumulative distribution function

$$F(t; \lambda, \beta) = 1 - (1 - G(t; \lambda))^{\beta}, \quad t \in \mathbb{R}, \tag{17.11}$$

where $G(\cdot; \lambda)$ has only the parameter λ. For instance, assuming that $G(\cdot; \lambda)$ can be written as

$$G(t; \lambda) = 1 - \mathrm{e}^{-h(t;\lambda)}, \quad t \in \mathbb{R},$$

with an increasing function $h(\cdot; \lambda)$, confidence intervals can be established provided that $h(b; \lambda)/h(a; \lambda)$, $b > a$, is either strictly increasing or decreasing in λ. Examples of such distributions discussed in Wang et al. [890] are Gompertz distribution, i.e., $h(t; \lambda) = \mathrm{e}^{\lambda t} - 1$, or Lomax distribution, i.e., $h(t; \lambda) = -\log(1 + \lambda t)$. In these cases, we get the same result as in Theorem 17.1.9, where

$$\tau(\mathbf{X}^{\mathscr{R}}, \lambda) = 2 \sum_{j=1}^{m-1} \log \left(\frac{\sum_{i=1}^{m} (R_i + 1) h(X_{i:m:n}; \lambda)}{\sum_{i=1}^{j-1} (R_i + 1) h(X_{i:m:n}; \lambda) + \gamma_j h(X_{j:m:n}; \lambda)} \right). \tag{17.12}$$

Notice that the bounds have to be reversed when the ratio $h(b; \lambda)/h(a; \lambda)$, $b > a$, is decreasing. This happens for the Lomax distribution. The case of Gompertz distribution [see (12.42)] has also been addressed in Wu et al. [914]. A bathtub-shaped lifetime distribution belonging to this family with cumulative distribution function F in (12.43) has recently been discussed by Wu [906] and Wu et al. [923]. In the latter, several proposals for the confidence regions are compared in a simulation study w.r.t. the criteria of highest power, minimum confidence width, and the minimum confidence region. Similar results for Burr XII are provided in Wu et al. [922].

Wang [883] considered a scale family of scaled half-logistic distributions. He established an exact confidence interval

$$\left[\tau^{-1}\left(\chi^2_{1-\alpha/2}(2m)\right), \tau^{-1}\left(\chi^2_{\alpha/2}(2m)\right)\right],$$

where $\tau^{-1}(\omega)$ is the unique solution of the equation $\tau(\mathbf{X}^{\mathscr{R}}, t) = \omega$, $t > 0$. $\tau(\mathbf{X}^{\mathscr{R}}, \cdot)$ is defined by

$$\tau(\mathbf{X}^{\mathscr{R}}, t) = 2\sum_{j=1}^{m}(R_j + 1)\log\left(1 + \mathrm{e}^{X_{j:m:n}/t}\right) - 2n\log(2), \quad t > 0.$$

This approach has been applied to a scale family of distributions including Weibull, log-logistic, and gamma distributions by Wang [885] for general progressively censored data. Nigm and Abo-Eleneen [692] established an exact confidence interval for the shape parameter of an inverse Weibull distribution by adapting the approach of Wu [904].

17.1.5 Nonparametric Confidence Intervals for Quantiles

Exact confidence intervals for parametric models are widely used. They are preferable when a parametric model is appropriate. If such model assumptions are questionable, a nonparametric approach can be considered. We now present nonparametric confidence intervals for quantiles based on a progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$. This issue was first addressed by Guilbaud [418] who transfers an idea well known in nonparametric inference for population quantiles to progressive censoring. This approach is explained in detail in Arnold et al. [58, pp. 183–184] and David and Nagaraja [327, Sect. 7.1]. The crucial idea is to use the quantile representation of the order statistics in terms of uniform order statistics. In fact, for a sample $X_{1:n}, \ldots, X_{n:n}$, the coverage probability of a confidence interval $[X_{k:n}, X_{\ell:n}]$, with $1 \le k < \ell \le n$ for ξ_p, $p \in (0, 1)$, is given by

$$\begin{aligned} P(X_{k:n} \le \xi_p \le X_{\ell:n}) &= P(F^{\leftarrow}(U_{k:n}) \le F^{\leftarrow}(p) \le F^{\leftarrow}(U_{\ell:n})) \\ &= P(U_{k:n} \le p \le U_{\ell:n}) = \sum_{i=k}^{\ell-1}\binom{n}{i}p^i(1-p)^{n-i}. \end{aligned}$$

This observation dates back to Thompson [844]. For a progressively Type-II censored sample, an analogous construction can be made as it was observed first by Guilbaud [418]. To compute the coverage probability $P(X_{k:m:n} \le \xi_p \le X_{\ell:m:n})$, he made use of the mixture representation of progressively Type-II censored order statistics in terms of order statistics (see Sect. 10.1 as well as Thomas and Wilson [843] and Fischer et al. [371]). However, as pointed out in Balakrishnan et al. [144], this approach is not recommended for computational purposes. Moreover, using the quantile representation of progressively Type-II censored order statistics given in Theorem 2.1.1, the necessary coverage probabilities can be easily calculated. Given a progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$, the coverage probability of a one-sided confidence interval $(-\infty, X_{\ell:m:n}]$ considered in Guilbaud [418] can be directly calculated using the explicit expression for the cumulative distribution function of $X_{\ell:m:n}$ as established by Kamps and Cramer [503] [see also (2.25)]:

$$P(\xi_p < X_{\ell:m:n}) = P(p < U_{\ell:m:n}) = 1 - F^{U_{\ell:m:n}}(p) \tag{17.13}$$

	ℓ							
p	1	2	3	4	5	6	7	8
0.1	0.1351	0.4203	0.7054	0.9006	0.9746	0.9962	0.9996	1.0000
0.2	0.0144	0.0829	0.2369	0.5003	0.7377	0.9114	0.9775	0.9957
0.3	0.0011	0.0104	0.0462	0.1663	0.3641	0.6492	0.8514	0.9523
0.4	0.0001	0.0008	0.0055	0.0343	0.1139	0.3280	0.5881	0.8028
0.5	0.0000	0.0000	0.0004	0.0042	0.0218	0.1120	0.2960	0.5399
0.6	0.0000	0.0000	0.0000	0.0003	0.0023	0.0241	0.0998	0.2618
0.7	0.0000	0.0000	0.0000	0.0000	0.0001	0.0028	0.0194	0.0792
0.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0015	0.0109
0.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003

Table 17.3 Coverage probabilities of one-sided confidence intervals $(-\infty, X_{\ell:m:n}]$ for ξ_p computed from (17.13) with censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$. Values for $p \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ can also be found in Guilbaud [418, Table 1]

$$= \Big(\prod_{i=1}^{\ell} \gamma_i\Big) \sum_{j=1}^{\ell} \frac{1}{\gamma_j} a_{j,\ell}(1-p)^{\gamma_j}, \quad p \in (0,1).$$

This formula yields quite an efficient method to compute coverage probabilities for several confidence intervals.

Example 17.1.16. We apply the method to Nelson's insulating fluid data 1.1.4 considered in Guilbaud [418]. A selection of coverage probabilities for the particular censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$ is presented in Table 17.3. The values show that the maximum coverage probability of a lower nonparametric confidence interval $(-\infty, X_{\ell:m:n}]$ is given by 0.5399. The corresponding interval is $(-\infty, X_{8:8:19}]$.

As mentioned above, the coverage probabilities may not yield a desired confidence level. In order to overcome this problem, Balakrishnan et al. [144] studied the problem of adding a second independent sample to increase the coverage probabilities. Suppose $X_{1:r:n}, \dots, X_{r:r:n}$ and $Y_{1:s:m}, \dots, Y_{s:s:m}$ are two independent progressively Type-II censored samples from a population cumulative distribution function F with censoring schemes $\mathscr{R} = (R_1, \dots, R_r)$ and $\mathscr{S} = (S_1, \dots, S_s)$, respectively. Without loss of any generality, let $r \geq s$. Then, the two samples are pooled and ordered leading to the ordered pooled sample

$$W^*_{(1)} \leq \cdots \leq W^*_{(r+s)}. \tag{17.14}$$

Since the quantile function $F^{\leftarrow}$ preserves ordering, it is clear that the ordered pooled sample exhibits the property

$$(W^*_{(1)}, \dots, W^*_{(r+s)}) \stackrel{\mathrm{d}}{=} \big(F^{\leftarrow}(W_{(1)}), \dots, F^{\leftarrow}(W_{(r+s)})\big),$$

where $W_{(1)}, \ldots, W_{(r+s)}$ is constructed by analogy but based on two samples from the uniform distributions. Hence, it follows that, for the calculation of the coverage probabilities, we can restrict ourselves to a uniform distribution.

Naturally, confidence intervals for a quantile ξ_p of the population cumulative distribution function F are constructed by, e.g., $(-\infty, W^*_{(\ell)}]$, where the coverage probability is given by

$$P(\xi_p \leq W^*_{(\ell)}) = P(p \leq W_{(\ell)}).$$

Therefore, an expression for the latter quantity has to be established. As presented in Theorem 2.4.1, the ith progressively Type-II censored order statistic $X_{i:r:n}$ depends only on the right truncated censoring scheme $\mathscr{R}_{\triangleright i-1} = (R_1, \ldots, R_{i-1})$ and the sample size n. Denoting by $G^{\mathscr{R}_{\triangleright i-1}}_{i:n}$ the cumulative distribution function of the ith uniform progressively Type-II censored order statistic, Balakrishnan et al. [144] established the following expressions for the marginal cumulative distribution function of $W_{(\ell)}$. To condense the notation, we introduce in addition to the notation used in Theorem 2.4.2 the quantities

$$\eta_j = \sum_{i=j}^{s}(S_i + 1),\ 1 \leq j \leq s, \qquad d_{j-1} = \prod_{i=1}^{j} \eta_i,\ 1 \leq j \leq s,$$

$$b_{j,s} = \prod_{\substack{i=1\\ i\neq j}}^{r} \frac{1}{\eta_i - \eta_j}.$$

Then, the following expressions for the coverage probabilities result.

Theorem 17.1.17. (i) For $i = 1, \ldots, s$ and $0 \leq p \leq 1$, $F^{W_{(i)}}(p)$ equals

$$\prod_{\ell=1}^{i} \frac{\gamma_\ell}{(\gamma_\ell + \eta_1)} G^{\mathscr{R}_{\triangleright i-1}}_{i:n+\eta_1}(p) + \prod_{\ell=1}^{i} \frac{\eta_\ell}{(\eta_\ell + \gamma_1)} G^{\mathscr{S}_{\triangleright i-1}}_{i:m+\gamma_1}(p)$$
$$+ \sum_{k=1}^{i-1} d_{i-k-1} c_{k-1} \left[\sum_{\ell=1}^{i-k+1} \frac{b_{\ell,i-k+1}}{\prod_{\nu=1}^{k}(\gamma_\nu + \eta_\ell)} G^{\mathscr{R}_{\triangleright k-1}}_{k:n+\eta_\ell}(p) \right]$$
$$+ \sum_{k=1}^{i-1} c_{i-k-1} d_{k-1} \left[\sum_{\ell=1}^{i-k+1} \frac{a_{\ell,i-k+1}}{\prod_{\nu=1}^{k}(\eta_\nu + \gamma_\ell)} G^{\mathscr{S}_{\triangleright k-1}}_{k:m+\gamma_\ell}(p) \right];$$

(ii) For $i = s+1, \ldots, r$ and $0 \leq p \leq 1$, $F^{W_{(i)}}(p)$ equals

$$G^{\mathscr{R}_{\triangleright i-s-1}}_{i-s:n}(p) - d_{s-1} \sum_{j=1}^{s} \frac{b_{j,s}}{\eta_j \prod_{\nu=1}^{i-s}(\gamma_\nu + \eta_j)} G^{\mathscr{R}_{\triangleright i-s-1}}_{i-s:n+\eta_j}(p)$$
$$+ \prod_{\ell=1}^{i} \frac{\gamma_\ell}{(\gamma_\ell + \eta_1)} G^{\mathscr{R}_{\triangleright i-1}}_{i:n+\eta_1}(p)$$

$$+ \sum_{k=i-s+1}^{i-1} d_{i-k-1} c_{k-1} \left[\sum_{\ell=1}^{i-k+1} \frac{b_{\ell,i-k+1}}{\prod_{\nu=1}^{k}(\gamma_\nu + \eta_\ell)} G_{k:n+\eta_\ell}^{\mathscr{R}\rhd k-1}(p) \right]$$

$$+ \sum_{k=1}^{s} c_{i-k-1} d_{k-1} \left[\sum_{\ell=1}^{i-k+1} \frac{a_{\ell,i-k+1}}{\prod_{\nu=1}^{k}(\eta_\nu + \gamma_\ell)} G_{k:m+\gamma_\ell}^{\mathscr{S}\rhd k-1}(p) \right];$$

(iii) For $i = r+1, \ldots, r+s$ and $0 \le p \le 1$, $F^{W_{(i)}}(p)$ equals

$$G_{i-s:n}^{\mathscr{R}\rhd i-s-1}(p) - d_{m-1} \sum_{j=1}^{s} \frac{b_{j,s}}{\eta_j \prod_{\nu=1}^{i-s}(\gamma_\nu + \eta_j)} G_{i-s:n+\eta_j}^{\mathscr{R}\rhd i-s-1}(p)$$

$$+ \sum_{k=i-s+1}^{r} d_{i-k-1} c_{k-1} \left[\sum_{\ell=1}^{i-k+1} \frac{b_{\ell,i-k+1}}{\prod_{\nu=1}^{k}(\gamma_\nu + \eta_\ell)} G_{k:n+\eta_\ell}^{\mathscr{R}\rhd k-1}(p) \right]$$

$$+ G_{i-r:m}^{\mathscr{S}\rhd i-r-1}(p) - c_{n-1} \sum_{j=1}^{r} \frac{a_{j,r}}{\gamma_j \prod_{\nu=1}^{i-r}(\eta_\nu + \gamma_j)} G_{i-r:m+\gamma_j}^{\mathscr{S}\rhd i-r-1}(p)$$

$$+ \sum_{k=i-r+1}^{s} c_{i-k-1} d_{k-1} \left[\sum_{\ell=1}^{i-k+1} \frac{a_{\ell,i-k+1}}{\prod_{\nu=1}^{k}(\eta_\nu + \gamma_\ell)} G_{k:m+\gamma_\ell}^{\mathscr{S}\rhd k-1}(p) \right].$$

Remark 17.1.18. The joint distribution of $W_{(1)}^*, \ldots, W_{(r+s)}^*$ in (17.14) has been established in Balakrishnan et al. [154] as a mixture of distributions of certain progressively Type-II censored order statistics (see also Theorem 17.1.23). In particular, they found that

$$(W_{(1)}^*, \ldots, W_{(r+s)}^*) \stackrel{\mathrm{d}}{=} \sum_{\mathscr{T} \in \mathscr{C}(\mathscr{R},\mathscr{S})} \pi_{\mathscr{T}} P^{(X_{1:r+s:n+m}^{\mathscr{T}}, \ldots, X_{r+s:r+s:n+m}^{\mathscr{T}})},$$

where $\mathscr{C}(\mathscr{R}, \mathscr{S})$ is a set of censoring schemes generated by merging the censoring schemes $\mathscr{R} = (R_1, \ldots, R_r)$ and $\mathscr{S} = (S_1, \ldots, S_s)$ such that the order of the censoring numbers in each scheme is preserved in $\mathscr{T}$. The mixing probability $\pi_{\mathscr{T}}$ is defined as

$$\pi_{\mathscr{T}} = \frac{\prod_{i=1}^{r} \gamma_i(\mathscr{R}) \prod_{i=1}^{s} \gamma_i(\mathscr{S})}{\prod_{i=1}^{r+s} \gamma_i(\mathscr{T})} \#(\mathscr{T}),$$

where $\#(\mathscr{T})$ is the number of all permutations of $R_1, \ldots, R_r, S_1, \ldots, S_s$ leading to $\mathscr{T}$ subject to respecting the order within the censoring plans $\mathscr{R}$ and $\mathscr{S}$.

The coverage probabilities of the ordered pooled sample exceed the coverage probabilities of the one-sample confidence intervals. The exact gain is given in the following lemma.

Lemma 17.1.19. Let $F_{j:r:n}$ denote the cumulative distribution function of a uniform progressively Type-II censored order statistic $U_{j:r:n}$. Then, the maximum gains in the coverage probabilities are given by

(i) $\Delta_u = (1-p)^{\max\{n,m\}}(1-(1-p)^{\min\{n,m\}}) > 0$,
(ii) $\Delta_\ell = \min\{F_{r:r:n}(p), F_{s:s:m}(p)\}(1-\max\{F_{r:r:n}(p), F_{s:s:m}(p)\}) > 0$,
(iii) $\Delta_2 = \min\{F_{r:r:n}(p)+(1-p)^n, F_{s:s:m}(p)+(1-p)^m\} - F_{r:r:n}(p)F_{s:s:m}(p) - (1-p)^{m+n} \geq \Delta_\ell + \Delta_u > 0$.

Proof. By construction of the pooled sample, we have the identity $W_{(1)} = \min\{X_{1:r:n}, Y_{1:s:m}\} = \min\{X_{1:n}, Y_{1:m}\}$. Hence,

$$P(W_{(1)} \leq \xi_p) = P(\min\{X_{1:n}, Y_{1:m}\} \leq \xi_p) = 1-(1-p)^{n+m}.$$

On the other hand, $W_{(r+s)} = \max\{X_{r:r:n}, Y_{s:s:m}\}$ so that

$$P(W_{(r+s)} \leq \xi_p) = P(\max\{X_{r:r:n}, Y_{s:s:m}\} \leq \xi_p) = F_{r:r:n}(p)F_{s:s:m}(p).$$

This yields the second assertion. Combining the previous results, we obtain for the two-sided interval

$$\begin{aligned}
\Delta_2 &= P(W_{(1)} \leq \xi_p \leq W_{(r+s)}) \\
&\qquad - \max\{P(X_{1:r:n} \leq p \leq X_{r:r:n}), P(Y_{1:s:m} \leq p \leq Y_{s:s:m})\} \\
&= 1-(1-p)^{n+m} - F_{r:r:n}(p)F_{s:s:m}(p) \\
&\qquad - \max\{1-(1-p)^n - F_{r:r:n}(p), 1-(1-p)^m - F_{s:s:m}(p)\} \\
&= \min\{F_{r:r:n}(p) + (1-p)^n, F_{s:s:m}(p) + (1-p)^m\} \\
&\qquad - F_{r:r:n}(p)F_{s:s:m}(p) - (1-p)^{m+n} \\
&\geq \min\{F_{r:r:n}(p), F_{s:s:m}(p)\} + \min\{(1-p)^n, (1-p)^m\} \\
&\qquad - F_{r:r:n}(p)F_{s:s:m}(p) - (1-p)^{m+n} \\
&= \min\{F_{r:r:n}(p), F_{s:s:m}(p)\}(1-\max\{F_{r:r:n}(p), F_{s:s:m}(p)\}) \\
&\qquad\qquad + (1-p)^{\max\{m,n\}}(1-(1-p)^{\min\{m,n\}}) \\
&= \Delta_\ell + \Delta_u > 0.
\end{aligned}$$

This proves the assertion. □

Notice that the gains are directly connected to the maximum coverage probabilities of a two-sample confidence interval which are calculated in the previous proof. In particular, these quantities are given by

(i) $\mathsf{CP}_u = 1-(1-p)^{n+m}$,
(ii) $\mathsf{CP}_\ell = F_{r:r:n}(p)F_{s:s:m}(p)$,
(iii) $\mathsf{CP}_2 = 1-(1-p)^{n+m} - F_{r:r:n}(p)F_{s:s:m}(p)$.

Additional progressively censored data sample		
$s = 4, m = 9$	$s = 5, m = 12$	$s = 7, m = 15$
$(S_1, \dots, S_4)$	$(S_1, \dots, S_5)$	$(S_1, \dots, S_7)$
$(0^{*3}, 5)$	$(0^{*4}, 7)$	$(0^{*6}, 8)$
0.6567	0.6291	0.6796
$(3, 0, 1^{*2})$	$(2, 0, 2^{*2}, 1)$	$(0^{*2}, 4, 0^{*2}, 2^{*2})$
0.8609	0.8525	0.8742
$(1, 2^{*2}, 0)$	$(2, 3, 0, 2, 0)$	$(0, 3, 0, 2^{*2}, 0, 1)$
0.9055	0.9361	0.9429
$(1, 4, 0^{*^2})$	$(2, 5, 0^{*3})$	$(0, 3^{*2}, 0^{*2}, 2, 0)$
0.9377	0.9695	0.9705
$(5, 0^{*3})$	$(7, 0^{*4})$	$(8, 0^{*6})$
0.9589	0.9791	0.9950

Table 17.4 Maximum one-sided coverage probabilities for the population median obtained by adding an independent sample $Y^{\mathscr{S}}_{1:s:m}, \dots, Y^{\mathscr{S}}_{s:s:m}$ to Nelson's insulating fluid data 1.1.4. The table is based on Table 4 in Balakrishnan et al. [144]

It is clear from these calculations that the coverage probability can be increased by adding more samples. In particular, the coverage probability of a lower confidence interval in the k-sample situation is given by

$$\mathsf{CP}_\ell = \prod_{i=1}^{k} F_{r_i:r_i:n}(p).$$

Since $F_{r_i:r_i:n}(p) \in (0, 1)$ for $p \in (0, 1)$, any desired confidence level $1 - \alpha$ can be ensured by adding more samples. An analogous argument holds for the other confidence intervals.

Example 17.1.20. Nonparametric confidence intervals for a quantile are constructed based on Nelson's progressively censored insulating fluid data 1.1.5 (see also Guilbaud [418]). The maximum one-sided confidence interval $[0, X_{8:8:19}]$ for the population median has only a coverage probability of 0.5399 (see Table 17.3). In Table 17.4, some values of the maximum one-sided coverage probabilities for the median are presented when Nelson's data is combined with another independent progressively Type-II censored sample with censoring scheme $\mathscr{S}$. It turns out that the censoring plan has a great impact on the total coverage probability and so it has to be planned carefully when the procedure is carried out by design.

Another possibility to improve the coverage probability with one sample has been discussed by Balakrishnan and Han [98]. Noticing that the censoring plan influences the coverage probability, they have addressed the problem of determining an optimal censoring design leading to a maximum coverage probability. Given that $m = 8$ and $n = 19$, they found that the censoring scheme $\mathscr{R}_{0.99} = (0, 11, 0^{*6})$ is optimal ensuring a confidence level of 99 % for the interval $[X_{2:8:19}^{\mathscr{R}_{0.99}}, X_{8:8:19}^{\mathscr{R}_{0.99}}]$ and an expected minimal width in the uniform case. For a confidence level of 95 %, they obtained the interval $[X_{4:8:19}^{\mathscr{R}_{0.95}}, X_{8:8:19}^{\mathscr{R}_{0.95}}]$ with $\mathscr{R}_{0.95} = (0^{*4}, 11, 0^{*3})$. Details are provided in Sect. 26.5.2.

Remark 17.1.21. As seen in the preceding example, the coverage probabilities typically differ from prefixed confidence levels. For this reason, different approaches have been proposed to construct exact or nearly exact confidence intervals (up to the maximum confidence level in the sample) (see Zieliński and Zieliński [947] and Hutson [467]). Obviously, these approaches can be applied to the two-sample case to construct confidence intervals using the sample $W_{(1)}, \ldots, W_{(r+s)}$.

An extension to the multi-sample case is presented in Volterman et al. [878].

17.1.6 An Excursus: Two-Sample Nonparametric Confidence Intervals from Type-II Censored Data

In this section, we address a particular case of the above problem discussed by Balakrishnan et al. [144] in detail. An extension to the multiple sample case is presented in Volterman and Balakrishnan [876]. Namely, let $X_{1:n}, \ldots, X_{r:n}$ and $Y_{1:m}, \ldots, Y_{s:m}$ be independent right censored samples from a uniform distribution with sample sizes n and m, respectively. Without loss of any generality, let $r \geq s$. As above, denote the ordered pooled sample by

$$W_{(1)} \leq \cdots \leq W_{(r+s)}.$$

According to Balakrishnan et al. [144], the following probabilities result.

Lemma 17.1.22. (i) $\pi_0 = P(Y_{s:m} < X_{1:n}) = \frac{\binom{n+m-s}{n}}{\binom{n+m}{n}}$;

(ii) For $i = 1, \ldots, r-1$, we have $\pi_i = P(X_{i:n} < Y_{s:m} < X_{i+1:n}) = \frac{\binom{s+i-1}{s-1}\binom{n+m-s-i}{n-i}}{\binom{n+m}{n}}$;

(iii) $\pi_0^* = P(X_{r:n} < Y_{1:m}) = \frac{\binom{n+m-r}{m}}{\binom{n+m}{m}}$;

(iv) For $i = 1, \ldots, s-1$, we have $\pi_i^* = P(Y_{i:m} < X_{r:n} < Y_{i+1:m}) = \frac{\binom{r+i-1}{r-1}\binom{n+m-r-i}{m-i}}{\binom{n+m}{m}}$.

Up to some sets of measure zero, these events form a decomposition of the probability space so that $\sum_{i=0}^{r-1} \pi_i + \sum_{i=0}^{s-1} \pi_i^* = 1$.

From Theorem 17.1.17, it is clear that the marginal cumulative distribution functions in the ordered pooled sample can be represented as a mixture of cumulative distribution functions of usual order statistics using Theorem 10.1.1. But, it turns out that the connection is much deeper. Namely, the distribution of the ordered pooled sample is a mixture of progressively Type-II censored order statistics. Before presenting this result due to Balakrishnan et al. [144], we introduce random variables $T_{i:r+s:n+m}^{\mathscr{R}_j}$, $1 \le i \le r+s$, $0 \le j \le r-1$, as uniform progressively Type-II censored order statistics based on the one-step censoring scheme

$$\mathscr{R}_j = (0^{*s+j-1}, m-s, 0^{*r-j-1}, n-r),$$

and $T_{i:r+s:n+m}^{\mathscr{R}_j^*}$ $(1 \le i \le r+s,\ 0 \le j \le s-1)$ as uniform progressively Type-II censored order statistics based on the censoring scheme

$$\mathscr{R}_j^* = (0^{*r+j-1}, n-r, 0^{*s-j-1}, m-s).$$

Now, Theorem 17.1.23 shows that the distribution of the ordered pooled sample $W_{(1)}, \ldots, W_{(r+s)}$ is indeed a mixture of uniform progressively Type-II censored order statistics with weights given in Lemma 17.1.22. The corresponding censoring schemes are one-step censoring schemes (with additional right censoring). The generalization to progressively Type-II censored data due to Balakrishnan et al. [154] is mentioned in Remark 17.1.18.

Theorem 17.1.23 (Balakrishnan et al. [144]). The joint distribution of the ordered pooled sample $W_{(1)}, \ldots, W_{(r+s)}$ is a mixture of uniform progressively Type-II censored order statistics given by

$$(W_{(1)}, \ldots, W_{(r+s)}) \stackrel{\mathrm{d}}{=} \sum_{j=0}^{r-1} \pi_j P^{(T_{1:r+s:n+m}^{\mathscr{R}_j}, \ldots, T_{r+s:r+s:n+m}^{\mathscr{R}_j})} + \sum_{j=0}^{s-1} \pi_j^* P^{(T_{1:r+s:n+m}^{\mathscr{R}_j^*}, \ldots, T_{r+s:r+s:n+m}^{\mathscr{R}_j^*})},$$

where π_j and π_j^* are as given in Lemma 17.1.22.

Proof. Let $\mathbf{x}_k = (x_1, \ldots, x_k) \in \mathbb{R}^k$. First, the γ's corresponding to the censoring schemes $\mathscr{R}_j$ and $\mathscr{R}_j^*$ are given by

$$\gamma_\ell(\mathscr{R}_j) = \begin{cases} n+m-\ell+1, & \ell = 1, \ldots, s+j \\ n+s-\ell+1, & \ell = s+j+1, \ldots, r+s \end{cases}, \quad j = 0, \ldots, r-1,$$

$$\gamma_\ell(\mathscr{R}_j^*) = \begin{cases} n+m-\ell+1, & \ell = 1, \ldots, r+j \\ m+r-\ell+1, & \ell = r+j+1, \ldots, r+s \end{cases}, \quad j = 0, \ldots, s-1. \tag{17.15}$$

This yields, for instance, the normalizing constant

$$\prod_{\ell=1}^{r+s} \gamma_\ell(\mathscr{R}_j) = \frac{(n+m)!(n-j)!}{(n+m-s-j)!(n-r)!}, \quad j = 0, \ldots, r-1.$$

It is sufficient to consider the sets $\{Y_{s:m} < X_{1:n}\}$ and $\{X_{i:n} < Y_{s:m} < X_{i+1:n}\}$ (for $1 \le i \le r-1$) only, since the remaining cases follow by symmetry arguments. Let $t_1 \le \cdots \le t_{r+s}$, $A = \{(\mathbf{y}_s, \mathbf{x}_r) | y_1 \le \cdots \le y_s < x_1 \le \cdots \le x_r\}$, and

$$g(\mathbf{y}_s, \mathbf{x}_r) = c_{r-1} c_{s-1} \prod_{k=1}^{s} f(y_k) \overline{F}^{m-s}(y_s) \prod_{\ell=1}^{r} f(x_\ell) \overline{F}^{n-r}(x_r) \mathbb{1}_A(\mathbf{y}_s, \mathbf{x}_r)$$

be the joint density function of the two (independent) samples, given $y_s < x_1$. Then, with $B = \times_{k=1}^{r+s}(-\infty, t_k]$, we have

$$\begin{aligned} P(W_{(j)} \le t_j, 1 \le j \le r+s, Y_{s:m} < X_{1:n}) &= \int_B g(\mathbf{y}_s, \mathbf{x}_r) d\mathbf{x}_r d\mathbf{y}_s \\ &= \frac{c_{s-1} c_{r-1}}{c(\mathscr{R}_0)} P(T^{\mathscr{R}_0}_{j:r+s:n+m} \le t_j, 1 \le j \le r+s) \\ &= \frac{\binom{n+m-s}{n}}{\binom{n+m}{n}} P(T^{\mathscr{R}_0}_{j:r+s:n+m} \le t_j, 1 \le j \le r+s), \end{aligned}$$

where $c(\mathscr{R}_0)$ is the normalizing constant corresponding to the joint density function of progressively Type-II censored order statistics with censoring scheme $\mathscr{R}_0$ [see (17.15)].

Let $1 \le i \le r-1$, $\mathfrak{S}_{s+i-1}$ be the set of permutations of $\{1, \ldots, s+i-1\}$ and $\mathfrak{S}^{\le}_{s+i-1} \subseteq \mathfrak{S}_{s+i-1}$ be the set of permutations σ with $\sigma(1) < \cdots < \sigma(s-1)$ and $\sigma(s) < \cdots < \sigma(s+i-1)$. Moreover, let $V_k = Y_{k:m}$ $(1 \le k \le s-1)$ and $V_{s+k-1} = X_{k:n}$ $(1 \le k \le i)$. For $\{X_{i:n} < Y_{s:m} < X_{i+1:n}\}$, we obtain with $A_i = \{(\mathbf{v}_{s+i-1}, y_s, x_{i+1}, \ldots, x_r) | v_1 \le \cdots \le v_{s+i-1} < y_s < x_{i+1} \le \cdots \le x_r\}$

$$\begin{aligned} &P(W_{(j)} \le t_j, 1 \le j \le r+s, X_{i:n} < Y_{s:m} < X_{i+1:n}) \\ &= \sum_{\sigma \in \mathfrak{S}_{s+i-1}} P(W_{(j)} \le t_j, 1 \le j \le r+s, \\ &\qquad V_{\sigma(1)} < \cdots < V_{\sigma(s+i-1)}, X_{i:n} < Y_{s:m} < X_{i+1:n}) \\ &= \sum_{\sigma \in \mathfrak{S}^{\le}_{s+i-1}} P(W_{(j)} \le t_j, 1 \le j \le r+s, \\ &\qquad V_{\sigma(1)} < \cdots < V_{\sigma(s+i-1)}, X_{i:n} < Y_{s:m} < X_{i+1:n}) \qquad (17.16) \\ &= \sum_{\sigma \in \mathfrak{S}^{\le}_{s+i-1}} \int_B g_i(v_{\sigma(1)}, \ldots, v_{\sigma(s+i-1)}, y_s, x_{i+1}, \ldots, x_r) dx_r \ldots dx_{i+1} dy_s d\mathbf{v}_{s+i-1}, \end{aligned}$$

where

$$g_i(\mathbf{v}_{s+i-1}, y_s, x_{i+1}, \ldots, x_r) = c_{s-1}c_{r-1} \prod_{k=1}^{s+i-1} f(v_k) f(y_s)\overline{F}^{m-s}(y_s)$$
$$\times \prod_{\ell=i+1}^{r} f(x_\ell)\overline{F}^{n-r}(x_r)\mathbb{1}_{A_i}(\mathbf{v}_{s+i-1}, y_s, x_{i+1}, \ldots, x_r).$$

The probabilities in (17.16) are zero if $\sigma \notin \mathfrak{S}^{\leq}_{s+i-1}$. Since g_i is invariant with respect to σ and $|\mathfrak{S}^{\leq}_{s+i-1}| = \binom{s+i-1}{s-1}$, we find using (17.15) that

$$\begin{aligned}
&P(W_{(j)} \leq t_j, 1 \leq j \leq r+s, X_{i:n} < Y_{s:m} < X_{i+1:n})\\
&= \binom{s+i-1}{s-1} \int_B g(\mathbf{v}_{s+i-1}, y_s, x_{i+1}, \ldots, x_r) dx_r \ldots dx_{i+1} dy_s d\mathbf{v}_{s+i-1}\\
&= \frac{\binom{s+i-1}{s-1}\binom{n+m-s-i}{n-i}}{\binom{n+m}{n}} P(T^{\mathcal{R}_i}_{j:r+s:n+m} \leq t_j, 1 \leq j \leq r+s),
\end{aligned}$$

as required. □

Intuitively, the mixture result can be explained as follows. Conditional on the event $\{Y_{s:m} < X_{1:n}\}$, the ordered sample $(W_{(1)}, \ldots, W_{(r+s)})$ is given by

$$(Y_{1:m}, \ldots, Y_{s:m}, X_{1:n}, \ldots, X_{r:n}).$$

Taking into account an IID sample of size $n+m$ from F, this particular sample can be seen as follows. First, the s smallest observations are noted. Then, $m-s$ larger variables of the original sample are randomly withdrawn. Afterwards, the next r observations are considered, which means that the largest $n-r$ variates are censored. This progressive censoring procedure is associated with the censoring scheme $\mathscr{R}_0$.

Similarly, conditional on the event $\{X_{i:n} < Y_{s:m} < X_{i+1:n}\}$, the ordered sample $(W_{(1)}, \ldots, W_{(r+s)})$ is given by

$$(V_1, \ldots, V_{s+i-1}, Y_{s:m}, X_{i+1:n}, \ldots, X_{r:n}),$$

where $V_1, \ldots, V_{s+i-1}$ is an arrangement of the random variables $X_{1:n}, \ldots, X_{i:n}$ and $Y_{1:m}, \ldots, Y_{s-1:m}$. Hence, $V_1, \ldots, V_{s+i-1}, Y_{s:m}$ equal the first $s+i$ order statistics in a sample of size $n+m$. Then, $m-s$ larger variables are withdrawn from the sample, and the remaining ordered observations follow. This sample corresponds to a progressively Type-II censored sample associated with the censoring scheme $\mathscr{R}_i$.

Remark 17.1.24. As mentioned before, Guilbaud [418, 419] has expressed the distribution of progressively Type-II censored order statistics as a mixture of distributions of order statistics. Hence, the distribution of $(W_{(1)}, \ldots, W_{(r+s)})$ is a mixture of mixtures of the usual order statistics from a sample of size $n + m$.

Computational aspects of the above procedure have been discussed in Balakrishnan et al. [155]. They proposed a branch and bound procedure for efficient computation of the confidence intervals. An extension to the multi-sample case has been obtained by Volterman et al. [878]. The approach has also been utilized by Beutner and Cramer [194] in the models of minimal repair and record values.

17.2 Conditional Statistical Intervals

So far, all the inferential methods we have discussed are unconditional in nature. In this section, we will demonstrate how exact confidence intervals or prediction intervals may be obtained using the conditional method. Conditional inference, first proposed by Fisher [373], has been successfully applied by Lawless [570, 571, 572, 574, 575] to develop inference based on complete as well as conventionally Type-II right censored samples. As a matter of fact, Lawless [574, pp. 199] indicated the use of conditional inference based on progressively Type-II right censored data, but a full length account of this topic has been provided by Viveros and Balakrishnan [875] which naturally forms a basis for much of the discussion in this section.

17.2.1 Inference in a General Location–Scale Family

Let $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ be a sample of progressively Type-II censored order statistics from a location–scale family of distributions as in (11.1), i.e.,

$$\mathscr{F}_{ls} = \left\{ F\left(\frac{\cdot - \mu}{\vartheta}\right) \middle| \mu \in \mathbb{R}, \vartheta > 0 \right\}. \tag{17.17}$$

Suppose $\widehat{\mu}$ and $\widehat{\vartheta}$ denote the MLEs of μ and ϑ. Then, $\widehat{\mu}$ and $\widehat{\vartheta}$ are equivariant estimators, i.e., they satisfy the identities

$$\widehat{\mu}(d\mathbf{X}^{\mathscr{R}} + c\mathbb{1}) = d\widehat{\mu}(\mathbf{X}^{\mathscr{R}}) + c \quad \text{and} \quad \widehat{\vartheta}(d\mathbf{X}^{\mathscr{R}} + c\mathbb{1}) = d\widehat{\vartheta}(\mathbf{X}^{\mathscr{R}})$$

for any constants $c \in \mathbb{R}$ and $d > 0$. This can be directly seen from the likelihood function

$$\mathsf{L}(\mu, \vartheta; \mathbf{x}) = \prod_{j=1}^{m} \left[\frac{\gamma_j}{\vartheta} f\left(\frac{x_j - \mu}{\vartheta}\right) \left[1 - F\left(\frac{x_j - \mu}{\vartheta}\right)\right]^{R_j} \right] \prod_{j=2}^{m} \mathbb{1}_{[x_{j-1}, \infty)}(x_j)$$

following the arguments as in Lawless [575, pp. 562]. Since equivariance is the crucial property of estimators in conditional inference, the MLEs may be replaced by any equivariant estimators $\widetilde{\mu}$ and $\widetilde{\vartheta}$ of μ and ϑ.

Now, the estimators

$$T_1 = \frac{\widehat{\mu} - \mu}{\widehat{\vartheta}} \quad \text{and} \quad T_2 = \frac{\widehat{\vartheta}}{\vartheta} \tag{17.18}$$

are pivotal quantities. Their joint distribution is free of the distribution parameters μ and ϑ. Moreover,

$$\mathbf{A}^{\mathscr{R}} = (A_{1:m:n}, \ldots, A_{m:m:n}) = \left(\frac{X_{1:m:n} - \widehat{\mu}}{\widehat{\vartheta}}, \ldots, \frac{X_{m:m:n} - \widehat{\mu}}{\widehat{\vartheta}} \right)$$

forms a set of ancillary statistics where only $m - 2$ are functionally independent (cf. Lawless [575, pp. 562]). Hence, inference for (μ, ϑ) may be established on the distribution of (T_1, T_2), conditional on an observation $\mathbf{a}$ of $\mathbf{A}^{\mathscr{R}}$. Writing

$$\frac{x_i - \mu}{\vartheta} = \frac{(x_i - \widehat{\mu}) + (\widehat{\mu} - \mu)}{\widehat{\vartheta}} \cdot \frac{\widehat{\vartheta}}{\vartheta} = a_i t_2 + t_1 t_2,$$

the density function of (T_1, T_2), given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m = (a_1, \ldots, a_m)$, is given by

$$f^{T_1,T_2|\mathbf{A}^{\mathscr{R}}}(t_1, t_2|\mathbf{a}_m) = k(\mathbf{a}_m) t_2^{m-1} \prod_{i=1}^{m} f(a_i t_2 + t_1 t_2)[1 - F(a_i t_2 + t_1 t_2)]^{R_i}, t_1 \in \mathbb{R}, t_2 > 0, \tag{17.19}$$

where $k(\mathbf{a}_m)$ depends on $\mathbf{a}_m$, m, $\mathscr{R}$, and n only. $k(\mathbf{a}_m)$ can be obtained by integrating the conditional density function w.r.t. t_1 and t_2. By integrating out the other variable, one gets the conditional marginal density functions $f^{T_1|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a}_m)$ and $f^{T_2|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a}_m)$. They can be used for conditional inference for the parameters μ and ϑ. In general, these integrals may only be computed by numerical methods. However, for the exponential distribution, explicit expressions result.

Conditional Confidence Intervals for Location and Scale Parameters

Conditional confidence intervals may be constructed via the marginal conditional density functions. For instance, confidence intervals for the location parameter may be obtained by the condition

$$P(\ell_1 \le T_1 \le \ell_2) = \int_{\ell_1}^{\ell_2} f^{T_1|\mathbf{A}^{\mathscr{R}}}(t|\mathbf{a}_m) dt = 1 - \alpha,$$

where $\alpha \in (0,1)$. ℓ_1 and ℓ_2 depend on $\mathbf{a}_m$ but not on μ and ϑ. They may be chosen as quantiles of the conditional cumulative distribution function $F^{T_1|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a}_m)$. Although these intervals are obtained by conditioning, the resulting intervals are common confidence intervals (see Lawless [575, pp. 564]).

Further inferential issues involve inference for quantiles, reliability, and prediction intervals (see also Lawless [572], Fraser [382]). By analogy with the above approach, conditional confidence intervals for quantiles and reliability as well as conditional prediction intervals for the lifetimes of future objects can be constructed.

Conditional Confidence Intervals for Quantiles

Given a quantile ξ_p, $p \in (0,1)$, of the baseline cumulative distribution function $F((\cdot-\mu)/\vartheta)$, we have $p = F((\xi_p-\mu)/\vartheta)$. Obviously,

$$\xi_p = \mu + \vartheta F^{\leftarrow}(p),$$

which can be estimated by

$$\widehat{\xi}_p = \widehat{\mu} + \widehat{\vartheta} F^{\leftarrow}(p).$$

Conditional inference can be developed based on

$$T_p = \frac{\xi_p - \widehat{\mu}}{\widehat{\vartheta}}$$

which is a pivotal quantity. It can be written as $T_p = \frac{F^{\leftarrow}(p)}{T_2} - T_1$ with T_1 and T_2 as in (17.18). Using the density transformation formula, we get the conditional density function of (T_p, T_2), given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m$, as

$$f^{T_p,T_2|\mathbf{A}^{\mathscr{R}}}(t_p, t_2|\mathbf{a}_m)$$
$$= k_p(\mathbf{a}_m) {t_2}^{m-1} \prod_{i=1}^{m} f(a_i t_2 - t_p t_2 + F^{\leftarrow}(p))[1 - F(a_i t_2 - t_p t_2 + F^{\leftarrow}(p))]^{R_i},$$
$$t_p \in \mathbb{R}, t_2 > 0.$$

Integrating w.r.t. t_2 yields the marginal conditional density function of T_p, given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m$, which can be utilized to establish conditional confidence intervals for the quantile ξ_p as described above for the distribution parameters.

Conditional Confidence Intervals for Reliability

Suppose we are interested in the reliability at a given mission time t_0. Then,

$$R(t_0) = 1 - F\Big(\frac{t_0 - \mu}{\vartheta}\Big)$$

denotes the reliability at t_0. In order to construct a lower confidence limit for $R(t_0)$, we consider a lower confidence bound $\ell_p(\mathbf{X}^{\mathscr{R}})$ for the quantile ξ_{1-p}, $p \in (0, 1)$, at level $\alpha \in (0, 1)$. Then, by construction,

$$P(\ell_p(\mathbf{X}^{\mathscr{R}}) \leq \xi_{1-p}) = 1 - \alpha.$$

Choosing p as a solution of the equation $\ell_p(\mathbf{X}^{\mathscr{R}}) = t_0$ yields the desired confidence limit. Denoting by $p_0 = p_0(\mathbf{X}^{\mathscr{R}})$ that solution, we get

$$P(R(t_0) \geq p_0(\mathbf{X}^{\mathscr{R}})) = P(\ell_p(\mathbf{X}^{\mathscr{R}}) \leq \xi_{1-p}) = 1 - \alpha$$

(see also Remark 17.1.6 and comments following this remark).

Conditional Prediction Intervals for Future Failure Times

A prediction problem considered often in the literature is that of predicting the minimal and maximal lifetimes in a future sample $Y_1, \ldots, Y_r$ from the same population cumulative distribution function $F(\cdot; \mu, \vartheta)$. This sample is supposed to be independent of the present data sample $\mathbf{X}^{\mathscr{R}}$. The density function of the minimum $Y_{1:r}$ is given by

$$f_{1:r}(t) = \frac{r}{\vartheta}\Big\{1 - F\Big(\frac{t - \mu}{\vartheta}\Big)\Big\}^{r-1} f\Big(\frac{t - \mu}{\vartheta}\Big). \qquad (17.20)$$

By assumption, it is independent of (T_1, T_2), given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m$. Therefore, the joint conditional density function of (T_1, T_2, T_3), with $T_3 = \frac{Y_{1:r} - \widehat{\mu}}{\widehat{\vartheta}}$, is given by

$$\begin{aligned} &f^{T_1,T_2,T_3|\mathbf{A}^{\mathscr{R}}}(t_1, t_2, t_3|\mathbf{a}_m) \\ &= k_*(\mathbf{a}_m) r t_2^{m-1} \prod_{i=1}^{m} f(a_i t_2 - t_p t_2 + F^{\leftarrow}(p))[1 - F(a_i t_2 - t_p t_2 + F^{\leftarrow}(p))]^{R_i} \\ &\qquad \times f((t_1 + t_3)t_2)[1 - F((t_1 + t_3)t_2)]^{r-1}, \quad t_1, t_3 \in \mathbb{R}, t_2 > 0. \end{aligned}$$

Integrating out t_1 and t_2, we arrive at the marginal conditional density function of T_3, given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m$. This density function can be used to construct conditional prediction intervals for $Y_{1:r}$ by analogy with the construction of conditional confidence intervals using $f^{T_3|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a}_m)$.

A conditional prediction for $Y_{r:r}$ can be obtained by replacing (17.20) by

$$f_{r:r}(t) = \frac{r}{\vartheta}\left\{F\left(\frac{t-\mu}{\vartheta}\right)\right\}^{r-1} f\left(\frac{t-\mu}{\vartheta}\right)$$

and proceeding as above. With $T_4 = \frac{Y_{r:r}-\widehat{\mu}}{\widehat{\vartheta}}$, the resulting conditional density function is given by

$$\begin{aligned} & f^{T_1,T_2,T_4|\mathbf{A}^{\mathscr{R}}}(t_1,t_2,t_4|\mathbf{a}_m) \\ &= k_{**}(\mathbf{a}_m) r t_2^{m-1} \prod_{i=1}^{m} f(a_i t_2 - t_p t_2 + F^{\leftarrow}(p))[1 - F(a_i t_2 - t_p t_2 + F^{\leftarrow}(p))]^{R_i} \\ &\qquad \times f((t_1+t_4)t_2)[F((t_1+t_4)t_2)]^{r-1}, \quad t_1, t_4 \in \mathbb{R}, t_2 > 0. \end{aligned}$$

17.2.2 Exponential Distribution

Suppose the baseline distribution is an $\mathsf{Exp}(\mu,\vartheta)$-distribution. Then, in the location–scale family (17.17), we have the cumulative distribution function $F(x) = 1 - \mathrm{e}^{-x}$, $x \geq 0$, and density function $f(x) = \mathrm{e}^{-x}$, $x \geq 0$, respectively. In this case, $(\widehat{\mu}, \widehat{\vartheta})$ forms a complete sufficient statistic for (μ, ϑ). Therefore, the pivotal quantities T_1, T_2, T_p, T_3, and T_4 are all statistically independent of the ancillary statistic $\mathbf{A}^{\mathscr{R}}$ according to Basu's theorem (see Basu [181], Boos and Hughes-Oliver [215], and Lehmann and Casella [582]). From Theorem 12.1.4, we have

$$\widehat{\mu}_{\mathsf{MLE}} = Z_{1:m:n} \quad \text{and} \quad \widehat{\vartheta}_{\mathsf{MLE}} = \frac{1}{m}\sum_{j=2}^{m}(R_j+1)(Z_{j:m:n} - Z_{1:m:n})$$

to be independent, $\frac{n(m-1)}{m} T_1 \sim \mathsf{F}_{2,2m-2}$, and $2mT_2 \sim \chi^2(2m-2)$. Using this result, we arrive at the same confidence intervals given in Corollary 17.1.1. Thus, the conditional inference approach leads to the same exact confidence intervals as the direct approach. The same comment applies to confidence intervals for both quantile ξ_p and the reliability $R(t_0)$ at a given mission time t_0.

17.2.3 Extreme Value Distribution

Viveros and Balakrishnan [875] considered the location–scale family of extreme value distributions which forms an important example in the area of lifetime modeling. For instance, log-times of a two-parameter Weibull distribution belong to such a family. This means that, given data from a two-parameter Weibull distribution, a log-transform leads to the above location–scale family with

$$f(t) = \mathrm{e}^{t-\mathrm{e}^t}, \quad t \in \mathbb{R}.$$

Substituting the cumulative distribution function and density function in (17.19) yields the joint conditional density function of T_1 and T_2, given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m$, as

$$f^{T_1,T_2|\mathbf{A}^{\mathscr{R}}}(t_1, t_2|\mathbf{a}_m)$$
$$= k(\mathbf{a}_m) t_2^{m-1} \mathrm{e}^{m t_1 t_2 + a_{\bullet m} t_2} \exp\Big\{\sum_{i=1}^{m} (R_i + 1)\mathrm{e}^{a_i t_2 + t_1 t_2}\Big\}, t_1 \in \mathbb{R}, t_2 > 0, \quad (17.21)$$

where $a_{\bullet m} = \sum_{j=1}^m a_j$ and $k(\mathbf{a}_m)$ is the normalizing constant. In order to compute the confidence intervals for μ, ϑ and a quantile ξ_p, the marginal conditional cumulative distribution functions have to be calculated. Viveros and Balakrishnan [875] provided expressions for $F^{T_i|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_m}$, $i \in \{p, 1, 2\}$, $p \in (0, 1)$. They have shown that these distributions can be expressed as

$$F^{T_i|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_m}(t) = \frac{H_i(t|\mathbf{a}_m)}{H_i(\infty|\mathbf{a}_m)},$$

where $H_i(\infty|\mathbf{a}_m) = \lim_{t\to\infty} H_i(t|\mathbf{a}_m)$, $i \in \{p, 1, 2\}$, $p \in (0, 1)$. For the location parameter μ, this expression reads

$$H_1(t|\mathbf{a}_m) = \int_0^\infty \frac{t_2^{m-1}\mathrm{e}^{a_{\bullet m} t_2} \mathsf{IG}\big(\sum_{i=1}^m (R_i + 1)\mathrm{e}^{(a_i+t)t_2}; m\big)}{\big\{\sum_{i=1}^m (R_i + 1)\mathrm{e}^{a_i t_2}\big\}^m} dt_2.$$

The scale case leads to a simpler expression for the conditional cumulative distribution function. Since t_1 can be integrated out in (17.21), the following expression results:

$$H_2(t|\mathbf{a}_m) = \int_0^t \frac{t_2^{m-1}\mathrm{e}^{a_{\bullet m} t_2}}{\big\{\sum_{i=1}^m (R_i + 1)\mathrm{e}^{a_i t_2}\big\}^m} dt_2.$$

For a quantile ξ_p, $p \in (0, 1)$, the resulting function $H_p(\cdot|\mathbf{a}_m)$ is given by

$$H_p(z|\mathbf{a}_m) = \int_0^\infty \frac{1 - \mathsf{IG}\big(\zeta_p(t_2)\mathrm{e}^{-z t_2}; m\big)}{\zeta_p^m(t_2)} t_2^{m-1} \mathrm{e}^{a_{\bullet m} t_2} dt_2,$$

where $\zeta_p(t_2) = -\log(1-p)\sum_{i=1}^m (R_i + 1)\mathrm{e}^{a_i t_2}$. These quantities have to be evaluated numerically as pointed out by Viveros and Balakrishnan [875]. They proposed an application of Simpson's rule to accomplish the necessary numerical integrations. The MLEs can be computed by a Newton–Raphson procedure (see Sect. 12.7.4).

i	1	2	3	4	5	6	7	8
$x_{i:8:19}$	-1.6608	-0.2485	-0.0409	0.2700	1.0224	1.5789	1.8718	1.9947

Table 17.5 Log-times of Nelson's progressively Type-II censored data with censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$

Example 17.2.1. For Nelson's progressively Type-II censored data 1.1.5, Viveros and Balakrishnan [875] applied the above method to the log-times of the observations (see Table 17.5) (see also Balakrishnan and Aggarwala [86, Example 9.1]). Assuming a location–scale model of extreme value distributions, they obtained the maximum likelihood estimates by numerical optimization:

$$\widehat{\mu} = 2.222, \qquad \widehat{\vartheta} = 1.026.$$

For a given level $1-\alpha = 95\,\%$, they found by numerical integration that $F^{T_1|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}(-1.76) = 0.025$ and $F^{T_1|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}(0.574) = 0.975$. Hence, an exact 95 % confidence interval for the location parameter μ is obtained from the equations

$$-1.76 \le \frac{\widehat{\mu}-\mu}{\widehat{\vartheta}} \le 0.574$$

as $1.63 \le \mu \le 4.03$. Similarly, they got for the scale parameter ϑ with $F^{T_2|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}(0.472) = 0.5$ and $F^{T_2|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}(1.418) = 0.95$ the exact 90 % confidence interval $[0.72, 2.18]$. For the 0.1th quantile $\xi_{0.1}$, they established the maximum likelihood estimate $\widehat{\xi}_{0.1} = \widehat{\mu} + F^{\leftarrow}(0.1)\widehat{\vartheta} = -0.087$. By numerical computations, they obtained the quantiles of $F^{T_{0.1}|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}$ from

$$F^{T_{0.1}|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}(-4.64) = 0.025 \quad \text{and} \quad F^{T_{0.1}|\mathbf{A}^{\mathscr{R}}=\mathbf{a}_8}(-1.345) = 0.975.$$

This yields the 95 % confidence interval $[-2.54, 0.84]$ for $\xi_{0.1}$. Finally, they considered the reliability $R(t_0)$ at mission time $t_0 = \log 2 = 0.693$. The resulting maximum likelihood estimate is $\widehat{R}(t_0) = 0.798$. For $p_0 = p_0(\mathbf{x}) = 0.637$, they found that $\ell_{p_0}(\mathbf{x}) = t_0$. Therefore, $(0.637, 1]$ becomes an upper 95 % confidence interval for $R(t_0)$.

Remark 17.2.2. As pointed out by Balakrishnan and Aggarwala [86], the exact 90 % confidence interval $[0.72, 2.18]$ for the scale parameter ϑ is quite close to the unconditional 90 % confidence interval for ϑ determined by Nelson [676, pp. 230] from the complete data set.

Remark 17.2.3. The confidence intervals obtained for the location–scale family of extreme value distributions can be transformed to confidence intervals for the two-parameter Weibull distribution. For instance, a 90 % confidence interval for the shape parameter $\beta = 1/\vartheta$ of the associated Weibull distribution is given by

$[0.46, 1.38]$. Similar to Example 17.1.10, the approach of Wu [904] yields a 90 % confidence interval $[0.3798, 1.5665]$ for β. Applying the method of Wang et al. [890], we end up with the 90 % confidence interval $[0.4492, 1.3677]$ which is quite close to the conditional confidence interval. The confidence region computed from Wu [904] comprises both of them and is wider.

Notice that all these intervals include the parameter $\beta = 1$. This provides some evidence that an exponential distribution might be appropriate to model the original progressively Type-II censored data.

17.2.4 Log-Gamma Distribution

Lin et al. [602] considered conditional inference for a location–scale family of three-parameter log-gamma distributions with known shape parameter α. The standard member is given by the density function

$$f(t) = \frac{\alpha^{\alpha-1/2}}{\Gamma(\alpha)} \exp\left\{\sqrt{\alpha}t - \alpha e^{t/\sqrt{\alpha}}\right\}, \quad t \in \mathbb{R}, \alpha > 0.$$

This distribution is deduced from (12.40) by a linear transformation of the random variables, i.e., $\sqrt{\alpha}(X - \log \alpha)$ (see, e.g., Lin et al. [603]). When $\alpha = 1$, the extreme value distribution is included as a particular case.

Lin et al. [602] established conditional confidence intervals for the location and scale parameters, quantiles, and reliability. The approach is along the lines of the extreme value distribution, but the resulting expressions are quite complicated. Therefore, we abstain from presenting details on the representations of the density functions. Lin et al. [602] conducted an extensive Monte Carlo simulation to compare the conditional confidence intervals with unconditional confidence intervals obtained by simulation.

Example 17.2.4. From data reported by Lieblein and Zelen [595] (see also Example 15.2.1), Lin et al. [602] generated several progressively Type-II censored data sets. The log-gamma model was fitted to the logarithms of the original data by Lawless [573] who proposed this distribution to study the effect of departures from a Weibull or log-normal model (see also Lawless [575, pp. 249–250]). The data are also analyzed in Balakrishnan and Chan [90, 91]. For illustration, we consider data set H generated in Lin et al. [602]:

2.884, 3.991, 4.017, 4.217, 4.229, 4.229, 4.232, 4.432,
4.534, 4.591, 4.655, 4.662, 4.851, 4.852, 5.156.

It consists of $m = 15$ observations. Eight measurements were progressively Type-II censored according to the first-step censoring plan $\mathscr{O}_{15} = (8, 0^{*14})$. Choosing $\alpha = 2$, the MLEs $\widehat{\mu} = 4.5365$ and $\widehat{\vartheta} = 0.4100$ result. The corresponding 95 % confidence intervals for μ and ϑ are given by $[4.3078, 4.7817]$ and $[0.3046, 0.6603]$, respectively. Results for other values of α and other censoring plans are presented by Lin et al. [602]. Moreover, a comparison with unconditional confidence intervals is also carried out.

17.2.5 Pareto Distribution

Conditional inference for Pareto distribution has been addressed by Aggarwala and Childs [15] (see also Balakrishnan and Aggarwala [86, Sect. 9.7]). They considered a location–scale family of Pareto distributions with standard member

$$F(t) = 1 - t^{-\alpha}, \quad t \geq 1,$$

where $\alpha > 0$ is a known parameter. Since the MLEs of μ and ϑ are not explicitly available in this case (see Sect. 12.5 with a different parametrization), Aggarwala and Childs [15] based their conditional inference on the BLUEs $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$ of μ and ϑ as given in Theorem 11.2.4. These are easily seen to be equivariant [see the general representation of BLUEs in (11.3)]. Since conditional confidence intervals are invariant w.r.t. the choice of the equivariant estimators, this yields a reasonable choice for computational purposes. The conditional density function given in (17.19) exhibits a quite simple form. Specifically, we have

$$\begin{aligned} &f^{T_1,T_2|\mathbf{A}^{\mathscr{R}}}(t_1, t_2|\mathbf{a}_m) \\ &\qquad = k(\mathbf{a}_m)\frac{t_1^m}{t_2^{\alpha n+1}} \prod_{i=1}^m \frac{1}{(a_i + t_1)^{\alpha(R_i+1)+1}}, t_1 \geq -a_1, (a_1 + t_1)t_2 > 1. \end{aligned}$$

The marginal conditional density function of T_1, given $\mathbf{A}^{\mathscr{R}} = \mathbf{a}_m$, is obtained via integration w.r.t. t_2, i.e.,

$$f^{T_1|\mathbf{A}^{\mathscr{R}}}(t_1|\mathbf{a}_m) = k(\mathbf{a}_m)\frac{\alpha^{m-1}(a_1 + t_1)^{\alpha n}}{n} \prod_{i=1}^m \frac{1}{(a_i + t_1)^{\alpha(R_i+1)+1}}, \ t_1 \geq -a_1.$$

$f^{T_2|\mathbf{A}^{\mathscr{R}}}(t_2|\mathbf{a}_m)$ has to be evaluated numerically.

Example 17.2.5. For the simulated Pareto data given in Example 11.2.7 with given $\alpha = 3$, $\mu = 0$, and $\vartheta = 5$, the estimates $\widehat{\mu}_{\mathsf{LU}} = 1.87680$ and $\widehat{\vartheta}_{\mathsf{LU}} = 3.16207$ result. For $\alpha = 0.1$, the necessary quantiles of $F^{T_1|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a})$ are given by $\ell_1 = -0.5233$ and $\ell_2 = 2.2215$ so that $\ell_1 \leq T_1 \leq \ell_2$ leads to the 90 % confidence interval $[-5.148, 3.531]$ for μ. In case of the location parameter, the quantiles $\ell_1 = 0.3155$ and $\ell_2 = 2.049$ result. The corresponding 90 % confidence interval is given by $[1.543, 10.021]$.

17.2.6 Laplace Distribution

Childs and Balakrishnan [258] developed conditional inference for the Laplace distribution as given in (12.25). The resulting expressions are quite involved. For completeness, we present only the necessary expressions, and for more details, one may refer to Childs and Balakrishnan [258]. The case of Type-II censored data is addressed in Childs and Balakrishnan [256]. For complete samples, Kappenman [508] has constructed conditional confidence intervals for μ and ϑ.

The conditional density function $f^{T_1,T_2|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a}_m)$ is given by

$$f^{T_1,T_2|\mathbf{A}^{\mathscr{R}}}(t_1,t_2|\mathbf{a}_m) = k(\mathbf{a}_m)2^{m-n}t_2^{m-1}\mathrm{e}^{-t_2\sum_{j=1}^m |t_1+a_j|} \times \sum_{i=0}^m \mathbb{1}_{(-a_{i+1},-a_i]}(t_1)\mathrm{e}^{-t_2\sum_{j=i+1}^m R_j(t_1+a_j)} \prod_{j=1}^i \left[2-\mathrm{e}^{t_2(t_1+a_j)}\right]^{R_j}, \quad t_1\in\mathbb{R}, t_2>0.$$

The constant $k(\mathbf{a}_m)$ can be explicitly calculated. For $\mathbf{k}_j = (k_1,\ldots,k_j) \in \mathbb{N}^j$, $j\geq 1$, we introduce the quantities

$$\kappa_1(j,\mathbf{k}_j) = \gamma_{j+1} - \sum_{i=1}^j (k_i+1),$$

$$\kappa_2(j,\mathbf{k}_j) = \sum_{i=j+1}^m (R_i+1)a_i - \sum_{i=1}^j (k_i+1)a_i,$$

$$\kappa_3(j,\mathbf{k}_j) = \left[\prod_{i=1}^j \binom{R_i}{k_i}\frac{(-1)^{k_i}}{2^{k_i}}\right]2^{-\sum_{i=j+1}^m R_i}.$$

For $j=0$, these expressions are given by

$$\kappa_1(0,\mathbf{k}_0) = n, \quad \kappa_2(0,\mathbf{k}_0) = \sum_{i=1}^m (R_i+1)a_i, \quad \kappa_3(0,\mathbf{k}_0) = 2^{m-n}.$$

With

$$\mathcal{M}_j = \{\mathbf{k}_j \mid 1\leq k_i\leq R_i, 1\leq i\leq j\} \text{ and } \mathcal{M}_j^0 = \{\mathbf{k}_j \in \mathcal{M}_j \mid \kappa_1(j,\mathbf{k}_j)=0\},$$

the normalizing constant exhibits the explicit representation

$$k(\mathbf{a}_m)^{-1} = H(\Gamma(m-1),\Gamma(m-1),\Gamma(m-1)),$$

where

$$H(x, y, z) = \sum_{j=0}^{m} \sum_{\mathbf{k}_j \in \mathcal{M}_j \setminus \mathcal{M}_j^0} \frac{\kappa_3(j, \mathbf{k}_j)}{\kappa_1(j, \mathbf{k}_j)} \left\{ \frac{x}{\left[\kappa_2(j, \mathbf{k}_j) - \kappa_1(j, \mathbf{k}_j) a_{j+1}\right]^{m-1}} - \frac{y}{\left[\kappa_2(j, \mathbf{k}_j) - \kappa_1(j, \mathbf{k}_j) a_j\right]^{m-1}} \right\} + z \sum_{j=0}^{m} \sum_{\mathbf{k}_j \in \mathcal{M}_j^0} \frac{\kappa_3(j, \mathbf{k}_j)}{\left[\kappa_2(j, \mathbf{k}_j)\right]^m} (a_{j+1} - a_j). \tag{17.22}$$

The conditional density function $f^{T_1|\mathbf{A}^{\mathscr{R}}}(\cdot|\mathbf{a}_m)$ can be written as

$$f^{T_1|\mathbf{A}^{\mathscr{R}}}(t_1|\mathbf{a}_m) = \sum_{j=1}^{m} k(\mathbf{a}_m) \mathbb{1}_{(-a_{j+1}, -a_j]}(t_1) \sum_{\mathbf{k}_j \in \mathcal{M}_j} \frac{\Gamma(m) \kappa_3(j, \mathbf{k}_j)}{[\kappa_1(j, \mathbf{k}_j) t_1 + \kappa_2(j, \mathbf{k}_j)]^m} + \mathbb{1}_{(-a_1, \infty)}(t_1) k(\mathbf{a}_m) \frac{\Gamma(m) 2^{m-n}}{[n t_1 + \kappa_2(0, \mathbf{k}_0)]^m}.$$

Denoting by $k \in \{0, \dots, m\}$ the unique number $-a_{k+1} < t_1 \leq -a_k$, the conditional cumulative distribution function can be expressed as

$$F^{T_1|\mathbf{A}^{\mathscr{R}}}(t_1|\mathbf{a}_m) = G_k(t_1|\mathbf{a}_m) + \sum_{j=k+1}^{m} G_k(-a_j|\mathbf{a}_m), \tag{17.23}$$

where

$$G_j(t|\mathbf{a}_m) = k(\mathbf{a}_m) \Gamma(m) \times \sum_{\mathbf{k}_j \in \mathcal{M}_j \setminus \mathcal{M}_j^0} \frac{\kappa_3(j, \mathbf{k}_j)}{\kappa_1(j, \mathbf{k}_j)(1-m)} \left\{ \frac{1}{\left[\kappa_2(j, \mathbf{k}_j) + \kappa_1(j, \mathbf{k}_j) t\right]^{m-1}} - \frac{1}{\left[\kappa_2(j, \mathbf{k_j}) - \kappa_1(j, \mathbf{k}_j) a_{j+1}\right]^{m-1}} \right\} + k(\mathbf{a}_m) \Gamma(m) \sum_{\mathbf{k}_j \in \mathcal{M}_j^0} \frac{\kappa_3(j, \mathbf{k}_j)}{\left[\kappa_2(j, \mathbf{k}_j)\right]^m} (t + a_{j+1}).$$

The desired confidence interval for the location parameter μ is given by the equations

$$F^{T_1|\mathbf{A}^{\mathscr{R}}}(L_1|\mathbf{a}_m) = 1 - \frac{\alpha}{2} \quad \text{and} \quad F^{T_1|\mathbf{A}^{\mathscr{R}}}(L_2|\mathbf{a}_m) = \frac{\alpha}{2},$$

with the confidence limits $L_1, L_2 \in \mathbb{R}$. Finally, the confidence interval

$$\left[\widehat{\mu} - L_1\widehat{\vartheta}, \widehat{\mu} - L_2\widehat{\vartheta}\right]$$

results, where $\widehat{\mu}$ and $\widehat{\vartheta}$ are the MLEs of μ and ϑ.

For the scale parameter, the conditional density function

$$\begin{aligned} f^{T_2|\mathbf{A}^{\mathscr{R}}}(t_2|\mathbf{a}_m) = k(\mathbf{a}_m) \sum_{j=0}^{m} \sum_{\mathbf{k}_j \in \mathcal{M}_j \setminus \mathcal{M}_j^0} & \frac{\kappa_3(j,\mathbf{k}_j)}{\kappa_1(j,\mathbf{k}_j)} \\ & \cdot t_2^{m-2} \left\{ \mathrm{e}^{-t_2\left[\kappa_2(j,\mathbf{k}_j) - \kappa_1(j,\mathbf{k}_j)a_{j+1}\right]} - \mathrm{e}^{-t_2\left[\kappa_2(j,\mathbf{k}_j) - \kappa_1(j,\mathbf{k}_j)a_j\right]} \right\} \\ & + k(\mathbf{a}_m) \sum_{j=0}^{m} \sum_{\mathbf{k}_j \in \mathcal{M}_j^0} \kappa_3(j,\mathbf{k}_j)(a_{j+1} - a_j) t_2^{m-1} \mathrm{e}^{-t_2\kappa_2(j,\mathbf{k}_j)} \end{aligned}$$

is needed. Using the function H defined in (17.22) and the incomplete gamma function $\mathsf{IG}(\cdot;\alpha)$, the corresponding conditional cumulative distribution function can be written as

$$\begin{aligned} F^{T_2|\mathbf{A}^{\mathscr{R}}}(t|\mathbf{a}_m) = k(\mathbf{a}_m) H\Big(& \mathsf{IG}\big([\kappa_2(j,\mathbf{k}_j) - \kappa_1(j,\mathbf{k}_j)a_{j+1}]t, m-1\big), \\ & \mathsf{IG}\big([\kappa_2(j,\mathbf{k}_j) - \kappa_1(j,\mathbf{k}_j)a_j]t, m-1\big), \mathsf{IG}\big(\kappa_2(j,\mathbf{k}_j)t, m\big)\Big). \end{aligned}$$

Solving the equations analogous to (17.23), the desired confidence bounds can be obtained.

Example 17.2.6. Childs and Balakrishnan [258] applied the method to simulated data taken from Balakrishnan and Aggarwala [86] (see Data B.1.3) which have been simulated from a Laplace(25, 5)-distribution with $n = 20$ and censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$, $m = 10$. The MLEs are given by $\widehat{\mu} = 26.31069$ and $\widehat{\vartheta} = 2.67091$, respectively. For the location case, Childs and Balakrishnan [258] obtained the bounds $L_1 = 0.66085$ and $L_2 = -0.77611$ which gives the 95 % confidence interval $[24.55, 28.38]$. In the scale case, the bounds are given by $L_1 = 1.60544$ and $L_2 = 0.45905$ yielding the 95 % confidence interval $[1.66, 5.82]$ for ϑ.

Example 17.2.7. For the data given in Table 12.3, Childs and Balakrishnan [258] postulated a Laplace model and considered conditional inference for the location and scale parameters. According to Example 12.6.9, the MLEs of μ and ϑ are given by $\widehat{\mu} = 1033.81667$ and $\widehat{\vartheta} = 182.86675$. A 95 % confidence interval for μ results from the bounds $L_1 = 0.38875$ and $L_2 = -0.61026$ as

[962.73, 1145.41]. From $L_1 = 1.42632$ and $L_2 = 0.60950$, a 95 % conditional confidence interval for ϑ is obtained as [128.21, 300.03].

17.2.7 Other Distributions

Maswadah [641] considered conditional inference for inverse Weibull distributions with cumulative distribution function

$$F(t) = \exp\Big\{ - (\alpha t)^{-\beta} \Big\}, \quad t \geq 0, \alpha, \beta > 0,$$

in terms of generalized order statistics. The resulting expressions for a progressively Type-II censored sample involve integrals which have to be evaluated numerically. For illustration, the author has presented a numerical study.

17.3 Asymptotic Confidence Intervals

Asymptotic statistical intervals are based on the asymptotic distribution of standardized estimates. In location–scale models, one often uses the pivotal quantities

$$\zeta_1 = \frac{\widehat{\mu} - \mu}{\widehat{\vartheta} \cdot \sqrt{I_{11}}}, \quad \zeta_2 = \frac{\widehat{\mu} - \mu}{\vartheta \cdot \sqrt{I_{11}}}, \quad \zeta_3 = \frac{\widehat{\vartheta} - \vartheta}{\widehat{\vartheta} \cdot \sqrt{I_{22}}}$$

which, under certain regularity conditions, are considered as asymptotically standard Gaussian distributed. Here, I_{jj}, $j = 1, 2$, are the respective components of the asymptotic variance–covariance matrix. For instance, Balakrishnan and Kannan [104] have considered a location–scale family of logistic distributions with standard cumulative distribution function as in (12.39), i.e.,

$$F(t) = \Big[1 + \mathrm{e}^{-\pi t/\sqrt{3}} \Big]^{-1}, \quad t \in \mathbb{R}.$$

They simulated the probability coverages

$$P(|\zeta_i| \leq 1.65) \quad \text{and} \quad P(|\zeta_i| \leq 1.96), \quad i = 1, 2, 3,$$

by Monte Carlo simulations leading to approximate confidence intervals with levels 90 and 95 %, respectively. Tables are provided for different censoring schemes. Instead of the MLEs, approximate maximum likelihood estimators are also used. Similar approaches are utilized by Balakrishnan et al. [134] for Gaussian distributions (see also Balakrishnan and Kim [107]), by Balakrishnan and Asgharzadeh [87] for a scaled half-logistic distribution, and by Asgharzadeh [61] for a generalized logistic distribution. Ng [681] considered a modified Weibull distribution with cumulative distribution function as in (12.35). He illustrated his results by the data presented in Table 12.4.

In the one-parameter case, a general result for the asymptotic distribution of a solution of the likelihood equation has been established by Lin and Balakrishnan [599]. Under some regularity conditions, they showed that the asymptotic distribution of a solution of the likelihood equation is Gaussian. This extends a result of Hoadley [442] to progressive censoring. For general progressively Type-II censored exponential data, Fernández [364] established the asymptotic normality of both the standardized MLE and BLUE of the scale parameter. Notice that in this case the likelihood equation has a unique solution as worked out by Balakrishnan et al. [130].

17.4 Prediction Intervals

17.4.1 Nonparametric Prediction Intervals

Using the mixture representation of progressively Type-II censored order statistics, Guilbaud [419] presented formulas for nonparametric two-sample prediction intervals based on progressively Type-II censored samples containing at least a specified number of observations in a future sample. This work has been followed up by Balakrishnan et al. [144] who established corresponding intervals based on two ordered pooled progressively Type-II censored samples. They computed exact prediction levels by applying the mixture representation of progressively Type-II censored order statistics. The approach has also been utilized by Beutner and Cramer [193] in the models of minimal repair and record values.

17.4.2 Parametric and Bayesian Prediction

Interval prediction issues with parametric models have been studied extensively in progressive Type-II censoring. Since the models and assumptions are quite different, we provide only a survey of the existing literature. The following presentation is organized w.r.t. the predicted object. In particular, we address the following three different issues:

(1) Prediction intervals for censored failure times;
(2) Prediction intervals for future observations in the same sample (this is a particular case of (1) in the sense that the lifetimes of the items removed in the final progressive censoring step are predicted);
(3) Prediction intervals for observations of an independent future sample from the same population.

The results are differentiated w.r.t. distributional assumptions and the method of prediction.

Prediction Intervals for Censored Failure Times

In Sect. 16.2, the point prediction of censored failure times is discussed extensively. Basak et al. [178] presented prediction intervals for the censored failure time $W_{j,k:R_j}$, $k = 1, \ldots, R_j$, $j = 1, \ldots, m$, assuming exponential and extreme value distributions, respectively. Using the mean squared predictive errors and standard errors, they established prediction intervals based on BLUP, MLP, MUP, and CMP (exponential distribution) and BLUP, MMLP, and CMP (extreme value distribution). In Basak and Balakrishnan [175], prediction intervals based on various predictors, including BLUP, MLP, and CMP, are discussed for the normal distribution. For Pareto distributions, we refer to Raqab et al. [746]. Generalized exponential and Rayleigh distributions are discussed in Madi and Raqab [627] and Raqab and Madi [745], respectively.

Extending the work of AL-Hussaini [27], Mohie El-Din and Shafay [652] considered Bayesian prediction intervals for the censored lifetimes assuming the lifetime distribution in (17.24).

Prediction Intervals for Future Observations in the Same Sample

Given a progressively censored sample $X^{\mathscr{R}}_{1:m:n}, \ldots, X^{\mathscr{R}}_{m:m:n}$ with $R_m > 0$, we might be interested in a predictive interval of the first failure time of the R_m items exceeding the final observation $X^{\mathscr{R}}_{m:m:n}$. For population cumulative distribution functions as in (12.10), i.e., for $F(t) = 1 - \exp\{-(d(t) - \mu)/\vartheta\}$, the following result holds. We use the notation $W_{m,1:R_m}$, introduced in Sect. 16.2, for this future observation.

Theorem 17.4.1. Let $\alpha \in (0, 1)$, $\mathbf{X}^{\mathscr{R}}$ be a progressively Type-II censored sample based on a population cumulative distribution functions as in (12.10) with strictly increasing d and $\mathscr{R}$ be a censoring scheme with $R_m > 0$. Then, a $100(1-\alpha)\%$ prediction interval for $W_{m,1:R_m}$ is given by

$$\mathscr{K} = \Big[d^{-1}\Big\{d(X^{\mathscr{R}}_{m:m:n}) + \tau(\mathbf{X}^{\mathscr{R}})\mathsf{F}_{\alpha/2}(2, 2(m-1))\Big\},$$
$$d^{-1}\Big\{d(X^{\mathscr{R}}_{m:m:n}) + \tau(\mathbf{X}^{\mathscr{R}})\mathsf{F}_{1-\alpha/2}(2, 2(m-1))\Big\}\Big],$$

where $\tau(\mathbf{X}^{\mathscr{R}}) = \frac{1}{(m-1)R_m} \sum_{i=2}^{m} (R_i + 1)(d(X^{\mathscr{R}}_{i:m:n}) - d(X^{\mathscr{R}}_{1:m:n}))$.

Proof. First, $W_{m,1:R_m}$ can be seen as a progressively Type-II censored order statistic in a progressively Type-II censored sample with concatenated censoring scheme $(R_1, \ldots, R_{m-1}, 0^{*R_m+1})$. In fact, the experimenter abstains from the right censoring and observes all failures exceeding $X^{\mathscr{R}}_{m:m:n}$. Then, with $V = \frac{d(W_{m,1:R_m}) - d(X^{\mathscr{R}}_{m:m:n})}{\tau(\mathbf{X}^{\mathscr{R}})}$, we have

$$P(W_{m,1:R_m} \in \mathscr{K}) = P\Big(\mathsf{F}_{\alpha/2}(2, 2(m-1)) \le V \le \mathsf{F}_{1-\alpha/2}(2, 2(m-1))\Big)$$
$$= 1 - \alpha,$$

since V has an F-distribution with degrees of freedom 2 and $2(m-1)$. This results from the comments in Remark 12.1.11 and the independence of the spacings $d(X^{\mathscr{R}}_{j:m:n}) - d(X^{\mathscr{R}}_{j-1:m:n})$, $2 \le j \le m$, and $d(W_{m,1:R_m}) - d(X^{\mathscr{R}}_{m:m:n})$. □

Remark 17.4.2. The above result can be found in Wu [909] for exponential distributions who also considered prediction intervals for future spacings $W_{m,1:R_m} - X^{\mathscr{R}}_{m:m:n}$ and $W_{m,j:R_m} - W_{m,j-1:R_m}$, $j = 2, \dots, R_m$. For Pareto distributions, the corresponding result is given in Wu [908]. Moreover, the results for the ratios $W_{m,1:R_m}/X^{\mathscr{R}}_{m:m:n}$ and $W_{m,j:R_m}/W_{m,j-1:R_m}$, $j = 2, \dots, R_m$, have also been established in this work.

Bayesian prediction intervals have also been proposed for future failure times. A general setting has been addressed in Abdel-Aty et al. [2]. They discussed Bayesian prediction intervals in the model of generalized order statistics based on a cumulative distribution function

$$F(t) = 1 - \mathrm{e}^{-\lambda_\theta(t)}, \quad t > 0, \tag{17.24}$$

where $\lambda_\theta : \mathbb{R} \longrightarrow \mathbb{R}$ denotes an appropriately chosen function. Moreover, they allowed the sample to be multiply censored. The prior is chosen according to

$$\pi_{\mathbf{a}}(\theta) \propto c_{\mathbf{a}}(\theta) \exp\{-d_{\mathbf{a}}(\theta)\}.$$

The results are also presented in the particular case of Pareto distributions. Results for Weibull distributions are discussed in Huang and Wu [463]. Linear hazard rate distributions [see (15.10)] are addressed in Lin et al. [604] for a general progressively Type-II censored sample.

Prediction Intervals for Observations of an Independent Future Sample from the Same Population

Prediction intervals for future observations of an independent sample from the same population have also been studied in the literature. In particular, Bayesian approaches have been studied extensively by choosing an appropriate prior. Since the results follow along the same lines, we only sketch the results here. For more details, we refer to the cited references. Ghafoori et al. [397] discussed the family (17.24) with parameter vector $\boldsymbol{\theta}$ and obtained prediction bounds and Bayes predictive estimators for the kth-future order statistic $Y_{k:N}$. They applied their results to one- and two-parameter Weibull and Pareto families (see also Ali Mousa [34]). Wu et al. [915] presented HPD prediction intervals for Rayleigh distributions.

For Weibull distributions, we refer to Huang and Wu [463] and Soliman et al. [816]. Prediction intervals for Burr-XII distributions are given in Ali Mousa and Jaheen [37]. One- and two-parameter Gompertz distributions are discussed in Jaheen [474]. For exponentiated modified Weibull distributions, see Klakattawi et al. [535].

Mohie El-Din and Shafay [652] presented Bayesian prediction intervals for a future sample of progressively Type-II censored order statistics $\mathbf{Y}^{\mathscr{S}}$ when the baseline cumulative distribution function is specified by (17.24). Ali Mousa and Al-Sagheer [35] considered the same problem for Rayleigh distributions.

17.5 Nonparametric Tolerance Intervals

Tolerance intervals are supposed to cover a certain proportion $p \in (0,1)$ of the probability mass of the population distribution. Given a confidence level $1-\alpha$, a two-sided tolerance interval $[X_{i:m:n}, X_{j:m:n}]$, $i < j$, is defined by the condition

$$P\big(F(X_{j:m:n}) - F(X_{i:m:n}) \geq p\big) \geq 1-\alpha, \tag{17.25}$$

where F denotes the population cumulative distribution function. Nonparametric tolerance intervals based on order statistics of an IID sample are well known (see, e.g., Krishnamoorthy and Mathew [553]). For progressively Type-II censored order statistics, it has been first addressed by Guilbaud [419] and further discussed by Balakrishnan et al. [144]. It turns out that the right hand side of Eq. (17.25) is given by

$$P\big(F(X_{j:m:n}) - F(X_{i:m:n}) \geq p\big) = 1 - F^{U_{j:m:n}-U_{i:m:n}}(p),$$

so that the cumulative distribution function of the generalized spacing $U_{j:m:n}-U_{i:m:n}$ from uniform progressively Type-II censored order statistics is necessary to compute the tolerance interval. The cumulative distribution function can be taken from (2.39) (see also Lemma 3 in Kamps and Cramer [503]). Guilbaud [419] applied the mixture representation of progressively Type-II censored order statistics to calculate these probabilities.

Remark 17.5.1. Pradhan [726] studied the performance of an approximate tolerance interval for the population distribution when the progressively Type-II censored lifetime data is observed from a k-unit parallel system. Conditional tolerance intervals for Pareto distributions are presented in Aggarwala and Childs [15].

17.6 Highest Posterior Density Credible Intervals

Based on the posterior density function $\pi_{\mathbf{a}}^*(\boldsymbol{\theta}|\mathbf{x})$ in (15.1), highest posterior density (HPD) credible intervals can be defined. According to Berger [190, pp. 140], a

$(1-\alpha)$ credible set for θ is defined by $\mathcal{C}_\alpha = \{t \mid \pi_{\mathbf{a}}^*(t|\mathbf{x}) > c_{1-\alpha}\}$, where $c_{1-\alpha}$ satisfies the condition $P(\theta \in \mathcal{C}_\alpha|\mathbf{x}) = 1-\alpha$. For unimodal posterior density functions, a two-sided HPD credible interval $[\widehat{\ell}, \widehat{u}]$ is defined by the conditions

$$\int_{\widehat{\ell}}^{\widehat{u}} \pi_{\mathbf{a}}^*(t|\mathbf{x})dt = 1-\alpha \quad \text{and} \quad \pi_{\mathbf{a}}^*(\widehat{\ell}|\mathbf{x}) = \pi_{\mathbf{a}}^*(\widehat{u}|\mathbf{x}). \tag{17.26}$$

We shall now review the available results. The results are quite sensitive regarding both the underlying parametrization of the lifetime distributions and the prior distribution. This leads to many different results although the approach adapted is usually similar.

HPD credible intervals for exponential distributions and Weibull distributions with known shape parameter, as discussed in Sect. 15.1, have been established in Schenk et al. [783] and Kundu [557]. Assuming a gamma prior, the posterior distribution is a particular gamma distribution and thus unimodal. Therefore, HPD credible sets for α can be easily obtained. HPD credible intervals assuming a Rayleigh distribution have been considered in Wu et al. [915] and Kim and Han [529] (see also Dey and Dey [338] for progressive Type-II censoring with random removals). Wu et al. [915] presented a $(1-\alpha)$ HPD credible interval $[\widehat{\ell}, \widehat{u}]$ which can be directly obtained from (17.26) by solving the equations

$$\mathsf{IG}(\nu_1, a+m) - \mathsf{IG}(\nu_2, a+m) = 1-\alpha, \quad \left(\frac{\widehat{u}}{\widehat{\ell}}\right)^{2(a+m)+1} = \mathrm{e}^{\nu_1-\nu_2},$$

where $\nu_1 = (2\widehat{\ell}^2)^{-1}(b+W)$ and $\nu_2 = (2\widehat{u}^2)^{-1}(b+W)$. A HPD credible interval for $R(t)$ can be constructed analogously.

Example 17.6.1. For the data in Example 15.2.1, Wu et al. [915] reported the 90 % HPD credible intervals as $(0.4534, 0.6989)$ and $(0.0856, 0.3568)$, respectively.

Kundu [557] discussed Weibull distributions with unknown shape and scale parameters (see Sect. 15.1).

Example 17.6.2. In the case of Example 15.1.2, Kundu [557] computed the 95 % HPD credible intervals $(0.0165, 0.1642)$ and $(0.3344, 1.0043)$ for the shape and scale parameters, respectively.

Inverse Weibull distributions are discussed in Sultan et al. [827]. HPD credible intervals for competing risks data from Weibull distributions have been addressed in Kundu and Pradhan [563]. Mokhtari et al. [654] studied Type-I progressively hybrid censored data. Further results on Bayesian inference are also presented in these works. Thus, we refer to the results presented in Chap. 15.

Chapter 18
Progressive Type-I Interval Censored Data

18.1 Parametric Inference

Progressive Type-I interval censoring has been introduced in Aggarwala [11] (see p. 14 and Fig. 1.8). She assumed a continuous lifetime distribution $F_{\boldsymbol{\theta}}$ leading to the likelihood function

$$\mathsf{L}(\boldsymbol{\theta}) \propto \prod_{j=1}^{k} \big(F_{\boldsymbol{\theta}}(T_j) - F_{\boldsymbol{\theta}}(T_{j-1})\big)^{d_j} \overline{F}_{\boldsymbol{\theta}}^{R_j}(T_j), \tag{18.1}$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)' \in \Theta \subseteq \mathbb{R}^p$ denotes the parameter vector and $d_1, \ldots, d_k$ are realizations of the number of observed failures $D_1, \ldots, D_k$ in the inspection intervals (see (1.8)). $T_0 = -\infty < T_1 < \cdots < T_k$ are the censoring times, and $\mathscr{R} = (R_1, \ldots, R_k)$ is the effectively applied censoring scheme. Obviously, the equation

$$\sum_{j=1}^{k} (d_j + R_j) = n$$

holds.

Aggarwala [11] considered likelihood inference for the scale parameter ϑ when the baseline distribution is exponential. In particular, assuming an $\mathsf{Exp}(\vartheta)$-distribution, the likelihood is proportional to ($T_0 = 0$)

$$\mathsf{L}(\vartheta) \propto \prod_{j=1}^{k} \left[1 - \mathrm{e}^{-(T_j - T_{j-1})/\vartheta}\right]^{d_j} \mathrm{e}^{-(d_j T_{j-1} + R_j T_j)/\vartheta}.$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_18,

This yields the log-likelihood function

$$\ell(\vartheta) = \text{const} + \sum_{j=1}^{k} d_j \log\Big[1 - \mathrm{e}^{-(T_j - T_{j-1})/\vartheta}\Big] - \frac{1}{\vartheta}\Big[\sum_{j=2}^{k} d_j T_{j-1} + \sum_{j=1}^{k} R_j T_j\Big]. \tag{18.2}$$

In general, an explicit expression for the MLE is not available. However, since

$$\frac{\partial \ell}{\partial \vartheta}(\vartheta) = \frac{1}{\vartheta^2}\Big(-\sum_{j=1}^{k} \frac{d_j(T_j - T_{j-1})}{\mathrm{e}^{(T_j - T_{j-1})/\vartheta} - 1} + \sum_{j=2}^{k} d_j T_{j-1} + \sum_{j=1}^{k} R_j T_j\Big),$$

we get $\lim_{\vartheta\to 0} \vartheta^2 \cdot \frac{\partial \ell}{\partial \vartheta}(\vartheta) = \sum_{j=2}^{k} d_j T_{j-1} + \sum_{j=1}^{k} R_j T_j > 0$ as well as $\lim_{\vartheta\to\infty} \vartheta^2 \cdot \frac{\partial \ell}{\partial \vartheta}(\vartheta) = -\infty$ provided that at least one failure is observed. Moreover, it is easily seen that $\vartheta^2 \cdot \frac{\partial \ell}{\partial \vartheta}(\vartheta)$ is a strictly decreasing function in ϑ showing that a unique solution of the likelihood equation exists. This has to be computed numerically in general. Cheng et al. [255] proposed an algorithm to compute the MLE which is based on "equivalent quantities" as discussed in Tan [836].

For constant inspection intervals, i.e., $T_i = i \cdot T$, $0 \le i \le k$, for some $T > 0$, the log-likelihood function in (18.2) can be written as

$$\ell(\vartheta) = \text{const} + \log\Big[1 - \mathrm{e}^{-T/\vartheta}\Big] d_{\bullet k} - \frac{1}{\vartheta}\Big[\sum_{j=2}^{k} (j-1) d_j T + \sum_{j=1}^{k} j R_j T\Big].$$

Then, as shown in Aggarwala [11], an explicit solution exists. The corresponding estimator is given by

$$\widehat{\vartheta} = \frac{T}{\log\Big[1 + \frac{D_{\bullet k}}{\sum_{j=2}^{k}(j-1)D_j + \sum_{j=1}^{k} jR_j}\Big]}.$$

Aggarwala [11] additionally presented a simulation algorithm to generate progressively Type-I interval censored data (see Sect. 8.3). She illustrated her results by some simulated data for constant inspection intervals. Furthermore, interval estimation and hypothesis testing procedures have also been discussed.

Ng and Wang [687] considered progressively Type-I interval censored data from a Weibull population $\mathsf{Weibull}(1/\vartheta, \beta)$. In order to estimate the parameters, they proposed several methods as follows:

(i) Mid-point approximation: It is assumed that failures observed in an interval $(T_{j-1}, T_j]$ occurred at the mid-point $m_j = \frac{1}{2}(T_{j-1} + T_j)$. Moreover, the units withdrawn at the jth step are considered with failure time T_j. This yields a pseudo-complete data set $y_1, \ldots, y_n$. Then, the likelihood equations for the Weibull distribution are solved assuming the complete sample $y_1, \ldots, y_n$. Similarly, the results for progressively Type-I censored data from the Weibull distribution, as given in Sect. 13.2, have been applied to the data $(m_j^{*D_j})$, $1 \le j \le k$.

Interval in months	Number at risk	Number of withdrawals
$[0, 5.5)$	112	1
$[5.5, 10.5)$	93	1
$[10.5, 15.5)$	76	3
$[15.5, 20.5)$	55	0
$[20.5, 25.5)$	45	0
$[25.5, 30.5)$	34	1
$[30.5, 40.5)$	25	2
$[40.5, 50.5)$	10	3
$[50.5, 60.5)$	3	2
$[60.5, \infty)$	0	0

Table 18.1 Survival times for patients with plasma cell myeloma. Data taken from Ng and Wang [687]

(ii) Maximum likelihood estimation/EM-algorithm: After determining the likelihood equations from (18.1), Ng and Wang [687] proposed an EM-algorithm to compute the MLEs. As for the exponential case, Cheng et al. [255] developed an alternative algorithm. They demonstrated that it is more efficient w.r.t. convergence rate. The results for the data in Table 18.1 correspond to those of Ng and Wang [687]. Wu et al. [921] and Lin et al. [609] considered MLEs for $\mathsf{Weibull}(\vartheta^{\beta}, \beta)$-lifetimes.

(iii) Method of moments: Using the fact that the νth moment of the Weibull distribution $\mathsf{Weibull}(1/\vartheta, \beta)$ is given by

$$\vartheta^{\nu/\beta}\Gamma\Big(1+\frac{\nu}{\beta}\Big),$$

they established an iterative procedure to compute the method of moments estimators.

(iv) Estimation based on Weibull probability plot: Using the Kaplan-Meier-type estimator

$$\widehat{F}(T_j) = 1 - \prod_{i=1}^{j}(1-\widehat{p}_i),$$

with $\widehat{p}_i = \frac{D_i}{n-D_{\bullet i-1}-R_{\bullet i-1}}$, $i = 1, \ldots, k$, estimators result by fitting the function $h(w) = -\log\vartheta + \beta w$ to the data (w_i, z_i), where $w_i = \log T_i$ and $z_i = \log(-\log(1-\widehat{F}(T_i)))$, $1 \le i \le k$.

(v) One-step approximate estimators using Newton–Raphson method: Choosing an initial value for the parameters, one step of the Newton–Raphson procedure is computed.

The resulting estimators were compared based on an extensive simulation study. Moreover, the methods were illustrated by the data in Table 18.1 which originally was reported by the National Cancer Institute. These data have also been analyzed in Lawless [575].

Generalized exponential distributions with cumulative distribution function as in (12.41) have been addressed by Chen and Lio [252] (see also Peng and Yan [716]). Adopting the results of Ng and Wang [687], they applied the midpoint approximation. Moreover, they presented a version of the EM-algorithm to approximate the MLEs of the parameters. Further, they used the method of moments to compute estimators. They illustrated their results by the data set given in Table 18.1. Similar results are presented in Lio et al. [614] for the generalized Rayleigh distribution and in Xiuyun and Zaizai [931] for the gamma distribution. For the Gompertz–Makeham distribution, we refer to Teimouri and Gupta [841]. Log-normal distributions are discussed in Amin [44]. Beta kernel distributions in (12.45) are considered by Teimouri et al. [842].

A general Bayesian approach has been discussed in Lin and Lio [601] using an MCMC-process. It is illustrated for both the two-parameter Weibull distribution and the generalized exponential distribution.

Remark 18.1.1. The model of grouped progressively censored data with random removals has been discussed in Xiang and Tse [924] based on Weibull lifetimes.

18.2 Optimal Inspection Times

Optimal grouping or inspection times for interval censored data have been studied in Kulldorff [555] and Vasudeva Rao et al. [873].

Lin et al. [606] discussed progressively Type-I interval censored data from a log-normal population. In particular, they addressed the problem of determining optimal inspection times $T_1 < \cdots < T_k$ provided that the censoring scheme is obtained from the proportions $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_m)$ as in (1.9). The resulting solution is called optimally spaced (OS) inspection times. The objective function is defined in terms of the expected Fisher information matrix (or the asymptotic variance–covariance matrix). Denoting by

$$\mathscr{I}(\mathbf{D}^{\boldsymbol{\pi}}; \mu, \vartheta) = \begin{pmatrix} I_{\mu\mu} & I_{\mu\vartheta} \\ I_{\mu\vartheta} & I_{\vartheta\vartheta} \end{pmatrix}, \quad \mathbf{D}^{\boldsymbol{\pi}} = (D_1, \ldots, D_k), \tag{18.3}$$

the expected Fisher information matrix, Lin et al. [606] computed expressions for the components of $\mathscr{I}(\mathbf{D}^{\boldsymbol{\pi}}; \mu, \vartheta)$. Optimal inspection times are determined by maximizing the determinant of the Fisher information matrix, i.e., Lin et al. [606] addressed the problem

$$\max_{T_1 < \cdots < T_k} \det \mathscr{I}(\mathbf{D}^{\boldsymbol{\pi}}; \mu, \vartheta) = \max_{T_1 < \cdots < T_k} \left(I_{\mu\mu} I_{\vartheta\vartheta} - I_{\mu\vartheta}^2 \right),$$

for given values of the parameters μ and ϑ and censoring proportions $\boldsymbol{\pi}$. Moreover, they considered optimally equi-spaced (OES) inspection times defined by

$$T_i = \mu - \frac{(k - 2i + 1)T}{2}, \quad i = 1, \ldots, k,$$

with an optimally chosen constant inspection time T. The corresponding optimization problem becomes

$$\max_{T>0} \left(I_{\mu\mu} I_{\vartheta\vartheta} - I_{\mu\vartheta}^2 \right).$$

As an alternative criteria, Lin et al. [606] proposed maximization of the trace of either the expected Fisher information matrix or the variance–covariance matrix of the MLEs. A third method to determine inspection times is called equal probability (EP) inspection scheme. Here, the inspection times are chosen such that the expected number of failures falling in each inspected interval is the same. Results are obtained by applying a simulated annealing procedure as described in Corana et al. [280,281].

The same approach has been employed by Lin et al. [609] to obtain optimal inspection schemes for the two-parameter Weibull(ϑ^β, β)-distribution.

18.3 Optimal Progressive Interval Censoring Proportions

Utilizing the expected Fisher information matrix in (18.3), Lin et al. [606] also considered the optimal choice of the inspection proportions $\boldsymbol{\pi}$ in the log-normal case. They considered the optimization problem

$$\max_{\pi_1,\ldots,\pi_k} \left(I_{\mu\mu} I_{\vartheta\vartheta} - I_{\mu\vartheta}^2 \right),$$

subject to the constraint

$$ED_1 + \sum_{i=2}^{k} E(D_i | D_1 = \eta_1, R_1 = \phi_1, \ldots, D_{i-1} = \eta_{i-1}, R_{i-1} = \phi_{i-1}) = n \cdot h,$$

where

$$\begin{aligned} \eta_1 &= ED_1, \quad \phi_1 = E(R_1 | D_1 = \eta_1), \qquad \text{and, for } 2 \le i \le k, \\ \eta_i &= E(D_i | D_1 = \eta_1, R_1 = \phi_1, \ldots, D_{i-1} = \eta_{i-1}, R_{i-1} = \phi_{i-1}), \\ \phi_i &= E(R_i | D_1 = \eta_1, R_1 = \phi_1, \ldots, D_{i-1} = \eta_{i-1}, R_{i-1} = \phi_{i-1}, D_i = \eta_i), \end{aligned}$$

and h is a proportion of failures prefixed in advance. Numerical results are presented for OS, OES, and EP inspection times. Moreover, the simultaneous optimization of inspection times and censoring proportions is addressed. Similar considerations can be found in Lin et al. [609] for Weibull(ϑ^β, β)-distribution.

Chapter 19
Goodness-of-Fit Tests in Progressive Type-II Censoring

In life testing, parametric distributional assumptions have been widely used. Therefore, it is essential to test whether the assumed distribution fits the given data. In the following sections, we present goodness-of-fit tests to decide whether the model assumption made is reasonable when the data are progressively Type-II censored.

19.1 Tests on Exponentiality

The first paper on goodness-of-fit tests for progressively Type-II censored data is by Balakrishnan et al. [131]. They considered both one- and two-parameter exponential distributions as null hypothesis. In particular, they addressed the problem

$$H_0 : P = \mathsf{Exp}(\vartheta), \vartheta > 0 \quad \longleftrightarrow \quad H_1 : P \neq \mathsf{Exp}(\vartheta). \tag{19.1}$$

Their test procedure is based on properties of the normalized spacings $S_r^{\mathscr{R}}$, $r = 1, \ldots, m$ (cf. (2.9)). Given an exponential distribution $\mathsf{Exp}(\vartheta)$, $\vartheta > 0$, the spacings $S_r^{\mathscr{R}}$, $r = 1, \ldots, m$, are independent $\mathsf{Exp}(\vartheta)$-distributed random variables (see Theorem 2.3.2). Then, Balakrishnan et al. [131] proposed the test statistic

$$T = \frac{\sum_{i=1}^{m-1} (m-i) S_i^{\mathscr{R}}}{(m-1) \sum_{i=1}^{m} S_i^{\mathscr{R}}} = \frac{1}{m-1} \sum_{j=1}^{m-1} \frac{\sum_{i=1}^{j} S_i^{\mathscr{R}}}{\sum_{i=1}^{m} S_i^{\mathscr{R}}}. \tag{19.2}$$

An analogue of this quantity has been suggested by Tiku [847] for complete and doubly Type-II censored samples. In order to use T as a test statistic, the null distribution has to be calculated. An expression for this distribution is based on the following result (see, e.g., Balakrishnan et al. [131] and Reiss [750, Corollary 1.6.9]).

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_19,

Lemma 19.1.1. The random variables

$$\frac{S_1^{\mathscr{R}}}{\sum_{i=1}^m S_i^{\mathscr{R}}}, \frac{S_1^{\mathscr{R}} + S_2^{\mathscr{R}}}{\sum_{i=1}^m S_i^{\mathscr{R}}}, \dots, \frac{\sum_{i=1}^{m-1} S_i^{\mathscr{R}}}{\sum_{i=1}^m S_i^{\mathscr{R}}}$$

are distributed as uniform order statistics $U_{1:m-1}, \dots, U_{m-1:m-1}$.

Hence, $(m-1)T$ is distributed as the sum of $m-1$ IID uniform random variables because $\sum_{i=1}^{m-1} U_{i:m-1} \stackrel{\mathrm{d}}{=} \sum_{i=1}^{m-1} U_i$ with IID uniform random variables $U_1, \dots, U_{m-1}$. This shows that $ET = \frac{1}{2}$ and $\mathrm{Var}\, T = \frac{1}{12(m-1)}$. Hence, H_0 is rejected if T has either too small or too large values:

$$\text{Reject } H_0 \text{ if } T < c_{\alpha/2}^T \text{ or } T > c_{1-\alpha/2}^T,$$

where c_β^T denotes the β-quantile of the distribution of T. The cumulative distribution function of a sum of independent uniform random variables can be found in Feller [361, p. 27] so that $(m-1)T$ has the cumulative distribution function (see also Bradley and Gupta [218])

$$P((m-1)T \le t) = \frac{1}{(m-1)!} \sum_{j=0}^{m-1} (-1)^j \binom{m-1}{j} [t-j]_+^{m-1}, \quad t \in \mathbb{R}.$$

Remark 19.1.2. It should be noted that the density function of the sum of k independent uniform random variables can be written as a cardinal B-spline. Using the convolution property of cardinal B-splines $N_{0,r}$ (see Schoenberg [785]), i.e.,

$$N_{0,r} = N_{0,r-1} * N_{0,1}, \quad r \ge 2,$$

where $N_{0,1}$ can be seen as the density function of a uniform random variable, it is clear that the cardinal B-spline $N_{0,k}$ equals the density function of the above sum. Therefore, we can write for the cumulative distribution function of $(m-1)T$

$$P((m-1)T \le t) = \int_0^t N_{0,m-1}(u)du, \quad t \ge 0.$$

This relation may be used to compute the percentiles of the distribution of T.

Balakrishnan and Lin [111] applied an algorithm of Huffer and Lin [466] to compute the exact distribution of T under H_0. A table of quantiles for selected values of m and level $1-\alpha$ has been presented. Marohn [638] pointed out that the distribution of T can be found in Kendall et al. [515, Example 11.9] and percentiles are available in Buckle et al. [224] for $m \le 30$.

Since it is difficult to obtain exact quantiles for large m, Balakrishnan et al. [131] and Marohn [638] suggested a normal approximation to test the hypothesis using the normalized test statistic $T^* = \sqrt{12(m-1)}(T - \frac{1}{2})$:

$$\text{Reject } H_0 \text{ if } |T^*| > z_{1-\alpha/2},$$

where z_β denotes the β-quantile of the standard normal distribution. Moreover, they presented an approximation of the power function for this asymptotic test. They conducted a simulation study for various alternatives including Weibull, Lomax, log-normal, and gamma distributions.

Applying Basu's theorem, Sanjel and Balakrishnan [768] showed that moments of T can be calculated from the identity

$$ET^k = \frac{E\Big(\sum_{i=1}^{m-1}(m-i)S_i^{\mathscr{R}}\Big)^k}{E\Big((m-1)\sum_{i=1}^{m} S_i^{\mathscr{R}}\Big)^k}.$$

Using these moments, they found that the density function of T can be expanded in terms of Laguerre orthogonal polynomials. Critical values obtained by this approximation provide quite accurate approximations to the exact percentiles given in Balakrishnan and Lin [111], as shown in Table 1 of Sanjel and Balakrishnan [768].

Balakrishnan et al. [131] also proposed two alternative procedures to test exponentiality. An adaption of a procedure presented in Spinelli and Stephens [821] uses the ordered spacings $S_{1:m}^{\mathscr{R}}, \ldots, S_{m:m}^{\mathscr{R}}$ and their mean $\overline{S} = \frac{1}{m}\sum_{i=1}^{m} S_{i:m}^{\mathscr{R}}$. The corresponding test statistic A^2 is defined by

$$A^2 = -\frac{1}{m}\sum_{i=1}^{m}(2i-1)[\log W_i + \log(1-W_i)] - m,$$

where $W_i = 1 - \exp\{-S_{i:m}^{\mathscr{R}}/\overline{S}\}$, $1 \le i \le m$. Secondly, they considered a Shapiro–Wilk test statistic comparing the squared mean $\overline{S}^2$ and the sum of squares $\sum_{i=1}^{m}(S_{i:m}^{\mathscr{R}})^2$ proposed in Shapiro and Wilk [801].

Wang [882] suggested an alternative test statistic which is also based on the spacings of the data. He used the statistic (cf. (19.2))

$$T_\star = -2\log\prod_{j=1}^{m-1}\frac{\sum_{i=1}^{j} S_i^{\mathscr{R}}}{\sum_{i=1}^{m} S_i^{\mathscr{R}}} = -2\sum_{j=1}^{m-1}\log\frac{\sum_{i=1}^{j} S_i^{\mathscr{R}}}{\sum_{i=1}^{m} S_i^{\mathscr{R}}}. \tag{19.3}$$

From Lemma 19.1.1, it follows that

$$T_\star \stackrel{\mathsf{d}}{=} -2\sum_{j=1}^{m-1}\log U_{j:m-1} \stackrel{\mathsf{d}}{=} 2\sum_{j=1}^{m-1} -\log U_j \stackrel{\mathsf{d}}{=} 2\sum_{j=1}^{m-1} Z_j,$$

where $Z_1, \ldots, Z_{m-1}$ are IID standard exponential random variables. Hence, $T_\star$ exhibits a χ^2-distribution with $2(m-1)$ degrees of freedom. Wang [882] conducted a simulation study to assess the power of the corresponding test. He used Weibull, log-normal, and gamma distributions as alternatives to the exponential distribution. The results show that the test performs very well. The results suggest that the test based on $T_\star$ has a better performance than that with test statistic T in most cases.

For small samples, Marohn [638] suggested Fisher's κ statistic

$$\kappa_m = m \cdot \max_{1 \le j \le m} (S_j^\diamond - S_{j-1}^\diamond),$$

where $S_0^\diamond = 0, S_m^\diamond = 1$, and $S_j^\diamond = \frac{\sum_{i=1}^j S_i^{\mathscr{R}}}{\sum_{i=1}^m S_i^{\mathscr{R}}}$, $j = 1, \ldots, m-1$. The null hypothesis is rejected if the maximum spacing of $S_1^\diamond, \ldots, S_m^\diamond$ is significantly large. The cumulative distribution function of κ_m is given in Marohn [638] and Brockwell and Davis [220].

Example 19.1.3. In order to illustrate the preceding tests, Nelson's progressively Type-II censored insulating fluid data 1.1.5 as presented in Viveros and Balakrishnan [875] has been considered by the abovementioned authors.

Balakrishnan et al. [131] computed the test statistic $T = 0.43251$ and the p-value of the asymptotic test as 0.53620. Therefore, the null hypothesis is not rejected by the test. This decision is consistent with the findings presented in Nelson [676] and Viveros and Balakrishnan [875]. According to Balakrishnan and Lin [111], the exact p-value of the test statistic T is 0.72799 showing that there is strong evidence that the population distribution is exponential. Notice that, due to the moderate value of $m = 8$, the normal approximation does not provide a good estimate of the p-value. Using the Laguerre orthogonal polynomial approximation, Sanjel and Balakrishnan [768] computed the approximation 0.72968 for the p-value which provides an extremely accurate approximation of the exact p-value.

The preceding results are also supported by the findings of Wang [882] using the test statistics $T_\star$. He calculated the test statistic $T_\star = 16.4308$ and the p-value 0.2878. Hence, the null hypothesis of an exponential distribution cannot be rejected, too.

Using Fisher's κ statistic, Marohn [638] computed $\kappa_8 = 1.941$ and the critical value $c_{0.05} = 4.125$ at level $\alpha = 0.05$. This result is in agreement with the preceding ones.

The procedures presented above have also been adapted to the two-parameter exponential case. In this case, the test problem reads (cf. (19.1))

$$H_0 : P = \mathsf{Exp}(\mu, \vartheta), \mu \in \mathbb{R}, \vartheta > 0 \quad \longleftrightarrow \quad H_1 : P \neq \mathsf{Exp}(\mu, \vartheta).$$

Given H_0, the first spacing $S_1^{\mathscr{R}} = \gamma_1(X_{1:m:n} - \mu)$ depends on the unknown parameter μ, whereas the remaining spacings $S_j^{\mathscr{R}} = \gamma_j(X_{j:m:n} - X_{j-1:m:n})$, $j = 2, \ldots, m$, do not depend on the location parameter and have the same property as in the one-parameter exponential case. Therefore, simple modifications of the preceding tests can be applied. Specifically, the modified test statistic results by deleting the first spacing. From (19.2) and (19.3), this leads to the statistics

$$T^* = \frac{1}{m-2}\sum_{j=2}^{m-1} \frac{\sum_{i=2}^{j} S_i^{\mathscr{R}}}{\sum_{i=2}^{m} S_i^{\mathscr{R}}}, \qquad T_\star^* = -2\sum_{j=2}^{m-1} \log \frac{\sum_{i=2}^{j} S_i^{\mathscr{R}}}{\sum_{i=2}^{m} S_i^{\mathscr{R}}}. \tag{19.4}$$

Since these quantities have the same structure as T and $T_\star$, respectively, they have similar null distributions. In particular, T^* is distributed as the sum of $m-2$ uniform random variables, whereas $T_\star^*$ has a χ^2-distribution with $2(m-2)$ degrees of freedom. Moreover, the above approximations may also be used by taking this change into account. A modified version of Fisher's κ may also be used as pointed out by Marohn [638].

Example 19.1.4. The preceding methods are illustrated by analyzing Spinelli's data B.1.5. According to Balakrishnan et al. [131], $T^* = 0.75391$ and the corresponding p-value computed by the normal approximation is given by 0.00019026. They concluded that the data provide enough evidence to reject the null hypothesis. Thus, it should not be assumed that the population distribution is a two-parameter exponential distribution. This conclusion is consistent with the results of Spinelli and Stephens [821] who used the complete sample in their analysis. The exact p-value has been computed by Balakrishnan and Lin [111] as 0.0000354 which supports the preceding result. The Laguerre orthogonal polynomial approximation by Sanjel and Balakrishnan [768] yields the p-value 0.00001.

For $T_\star^*$, Wang [882] found $T_\star^* = 15.0875$ and the corresponding p-value $P(T_\star^* < 15.0875) = 0.00084$ so that the null hypothesis H_0 of a two-parameter exponential distribution is rejected, too. Thus, all the above findings suggest that the two-parameter exponential distribution does not provide a good model for the present data.

19.2 Goodness-of-Fit Tests for Other Distributional Assumptions

19.2.1 Methods Based on Spacings and Deviation from the Uniform Distribution

Balakrishnan et al. [135] adopted the spacing-based tests for exponentiality to a general location–scale family of distributions. Given a cumulative distribution function G with density function g, they addressed the problem

$$H_0 : F = G_{\mu,\vartheta}, \mu \in \mathbb{R}, \vartheta > 0 \quad \longleftrightarrow \quad H_1 : F \neq G_{\mu,\vartheta}, \tag{19.5}$$

where F denotes the population cumulative distribution function and $G_{\mu,\vartheta}$ is defined by $G_{\mu,\vartheta} = G((\cdot - \mu)/\vartheta)$.

Analogous to the above tests, they proposed an adjusted version of the test statistic T^* as given in (19.4):

$$T_\ell^* = \frac{1}{m-2}\sum_{j=2}^{m-1}\frac{\sum_{i=2}^{j} S_i^{\mathrm{e},\mathscr{R}}}{\sum_{i=2}^{m} S_i^{\mathrm{e},\mathscr{R}}},$$

where the spacings $S_i^{\mathscr{R}}$ are replaced by the normalized spacings $S_i^{\mathrm{e},\mathscr{R}} = S_i^{\mathscr{R}}/ES_i^{\mathscr{R}}$. Notice that T_ℓ^* is location and scale invariant. The null hypothesis is rejected for either small or large values of T_ℓ^*. Since the distribution of T_ℓ^* is intractable except for the exponential distribution, Balakrishnan et al. [135] conducted a Monte Carlo simulation to examine the null distribution for normal and Gumbel families. They also presented approximations of the mean and variance of T_ℓ^* using first-order approximations of the single and product moments as given in Balakrishnan and Rao [115] (see also Sect. 7.6).

Example 19.2.1. Considering Nelson's progressively Type-II censored insulating fluid data 1.1.5 in log-time scale, Balakrishnan et al. [135] tested the data for a Gumbel population distribution. Notice that Nelson [676] and Viveros and Balakrishnan [875] assumed a Weibull distribution in the original time scale which yields a Gumbel distribution in the log-time scale. They assumed a normal approximation of the distribution of T_ℓ^* with mean 0.5 and variance 0.013597 leading to the p-value 0.73405. This provides strong evidence that the population distribution of the log data is indeed Gumbel.

This topic has been further examined by Pakyari and Balakrishnan [701] through two approaches. Assuming H_0 as in (19.5), they considered the progressively Type-II censored sample $U_{1:m:n}^*, \ldots, U_{m:m:n}^*$ where $U_{j:m:n}^* = G_{\mu,\vartheta}(X_{j:m:n})$. Notice that, provided that the null hypothesis is true, the random variables $U_{1:m:n}^*, \ldots, U_{m:m:n}^*$ are distributed as uniform progressively Type-II censored order statistics $U_{1:m:n}, \ldots, U_{m:m:n}$. The first approach is based on normalized spacings

$$S_{j;G}^{\mathscr{R}} = \gamma_j(U_{j:m:n}^* - U_{j-1:m:n}^*), \quad j = 1, \ldots, m,$$

where $U_{0:m:n}^* = 0$. An extension to k-spacings is also described. They used the statistics

$$G_{m:n} = \sum_{j=1}^{m}(S_{j;G}^{\mathscr{R}})^2, \quad Q_{m:n} = \sum_{j=1}^{m}(S_{j;G}^{\mathscr{R}})^2 + \sum_{j=1}^{m-1} S_{j;G}^{\mathscr{R}} S_{j+1;G}^{\mathscr{R}}$$

previously proposed by Greenwood [414] and Quesenberry and Miller [735] in the setting of order statistics.

Secondly, define the normalized random variables $Y_{j:m:n} = U^*_{j:m:n} - EU_{j:m:n}$, $j = 1, \ldots, m$, where an explicit expression of $EU_{j:m:n}$ is given in Theorem 7.2.3. The following tests are based on the deviation of the progressively Type-II censored order statistics $U^*_{1:m:n}, \ldots, U^*_{m:m:n}$ from the expected values of uniform progressively Type-II censored order statistics. Pakyari and Balakrishnan [701] considered the following location–scale invariant statistics adopting methods previously used for order statistics (see Brunk [222], Stephens [824], Durbin [344], and Hegazy and Green [437]):

$$\begin{aligned} C^+_{m:n} &= \max_{1\le j\le m} Y_{j:m:n}, & C^-_{m:n} &= \max_{1\le j\le m} (-Y_{j:m:n}), \\ C_{m:n} &= \max\{C^+_{m:n}, C^-_{m:n}\}, & K_{m:n} &= C^-_{m:n} + C^+_{m:n}, \\ T^{(1)}_{m:n} &= \frac{1}{m}\sum_{j=1}^m Y_{j:m:n}, & T^{(2)}_{m:n} &= \frac{1}{m}\sum_{j=1}^m |Y_{j:m:n}|. \end{aligned} \tag{19.6}$$

If the null hypothesis is true, the distances should be rather small and so large values of the distance would support the alternative. Critical values are determined by Monte Carlo simulations. Pakyari and Balakrishnan [701] tested normal and Gumbel models versus Student's t and log-gamma models, respectively.

Example 19.2.2. Pakyari and Balakrishnan [701] generated a new progressively Type-II censored sample with $m = 10$ observations from Nelson's insulating fluid data (see data B.1.6). They found that the goodness-of-fit tests based on the statistics given in (19.6) support the hypothesis of a Gumbel distribution for the log-time scale or of a Weibull distribution for the original data. This agrees with the above findings.

For King's wire strength connection data B.1.7, Pakyari and Balakrishnan [701] argued that a normal model would be appropriate for the data. This result is supported by all tests as can be seen from the computed p-values. This result is consistent with the analysis of Nelson [676].

Remark 19.2.3. Marohn [638] mentioned that Fisher's κ may be used for all failure time distributions that can be transformed into an exponential model, e.g., for generalized Pareto distributions and Gumbel distribution. For a Gumbel distribution with location parameter μ, the transformation $t \mapsto e^t$ may be used. The transformation $t \mapsto \log t$ is appropriate for Pareto data with unknown shape parameter.

An approach for testing goodness of fit for censored samples based on transformations has also been proposed by Michael and Schucany [647]. They applied transformations to Type-II censored data so that the transformed censored sample behaves under the null hypothesis like a complete sample from the uniform distribution. An appropriate transformation for progressively Type-II censored order statistics with a one-step censoring scheme is given on p. 439/440 in Michael and Schucany [647]. A detailed discussion on transformation-based approaches has recently been presented by Fischer and Kamps [370].

19.2.2 Tests Based on Empirical Distribution Function

Marohn [638] proposed a Kolmogorov–Smirnov-type goodness-of-fit tests for the exponential distribution, which can be used for large samples ($m > 30$). With $S_j^{\diamond} = \frac{\sum_{i=1}^{j} S_i^{\mathscr{R}}}{\sum_{i=1}^{m} S_i^{\mathscr{R}}}$, $j = 1, \ldots, m-1$, and the corresponding empirical distribution function $\widehat{F}_{m-1}$ defined by

$$\widehat{F}_{m-1}(t) = \frac{1}{m-1} \sum_{i=1}^{m-1} \mathbb{1}_{[0,t]}(S_i^{\diamond}), \quad t \in \mathbb{R},$$

he considered the Kolmogorov–Smirnov statistic

$$\Delta_{m-1} = \max_{t \in [0,1]} |\widehat{F}_{m-1}(t) - t|$$

to measure the deviation from the exponential distribution. H_0 is rejected if $\sqrt{m-1}\Delta_{m-1} > c_\alpha$, where c_α denotes the α-percentile of the limiting distribution of $\sqrt{m-1}\Delta_{m-1}$. The critical values $c_{0.01} = 1.63$ and $c_{0.05} = 1.36$ are given in Marohn [638].

Pakyari and Balakrishnan [700] adopted an approach by Chen and Balakrishnan [248] for testing the composite (parametric) null hypothesis

$$H_0 : F \in \mathcal{F}_\theta, \theta \in \Theta.$$

This method is based on a transformation to normality and an application of tests using goodness-of-fit tests based on empirical distribution functions. It should be noted that this approach is not only restricted to location–scale families of distributions. The proposed goodness-of-fit procedures are constructed via modified versions of the Kolmogorov–Smirnov statistic, the Cramér–von Mises statistic, and the Anderson–Darling statistic. For details on these statistics, we refer to D'Agostino and Stephens [318]. As an analogue to the empirical distribution function, Pakyari and Balakrishnan [700] proposed the nonparametric estimate

$$\widehat{F}_{m:n}(t) = \begin{cases} 0, & t < X_{1:m:n}, \\ \alpha_{j:m:n}, & X_{j:m:n} \le t < X_{j+1:m:n}, j = 1, \ldots, m-1 \\ \alpha_{m:m:n}, & X_{m:m:n} \ge t \end{cases}$$

for the population cumulative distribution function F given a progressively Type-II censored sample $X_{1:m:n}, \ldots, X_{m:m:n}$ (an explicit expression of $\alpha_{j:m:n} = EU_{j:m:n}$ is given in Theorem 7.2.3). Notice that $\widehat{F}_{m:n}$ will be degenerated except when $n = m$ holds (complete sample case). Under the null hypothesis, the cumulative distribution function F is estimated by $F(\cdot; \widehat{\theta})$, where $\widehat{\theta}$ denotes the maximum likelihood estimator of the parameter θ. Introducing the transformed data $U_{j:m:n}^* = F(X_{j:m:n}; \widehat{\theta})$,

$j = 1, \dots, m$, Pakyari and Balakrishnan [700] proposed extended versions of quantities used for Type-II right censored data (see, e.g., Lin et al. [605]). First, a Kolmogorov–Smirnov-type statistic is given by

$$D_{m:n} = \max\{D_{m:n}^+, D_{m:n}^-\}, \tag{19.7}$$

where $D_{m:n}^+ = \max_{1\le j\le m}(\alpha_{j:m:n} - U_{j:m:n}^*)$, $D_{m:n}^- = \max_{1\le j\le m}(U_{j:m:n}^* - \alpha_{j-1:m:n})$, and $EU_{0:m:n} = 0$. A modified Cramér–von Mises statistic is defined via the quadratic deviation

$$W_{m:n}^2 = n\int_0^{U_{m:m:n}^*} (\widehat{F}_{m:n}(F^{\leftarrow}(t;\widehat{\theta})) - t)^2 dt$$

which can be evaluated to

$$\begin{aligned} W_{m:n}^2 &= \frac{nU_{m:m:n}^{*3}}{3} \\ &\quad + n\sum_{j=0}^{m-1} \alpha_{j:m:n}(U_{j+1:m:n}^* - U_{j:m:n}^*)(\alpha_{j:m:n} - (U_{j+1:m:n}^* - U_{j:m:n}^*)), \end{aligned} \tag{19.8}$$

where $U_{m+1:m:n}^* = 1$. Similarly, the modified Anderson–Darling statistic reads

$$A_{m:n}^2 = n\int_0^{U_{m:m:n}^*} \frac{(\widehat{F}_{m:n}(F^{\leftarrow}(t;\widehat{\theta})) - t)^2}{t(1-t)} dt$$

which results in the sample version

$$\begin{aligned} A_{m:n}^2 &= n\sum_{j=1}^{m-1} \Big\{\alpha_{j:m:n}^2 \log\Big(\frac{U_{j+1:m:n}^*(1-U_{j:m:n}^*)}{U_{j:m:n}^*(1-U_{j+1:m:n}^*)}\Big) \\ &\quad + 2\alpha_{j:m:n}\log\Big(\frac{1-U_{j+1:m:n}^*}{1-U_{j:m:n}^*}\Big)\Big\} - n\big\{\log(1-U_{m:m:n}^*) + U_{m:m:n}^*\big\}. \end{aligned} \tag{19.9}$$

Adopting the approach of Chen and Balakrishnan [248], Pakyari and Balakrishnan [700] proposed a five-step procedure to test H_0 (Φ denotes the cumulative distribution function of a standard normal distribution):

① Compute the maximum likelihood estimate $\widehat{\theta}$ for the parameter θ and the values $U_{j:m:n}^* = F(X_{j:m:n};\widehat{\theta})$, $j = 1, \dots, m$;
② Calculate $Y_{j:m:n}^* = \Phi^{\leftarrow}(U_{j:m:n}^*)$, $j = 1, \dots, m$;
③ Considering $Y_{1:m:n}^*, \dots, Y_{m:m:n}^*$ as progressively Type-II censored order statistics from a standard normal distribution with mean μ and standard deviation σ, compute the maximum likelihood estimates $\widehat{\mu}$ and $\widehat{\sigma}$ (see Sect. 12.7.2);

④ Calculate $U^{**}_{j:m:n} = \Phi(Y^{*}_{j:m:n})$, $j = 1, \ldots, m$;

⑤ Compute $D_{m:n}$, $W^2_{m:n}$, $A^2_{m:n}$ as given in (19.7), (19.8), and (19.9) and reject H_0 at significance level $\alpha \in (0, 1)$, if the corresponding test statistic exceeds the critical value.

Critical values for the above procedures can be obtained via Monte Carlo simulation. Pakyari and Balakrishnan [700] provided a table for $\alpha = 0.1$. Moreover, they presented a power analysis for normal distributions against several Student's t distributions as well as for Gumbel distributions versus log-gamma distributions and vice versa. The method is applied to Nelson's fluid data as in Example 19.2.1 and to Spinelli and Stephens' data as in Example 19.1.4. All the tests supported the null hypothesis of a Gumbel distribution for Nelson's log-scaled data (or a Weibull distribution in the original scale). For Spinelli and Stephens' data, Pakyari and Balakrishnan [700] found that a Weibull model is supported by the three tests.

19.2.3 Tests Based on Kullback–Leibler Distance

Balakrishnan et al. [138] and Rad et al. [736] proposed goodness-of-fit tests based on Kullback–Leibler distance (9.27)

$$\mathscr{I}_{\mathscr{R}}(f \,\|\, f_0) = \int_S f^{\mathbf{X}^{\mathscr{R}}}(x) \log \frac{f^{\mathbf{X}^{\mathscr{R}}}(x)}{f_0^{\mathbf{Y}^{\mathscr{R}}}(x)} dx,$$

where $f_0 = f_0(\cdot; \theta)$ denotes the density function under the null hypothesis. Using (9.28) and the data $x_1, \ldots, x_m$, Balakrishnan et al. [138] presented the approximation

$$\mathscr{I}_{1,\ldots,m:m:n} = -n\overline{H}_{1,\ldots,m:m:n} - \sum_{j=1}^{m} [\log f_0(x_j; \theta) + R_j \log \overline{F}_0(x_j; \theta)]$$

of $\mathscr{I}_{\mathscr{R}}(f \,\|\, f_0)$, where $\overline{H}_{1,\ldots,m:m:n} = (\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n}(f) + \log c(\mathscr{R}))/n$ (see (9.24)). Analogous to Park [706], Balakrishnan et al. [138] gave the nonparametric estimate

$$H(w, m, n) = \frac{1}{n} \sum_{j=1}^{m} \log \left(\frac{X_{j+w:m:n} - X_{j-w:m:n}}{EU_{j+w:m:n} - EU_{j-w:m:n}} \right)$$

of $\overline{H}_{1,\ldots,m:m:n}$, where the *window size* w is a positive integer with $1 \le w \le m/2$ and $X_{j+w:m:n} = X_{m:m:n}$, $j + w \ge m$, $X_{j-w:m:n} = X_{1:m:n}$, $j - w \le 1$. Then, they proposed the test statistic

$$T(w,m,n) = -H(w,m,n) - \sum_{j=1}^{m}[\log f_0(x_j;\widehat{\theta}) + R_j \log \overline{F}_0(x_j;\widehat{\theta})],$$

where $\widehat{\theta}$ is an estimator of θ.

Clearly, the null distribution is present only in the second term of $T(w,m,n)$. For an exponential distribution, Balakrishnan et al. [138] used the maximum likelihood estimator $\widehat{\vartheta} = \frac{1}{m}\sum_{j=1}^{m}(R_j + 1)X_{j:m:n}$ as given in (12.4) to construct the test statistic

$$T(w,m,n) = -H(w,m,n) + \frac{m}{n}\Big(\log\widehat{\vartheta} + 1\Big).$$

A power analysis for several (IFR, DFR, and non-monotone hazard rate) alternatives has been conducted showing that first-step censoring has a higher power than the other used censoring plans for IFR alternatives. For DFR alternatives, right censoring has the best power. For alternatives with non-monotone hazard rate, the best scheme varies with the particular alternative.

Applying the test to Nelson's insulating fluid data 1.1.5, they obtained $T(3,8,19) = 0.0724$. For the estimate $\widehat{\vartheta} = 9.09$ of ϑ, they calculated the p-value as

$$P\big(T(3,8,19) > 0.0724 \,\big|\, H_0 : F = \mathsf{Exp}(9.09)\big) = 0.9921.$$

This shows strong evidence for an exponential distribution in agreement with previous findings.

The preceding approach has been applied by Rad et al. [736] to other distributions including two-parameter Pareto distributions, log-normal distributions, and Weibull distributions. They presented a power analysis of the constructed tests versus several alternatives with increasing, decreasing, and non-monotone hazard rates. Once again, these tests have been applied to Nelson's insulating fluid data 1.1.5. It turned out that neither the Pareto nor the log-normal model provides a good fit to the data. In accordance with previous findings (see Example 19.2.2), the Weibull distribution provides an excellent fit.

Chapter 20
Counting and Quantile Processes and Progressive Censoring

20.1 Counting Process Approach

Bordes [216] has addressed nonparametric inference with progressively Type-II censored lifetime data for the population cumulative distribution function F as well as for the corresponding survival function $R = \overline{F}$, the density function f, the hazard rate function λ, and the cumulative hazard function Λ. Given a sample $X_{1:m:n}, \ldots, X_{m:m:n}$ with observations $x_1 < \cdots < x_m$, the likelihood function is given by ($x_0 = 0$)

$$\mathsf{L}(\lambda|\mathbf{x}_m) = c(\mathscr{R}) \prod_{i=1}^{m} f(x_i) \overline{F}^{R_i}(x_i).$$

Therefore, the log-likelihood function can be written in terms of the hazard rate λ and the cumulative hazard function Λ as

$$\ell(\lambda|\mathbf{x}_m) = \log c(\mathscr{R}) + \sum_{i=1}^{m} \Big(\log \lambda(x_i) - (R_i + 1)\Lambda(x_i) \Big).$$

In order to maximize this function with regard to a discrete measure $\widehat{\lambda} = \sum_{i=1}^{m} \widehat{\lambda}_i \delta_{x_i}$, where δ_x denotes the Dirac measure in $x \in \mathbb{R}$, Bordes [216] obtained that the optimal weights are given by

$$\widehat{\lambda}_i = \frac{1}{\gamma_i}, \quad i = 1, \ldots, m.$$

The corresponding nonparametric estimator $\widehat{\Lambda}$ of the cumulative hazard function Λ is given by

$$\widehat{\Lambda}(t) = \widehat{\lambda}([0,t]) = \sum_{i=1}^{m} \frac{1}{\gamma_i} \mathbb{1}_{[X_{i:m:n},\infty)}(t), \quad t \geq 0.$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_20,

Applying the connection between the survival function and the cumulative hazard function, Bordes [216] presented the product limit estimator $\widehat{R}$ of R (cf. Andersen et al. [46]) as

$$\widehat{R}(t) = \prod_{\substack{i=1\\ X_{i:m:n}\le t}}^{m} \frac{\gamma_i - 1}{\gamma_i}.$$

Notice that, for $\mathscr{R} = (0^{*m})$, $\widehat{R} = 1 - \widehat{F}$, where $\widehat{F}$ denotes the empirical cumulative distribution function of the data $X_{1:n}, \ldots, X_{n:n}$ $(m = n)$.

Bordes [216] introduced the associated counting process $N = (N(t))_{t\ge 0}$ by

$$N(t) = \sum_{i=1}^{m} \mathbb{1}_{[0,t]}(X_{i:m:n}^{\mathscr{R}_m})$$

with natural filtration $\mathscr{F}^N = (\mathscr{F}_t^N)_{t\ge 0}$ generated by N and censoring numbers $(R_j)_{j\in\mathbb{N}}$, i.e., $\mathscr{F}_t^N$ is defined as the σ-field $\sigma\big(\mathbb{N}(s); s \le t\big)$, $\mathscr{R}_m = (R_1, \ldots, R_m)$. Then, he established the following result.

Theorem 20.1.1. The process $M = (M(t))_{t\ge 0}$ defined via

$$M(t) = N(t) - \int_0^t Y(s)\lambda(s)ds, \tag{20.1}$$

where $Y(s) = \sum_{i=1}^m (R_i + 1)\mathbb{1}_{[s,\infty)}(X_{i:m:n})$ is a martingale w.r.t. the filtration $\mathscr{F}^N$.

Bordes [216] applied this connection to establish an estimator of Λ by neglecting the martingale part in (20.1)

$$\widehat{\Lambda}(t) = \int_0^t \frac{dN(s)}{Y(s)}, \quad t \ge 0. \tag{20.2}$$

Remark 20.1.2. Balakrishnan and Bordes [89] presented a smoothed version of the hazard rate estimator using a kernel function κ:

$$\widehat{\lambda}^*(t) = \frac{1}{b}\int_0^\tau \kappa\Big(\frac{t-s}{b}\Big)d\widehat{\Lambda}(s) = \frac{1}{b}\sum_{i=1}^m \frac{1}{\gamma_i}\kappa\Big(\frac{t - X_{i:m:n}}{b}\Big)\mathbb{1}_{[0,\tau]}(X_{i:m:n}),$$

where b is a bandwidth parameter, $t \in [0,\tau]$, and $[0,\tau]$ is a fixed interval with $\tau \le \infty$. For the case of order statistics, the estimator $\widehat{\lambda}^*$ reduces to the estimator

$$\widehat{\lambda}^*(t) = \frac{1}{b}\sum_{i=1}^m \frac{1}{n-i+1}\kappa\Big(\frac{t - X_{i:n}}{b}\Big)\mathbb{1}_{[0,\tau]}(X_{i:n})$$

proposed by Watson and Leadbetter [892, 893].

Under some regularity conditions, it has been shown that the kernel estimator $\widehat{\lambda}^*$ converges in probability to λ. Moreover, asymptotic normality has been established. Furthermore, Balakrishnan and Bordes [89] discussed the problem of optimal bandwidth.

Assuming the following regularity conditions for the sequence $(R_j)_{j\geq 1}$ of censoring numbers and for $\tau \in \mathbb{R}$,

Regularity Condition 20.1.3.

(1) $\sup_{m\geq 1} R_m \leq K < +\infty$,
(2) $\frac{1}{m}\sum_{i=1}^{m} R_i \xrightarrow{m\to\infty} c$,
(3) $\tau \in \mathbb{R}$ satisfies $F(\tau) < 1$,

Bordes [216] studied the asymptotic behavior of the estimators $\widehat{R}$ and $\widehat{\Lambda}$. Then, the following consistency results hold.

Proposition 20.1.4. Under Regularity Condition 20.1.3, the normalized versions $N^{(m)} = N/m$ and $Y^{(m)} = Y/m$ satisfy the following asymptotic results:

$$\sup_{0\leq t\leq \tau} \left|N^{(m)}(t) - (1 - R^{c+1}(t))\right| \xrightarrow{\text{a.s.}} 0, \tag{20.3}$$

$$\sup_{0\leq t\leq \tau} \left|Y^{(m)}(t) - (c+1)R(t)\right| \xrightarrow{\text{a.s.}} 0. \tag{20.4}$$

Remark 20.1.5. Under some additional conditions, Alvarez-Andrade and Bordes [41] showed that the counting process N can be approximated by Poisson processes M_m with $EM_m(t) = t$ such that

$$\sup_{0\leq t<\infty} \left|N(t) - \frac{M_m(t)}{M_m(m)}\right| = \mathcal{O}\Big(\sqrt{\frac{\log_2 m}{m}}\Big).$$

Moreover, they established a Brownian motion approximation for concomitants of progressively Type-II censored order statistics.

Using the asymptotic expressions given in Proposition 20.1.4, Bordes [216] established the weak consistency of the estimators $\widehat{R}$ and $\widehat{\Lambda}$.

Theorem 20.1.6 (Bordes [216]). Under Regularity Conditions 20.1.3,

(i) $\sup_{0\leq t\leq \tau} \left|\widehat{\Lambda}(t) - \Lambda(t)\right| \xrightarrow{\text{P}} 0$,
(ii) $\sup_{0\leq t\leq \tau} \left|\widehat{R}(t) - R(t)\right| \xrightarrow{\text{P}} 0$,

Moreover, the following weak convergence of the standardized processes is true. Let $D[a,b]$ denote the space of Cadlag functions on an interval $[a,b]$ (see, e.g., Pollard [724, Chap. V]).

Theorem 20.1.7. Let B be the Brownian motion on $[0,\infty)$. Then, under Regularity Conditions 20.1.3,

(i) The estimator $\widehat{\Lambda}$ converges to a Brownian motion, i.e.,

$$\sqrt{m}\big(\widehat{\Lambda}(t)-\Lambda(t)\big) \xrightarrow{\text{d}} B\circ v(t) \quad \text{in } D[0,\tau],$$

where the covariance function v is defined by

$$\operatorname{Cov}\big(B\circ v(s), B\circ v(t)\big) = v(s\wedge t) = \frac{1-R^{c+1}(s\wedge t)}{(c+1)^2R^{c+1}(s\wedge t)}, \quad s,t\in[0,\tau].$$

The covariance function v may be consistently (uniformly) estimated on the interval $[0,\tau]$ by

$$\widehat{v}(t) = m\int_0^t \frac{dN(s)}{Y^2(s)}.$$

(ii) The estimator $\widehat{R}$ converges to a Brownian motion, i.e.,

$$\sqrt{m}\big(\widehat{R}(t)-R(t)\big) \xrightarrow{\text{d}} R(t)B\circ v(t) \quad \text{in } D[0,\tau],$$

with $\operatorname{Cov}\big(R(s)B\circ v(s), R(t)B\circ v(t)\big) = R(s)R(t)v(s\wedge t)$, $s,t\in[0,\tau]$.

(iii) Finally,

$$\sup_{0\le t\le\tau}\sqrt{\frac{m}{\widehat{v}(t)}}\frac{\big|\widehat{R}(t)-R(t)\big|}{\widehat{R}(t)} \xrightarrow{\text{d}} \sup_{t\in[0,1]}|B(t)|.$$

Bordes [216] used these results to establish asymptotic confidence bounds for Λ and R as well as Gill-type confidence bands (see Fleming and Harrington [376, p. 240]). He applied the results to Nelson's insulating fluid data 1.1.5 as presented by Viveros and Balakrishnan [875].

The results of Bordes [216] have been utilized by Alvarez-Andrade et al. [42] to construct homogeneity tests for several samples of progressively Type-II censored order statistics. In particular, they compared the hazard rates of K independent samples testing the null hypothesis

$$H_0:\lambda_1=\cdots=\lambda_K=\lambda_0 \text{ or, equivalently, } H_0:F_1=\cdots=F_K=F_0,$$

where λ_0 and F_0 are the given hazard rate and cumulative distribution function, respectively, under the null hypothesis.

Suppose that K samples $X_{j:m_i:n_i;i}$, $1\le j\le m_i$, of progressively Type-II censored order statistics are given with censoring numbers $R_{1,i},\ldots,R_{m_i,i}$, $1\le i\le K$ and that Regularity Conditions 20.1.8 are satisfied (see Regularity Condition 20.1.3).

Regularity Condition 20.1.8.

(1) There exists a universal bound C with

$$\text{for all } m_1, \ldots, m_K: \sup_{1 \le i \le K} \sup_{1 \le j \le m_i} R_{j,i} \le C < +\infty;$$

(2) There exist $\theta_1, \ldots, \theta_K$ such that $\sum_{i=1}^{K} \theta_i = 1$, and, for $1 \le k \le K$,

$$\frac{m_k}{\sum_{i=1}^{K} m_k} \longrightarrow \theta_k \text{ if } \sum_{i=1}^{K} m_i \longrightarrow \infty;$$

(3) For $1 \le k \le K$,

$$\frac{1}{m_k} \sum_{j=1}^{m_k} R_{j,k} \longrightarrow c_k, \text{ if } \sum_{i=1}^{K} m_i \longrightarrow \infty;$$

(4) $\tau \in \mathbb{R}$ satisfies $F_0(\tau) < 1$,

Alvarez-Andrade et al. [42] considered the Nelson–Aalen-type estimator of the cumulative hazard rate Λ_i given by (see (20.2))

$$\widehat{\Lambda}_i(t) = \int_0^t \frac{dN_i(s)}{Y_i(s)}, \quad t \ge 0.$$

For two-sample tests, i.e., $K = 2$, the test statistic is based on the discrepancy process $\mathfrak{D}_m$ defined by

$$\mathfrak{D}_m(t) = \sqrt{m} \int_0^t W_m(s)\big(d\widehat{\Lambda}_1(s) - d\widehat{\Lambda}_2(s)\big), \quad t \ge 0,$$

where W_m is an appropriate nonnegative locally bounded predictable weight process. These tests are linear in $\widehat{\Lambda}_1 - \widehat{\Lambda}_2$ and therefore referred to as linear tests. Then, Alvarez-Andrade et al. [42] showed the following result.

Theorem 20.1.9. Suppose that Regularity Condition 20.1.8 is satisfied. Assume that a deterministic function w defined on $[0, \tau]$ exists with

$$\sup_{t \in [0,\tau]} |W_m(t) - w(t)| \xrightarrow{\mathsf{P}} 0, \quad m \to \infty.$$

Then, under H_0,

$$\mathfrak{D}_m \xrightarrow{\mathsf{d}} \mathfrak{G} \circ v \quad \text{in } D[0, \tau],$$

where $\mathfrak{G}$ is a centered Gaussian martingale with covariance function v defined by

$$v(t) = \int_0^t w^2(s)\Big(\frac{1}{\theta_1 y_1(s)} + \frac{1}{\theta_2 y_2(s)}\Big)\lambda_0(s)ds$$

and $y_i(s) = (c_i + 1)(1 - F_i(s))^{c_i+1}$, $i = 1, 2$, $s \in [0, t]$.

Alvarez-Andrade et al. [42] pointed out that tests based on linear functionals may lead to a poor performance depending on the choice of τ. Therefore, as an alternative to $\mathfrak{D}_m$, they proposed a quadratic functional based on the spacings $\mathfrak{D}_m(t_k) - \mathfrak{D}_m(t_{k-1})$, $k = 1, \ldots, p$, with $0 = t_0 < t_1 < \cdots < t_p = \tau$. Since

$$V_m = \big(\mathfrak{D}_m(t_0), \mathfrak{D}_m(t_2) - \mathfrak{D}_m(t_1), \ldots, \mathfrak{D}_m(t_p) - \mathfrak{D}_m(t_{p-1})\big)'$$

is asymptotically normal with covariance matrix

$$\Sigma = \mathsf{diag}\,\big(v(t_1), v(t_2) - v(t_1), \ldots, v(t_p) - v(t_{p-1})\big),$$

Alvarez-Andrade et al. [42] showed that under Regularity Condition 20.1.8 and the assumption $v(t_k) - v(t_{k-1}) > 0$, $k = 1, \ldots, p$, the result

$$V_m' \widehat{\Sigma}^{-1} V_m \xrightarrow{\mathsf{d}} \chi^2(p), \text{ as } m \to \infty,$$

holds, where $\widehat{\Sigma}^{-1} = \mathsf{diag}(1/\widehat{v}(t_1), 1/[\widehat{v}(t_2) - \widehat{v}(t_1)], \ldots, 1/[\widehat{v}(t_p) - \widehat{v}(t_{p-1})])$ and $\widehat{v}$ is a uniformly consistent estimator of v defined by

$$\widehat{v}(t) = m\int_0^t W_m^2(s)\Big(\frac{J_1(s)}{Y_1(s)} + \frac{J_2(s)}{Y_2(s)}\Big)\frac{d(N_1(s) + N_2(s))}{Y_1(s) + Y_2(s)}$$

with $J_i(s) = \mathbb{1}_{(0,\infty)}(Y_i(s))$, $i = 1, 2$.

Remark 20.1.10. Further proposals like Kolmogorov–Smirnov-, Cramér–von Mises-, and Anderson–Darling-type statistics are presented by Alvarez-Andrade et al. [42]. Details on the corresponding limit results are also provided by these authors.

In the K-sample setting, Alvarez-Andrade et al. [42] considered a proportional hazards model, i.e., $\lambda_k = \alpha_k \lambda_0$ with some $\alpha_k > 0$, $k = 1, \ldots, K$. Defining $N = \sum_{i=1}^K N_i$ and $Y = \sum_{i=1}^K Y_i$, they considered a weighted score statistic $\widehat{\mathcal{U}}^{(w)}$ with components

$$\widehat{\mathcal{U}}_k^{(w)}(\tau) = \int_0^\tau W_k(s)dN_k(s) - \int_0^\tau \frac{W_k(s)Y_k(s)}{Y(s)}dN(s)$$

with a predictable weight function W_k which simplifies under H_0 to

$$\widehat{\mathcal{U}}_k^{(w)}(\tau) = \sum_{i=1}^{K} \int_0^\tau W_k(s)\Big(\delta_{ik} - \frac{Y_k(s)}{Y(s)}\Big) dM_i(s).$$

δ_{ik} denotes the Kronecker symbol and M_i is as defined in (20.1). Then, the following asymptotic result is true.

Theorem 20.1.11. Assume that Regularity Condition 20.1.8 is true and that W_k, $2 \le k \le K$, are bounded predictable processes such that

$$\max_{2\le k\le K} \sup_{s\in[0,\tau]} \left|W_k(s) - w_k(t)\right| \xrightarrow{\mathsf{P}} 0,\ m \to \infty,$$

for some deterministic functions w_k, $2 \le k \le K$.

Then, the process $\sqrt{m}\widehat{\mathcal{U}}^{(w)}$ converges in distribution to a centered Gaussian martingale $\mathfrak{G} \circ \Sigma$ in $\times_{i=2}^{K} D[0,\tau]$ with covariance matrix $\Sigma = (\sigma_{ij})_{2\le i,j\le K}$ on $[0,\tau]$ given by

$$\sigma_{ij} = \sum_{l=1}^{K} \theta_l \int_0^\tau w_i(s) w_j(s)\Big(\delta_{il} - \frac{y_i(s)}{y(s)}\Big)\Big(\delta_{lj} - \frac{y_j(s)}{y(s)}\Big) y_l(s)\lambda(s)ds,$$

where $y(s) = \sum_{l=1}^{K} \theta_l y_l(s)$, $s \in [0,\tau]$.

Moreover, for $t \in [0,\tau]$, the covariance matrices $\Sigma(t)$ are consistently estimated by $\widehat{\Sigma}(t) = (\widehat{\sigma}_{ij}(t))_{2\le i,j\le K}$, with

$$\widehat{\sigma}_{ij}(t) = \frac{1}{m} \sum_{i=1}^{K} \int_0^t W_i(s) W_j(s)\Big(\delta_{il} - \frac{Y_i(s)}{Y(s)}\Big)\Big(\delta_{lj} - \frac{Y_j(s)}{Y(s)}\Big) dN(s).$$

This theorem implies that, for $t \in [0,\tau]$ such that $\det \Sigma(t) \ne 0$,

$$\widehat{\kappa}(t) = (\widehat{\mathcal{U}}^{(w)}(t))'(\widehat{\Sigma}(t))^{-1}\widehat{\mathcal{U}}^{(w)}(t) \xrightarrow{\mathsf{d}} \chi^2(K-1), m \to \infty.$$

Hence, for given $t \in [0,\tau]$, a homogeneity test proceeds by checking

$$\widehat{\kappa}(t) \le \chi^2_{1-\alpha}(K-1).$$

Obviously, the performance of the proposed test depends heavily on the choice of $t \in [0,\tau]$. In order to overcome this drawback, Alvarez-Andrade et al. [42] constructed a quadratic functional that is asymptotically χ^2-distributed (see two-sample problem). Introducing $0 = t_0 < t_1 < \cdots < t_p = \tau$ and defining $\widehat{\Gamma}_i = \widehat{\Sigma}(t_i) - \widehat{\Sigma}(t_{i-1})$, $i = 1, \ldots, p$, with $\widehat{\Sigma}(0) = 0$, a consistent estimator of $\Gamma = \mathrm{diag}(\Gamma_1, \ldots, \Gamma_p)$ is given by $\widehat{\Gamma} = \mathrm{diag}(\widehat{\Gamma}_1, \ldots, \Gamma_p)$. Moreover, $\widehat{\mathcal{V}} = (\widehat{\mathcal{V}}_1', \ldots, \widehat{\mathcal{V}}_p')'$ with $\widehat{\mathcal{V}}_i = \widehat{\mathcal{U}}(t_i) - \widehat{\mathcal{U}}(t_{i-1})$ satisfies that $\sqrt{m}\,\widehat{\mathcal{V}}$ is asymptotically normal. Then,

$$\widehat{\mathcal{V}}'\widehat{\Gamma}^{-1}\widehat{\mathcal{V}} = \sum_{i=1}^{K} \widehat{\mathcal{V}}_i'\widehat{\Gamma}_i^{-1}\widehat{\mathcal{V}}_i \stackrel{\mathrm{d}}{\longrightarrow} \chi^2(p(K-1)), m \to \infty.$$

Details on the performance of the proposed estimators as well as an extensive simulation study are presented in Alvarez-Andrade et al. [42].

Remark 20.1.12. Nonparametric estimation of the survival function $R = \overline{F}$ with progressive Type-I censored data (see Sect. 13) has been discussed in Balakrishnan et al. [145] and Burke [225]. Burke [225] studied the asymptotic behavior of product limit estimators of R for two types of progressive Type-I censoring by means of martingales and empirical processes. As for the Type-II censoring scenario, limiting results are established. In particular, he showed that the presented estimators are asymptotically equivalent to those estimators applied for Type-II right censored data. It is illustrated that results for confidence bands and statistical tests can be directly applied to the progressive censoring setting.

Balakrishnan et al. [145] discussed nonparametric estimation of R with progressive Type-I interval censored data which involves only the censoring number R_i (at time T_i) and the number of observations D_i in a time interval $(T_{i-1}, T_i]$ (see Chap. 18). They proposed Nelson–Aalen- and Kaplan–Meier-type estimators when the population cumulative distribution functions are specified by a parametric family $(F_\theta)_{\theta\in\Theta}$, $\Theta \subseteq \mathbb{R}$. Moreover, asymptotical properties have been established.

Balakrishnan et al. [145] pointed out that the reliability at T_i can be estimated nonparametrically by

$$\widehat{\overline{F}}_i = \prod_{j=1}^{i} \frac{\alpha_j^-}{\alpha_{j-1}^+} = \prod_{j=1}^{i} \Big(1 - \frac{D_j}{\alpha_{j-1}^+}\Big),$$

where

$$\alpha_j^- = n - D_{\bullet j} - R_{\bullet j-1}, \quad \alpha_j^+ = n - D_{\bullet j} - R_{\bullet j} = \alpha_j^- - R_j.$$

As a result, θ can be estimated by the minimum-distance estimator

$$\widehat{\theta} = \arg\min_{\theta\in\Theta} \sum_{i=1}^{m} \Big(\overline{F}_\theta(T_i) - \widehat{\overline{F}}_i\Big).$$

Balakrishnan et al. [145] established weak consistency and asymptotic normality of $\widehat{\theta}$ under some regularity conditions.

20.1.1 Semiparametric Proportional Hazards Model

Alvarez-Andrade et al. [43] investigated a semiparametric proportional hazards model for progressively Type-II censored data assuming the hazard rate representation

$$\lambda(t;T) = \exp\left\{\beta_0' Z\right\}\lambda_0(t), \quad t \geq 0,$$

where $\beta_0 \in \mathbb{R}^p$ is an unknown regression parameter. λ_0 is an unknown hazard rate and Z is a vector of p covariates.

For $1 \leq j \leq m$, let $I_j = (i_{n-\gamma_j+1}, \ldots, i_{n-\gamma_{j+1}})$, $j = 1, \ldots, m$, denote the set of indices which assigns failures and removals to component numbers in the following way: $i_{n-\gamma_j+1}$ denotes the number of the jth failed unit, whereas $i_{n-\gamma_j+2}, \ldots, i_{n-\gamma_{j+1}}$ denote the numbers of progressively censored units in the jth censoring step (see also the construction in the proof of Theorem 10.2.1 and Fig. 10.1).

Assume that the index sets $I_1, \ldots, I_m$ are known. Defining the filtration $\mathscr{F}^N = (\mathscr{F}_t^N)_{t\geq 0}$ defined as the σ-field $\sigma\big((X_{i:m:n}, I_i), 1 \leq i \leq m; X_{i:m:n} \leq t\big)$ and the counting process

$$N = \sum_{i=1}^{m} N_i(t), \quad N_i(t) = \mathbb{1}_{[0,t]}(X_{i:m:n}), t \geq 0,$$

Alvarez-Andrade et al. [43] established martingale properties, e.g., for the processes defined by

$$M_i(t) = N_i(t) - \int_0^t \sum_{j\in I_i} \exp\{\beta_0' Z_j)\mathbb{1}_{(s,\infty)}(X_{i:m:n})\lambda_0(s)ds.$$

Z_j denotes the covariate of the jth lifetime X_j, $j = 1, \ldots, n$.

Conditional on the covariate Z, the cumulative hazard rate $\Lambda(\cdot; Z)$ is estimated by $\widehat{\Lambda}(\cdot; Z)$ defined as

$$\widehat{\Lambda}(t;Z) = \exp\left\{\widehat{\beta}_n' Z\right\}\widehat{\Lambda}_0(t), \quad t \geq 0.$$

Here, the estimator $\widehat{\beta}_n$ of β is given by $\widehat{\beta}_n = \arg\max_{\beta\in\mathbb{R}^p} C_n(\beta)$, where

$$C_n(\beta) = \sum_{i=1}^{m}\left\{\beta' Z_{n-\gamma_i+1} - \log\left(\sum_{j\in I_i^m}\exp\{\beta' Z_j\}\right)\right\}, \quad I_i^m = \bigcup_{j=i}^{m} I_j.$$

The nonparametric estimator of the cumulative hazard rate Λ_0 is given by

$$\widehat{\Lambda}_0(t) = \sum_{\substack{i=1 \\ X_{i:m:n} \le t}}^{m} \Big(\sum_{j \in I_i^m} \exp\{\beta_n' Z_j\} \Big)^{-1}.$$

The estimator of the survival function is defined by the product limit estimator $\widehat{R}$ given by

$$\widehat{R}(t; Z) = \prod_{\substack{i=1 \\ X_{i:m:n} \le t}}^{m} \Big(1 - \frac{\exp\{\beta_n' Z\}}{\sum_{j \in I_i^m} \exp\{\beta_n' Z_j\}} \Big), \quad t \ge 0.$$

Under some regularity conditions, Alvarez-Andrade et al. [43] obtained asymptotic properties:

(i) $\widehat{\beta}_n \xrightarrow{\mathsf{P}} \beta_0$,
(ii) $\sqrt{n}(\widehat{\beta}_n - \beta_0)$ converges in distribution to a centered Gaussian distribution with covariance matrix $\Sigma^{-1}(\tau)$ where $\tau > 0$ satisfies $\int_0^\tau \lambda_0(s)ds < \infty$ and

$$\Sigma(\tau) = \int_0^\tau v(s, \beta_0) s^{(0)}(s, \beta_0) \lambda_0(s) ds,$$

with appropriate functions $v(\cdot, \beta_0)$ and $s^{(0)}(\cdot, \beta_0)$,
(iii) $\Sigma(\tau)$ can be consistently estimated.

Moreover, they found that $\sqrt{n}(\widehat{\Lambda}(\cdot; Z) - \Lambda(\cdot; Z))$ converges weakly to a centered Gaussian process in $D[0, \tau]$. The corresponding variance function $v(\cdot; Z)$ has an explicit integral representation and can be consistently estimated on $[0, \tau]$. More details as well as simulation results can be found in Alvarez-Andrade et al. [43].

20.2 Quantile Process Approach

The notion of quantile processes has been introduced to progressively Type-II censored order statistics by Alvarez-Andrade and Bordes [40]. They studied the asymptotic behavior of the quantile process

$$\big(X_{\lfloor \alpha m \rfloor:m:n}\big)_{\alpha \in [0,1]}$$

with $X_{0:m:n} \equiv 0$ given the following assumptions.

Regularity Condition 20.2.1.

(C1) $\alpha \in [0, a]$ for some $a \in (0, 1)$,
(C2) The sequence of censoring numbers $(R_i)_{i\in\mathbb{N}}$ is bounded by K, i.e., $R_i \leq K$ for $i \in \mathbb{N}$,
(C3) $\overline{R} = \frac{1}{m}\sum_{i=1}^{m} R_i = c + \eta_m$, where $c \geq 0$ and $(\eta_m)_{m\in N} \subseteq \mathbb{R}$ satisfies either (i) $\eta_m = o(1)$ or (ii) $\eta_m = o(m^{-1/2})$,
(C4) Let $G = 1 - \overline{F}^{c+1}$. Then, $\varepsilon \in [0, a)$ exists with $G^{\leftarrow}([\varepsilon, a]) \subseteq (c, b) \subseteq (0, \infty)$ for some $b > c$ and the hazard rate λ of F is continuous with $\lambda(t) > 0$, $t \in (c, b)$.

For $\alpha \in [0, a]$, Alvarez-Andrade and Bordes [40] defined the process $\big(\widetilde{Y}^{(m)}(\alpha)\big)_{\alpha\in[0,a]}$ by

$$\widetilde{Y}^{(m)}(\alpha) = \sqrt{m}\sum_{j=1}^{\lfloor \alpha m\rfloor} \frac{Z_j^{(m)} - 1}{\gamma_{j,m}},$$

where $\gamma_{j,m} = \sum_{i=j}^{m}(R_i + 1)$, $1 \leq j \leq m$, $m \in \mathbb{N}$. $\big(Z_j^{(m)}\big)_{1\leq j\leq m, m\in\mathbb{N}}$ defines a triangular array of IID exponential random variables. Given (C1)–(C3)(i) from Regularity Condition 20.2.1, they showed that $\big(\widetilde{Y}^{(m)}(\alpha)\big)_{\alpha\in[0,a]}$ converges to a centered Gaussian process on $[0, a]$ with variance function v defined by

$$v(\alpha) = \frac{\alpha}{(c+1)^2(1-\alpha)},$$

i.e., $\big(\widetilde{Y}^{(m)}(\alpha)\big)_{\alpha\in[0,a]} \xrightarrow{\mathsf{d}} \mathfrak{G}_v$ on $D[0, a]$. They used this limiting result to establish the following result.

Theorem 20.2.2. Given (C1)–(C3)(i) from Regularity Condition 20.2.1, the following results hold:

(i) $\sup_{\alpha\in[0,a]} |X_{\lfloor\alpha m\rfloor:m:n} - G^{\leftarrow}(\alpha)| \xrightarrow{\mathsf{P}} 0$,
(ii) Let $X^{(m)} = \big(\sqrt{m}(X_{\lfloor\alpha m\rfloor:m:n} - G^{\leftarrow}(\alpha))\big)_{\alpha\in[0,a]}$. If additionally (C3)(ii) and (C4) hold, then

$$X^{(m)} \xrightarrow{\mathsf{d}} \mathfrak{G}_v \quad \text{on } D[\varepsilon, a],$$

where $\mathfrak{G}_v$ is a centered Gaussian process on $[\varepsilon, a]$ with variance function v defined by

$$v(\alpha) = \frac{\alpha}{(c+1)^2(1-\alpha)\lambda^2\big(G^{\leftarrow}(\alpha)\big)}, \quad \alpha \in [\varepsilon, a].$$

Alvarez-Andrade and Bordes [40] proposed the estimator

$$\widehat{v}(\alpha) = \frac{\alpha}{(\overline{R}+1)^2(1-\alpha)\widehat{\lambda}^2\big(X_{\lfloor \alpha m\rfloor:m:n}\big)}, \quad \alpha \in (0, a),$$

where $\overline{R} = \frac{1}{m}\sum_{j=1}^m R_j$ and $\widehat{\lambda}$ is an estimator of the hazard rate λ. Using the results of Bordes [216], they introduced the kernel estimator

$$\widehat{\lambda}^{(m)} = \frac{1}{b_m}\int_0^\tau \kappa\Big(\frac{t-s}{b_m}\Big)\widehat{\Lambda}^{(m)}(ds),$$

where κ is a kernel function, b_m is the bandwidth, and $\widehat{\Lambda}^{(m)}$ is a nonparametric estimator of the cumulative hazard function Λ as defined in (20.2). They showed that under assumptions (C1)–(C4) from Regularity Condition 20.2.1 and for all $\alpha \in [0, a]$, the estimator $\widehat{v}(\alpha)$ converges in probability to the true value of the variance function $v(\alpha)$.

Introducing the function $g^{(m)}(\alpha) = 1-(1-\alpha)^{\overline{R}+1}$, they showed that, given (C1)–(C3)(i) from Regularity Condition 20.2.1,

$$\widehat{q}_\alpha^{(m)} = X_{\lfloor g^{(m)}(\alpha)m\rfloor:m:n} \xrightarrow{\mathsf{P}} F^{\leftarrow}(\alpha).$$

Moreover, assuming additionally (C3)(ii) and (C4), they established asymptotic normality, i.e., for $\alpha \in (\varepsilon, a)$,

$$\sqrt{m}\big(\widehat{q}_\alpha^{(m)} - F^{\leftarrow}(\alpha)\big) \xrightarrow{\mathsf{d}} \mathsf{N}(0, v(g(\alpha)))$$

with $g(\alpha) = 1-(1-\alpha)^{c+1}$. Further results and an extensive discussion of these results are provided by Alvarez-Andrade and Bordes [40] (see also Alvarez-Andrade and Bordes [41]).

Chapter 21
Nonparametric Inferential Issues in Progressive Type-II Censoring

21.1 Precedence-Type Nonparametric Tests

Precedence-type tests are based on two samples $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ with population cumulative distribution functions F_1 and F_2, respectively. They are applied to test the null hypothesis

$$H_0 : F_1 = F_2. \tag{21.1}$$

In a life test, F_1 represents the distribution of products generated in a standard process, whereas F_2 would represent the outcome of a new process. Both samples are tested simultaneously and the subsequent failures are monitored. If additionally right censoring is employed to the Y-sample, this scenario results in two Type-II right censored samples $X_{1:n_1} \le \cdots \le X_{r_1:n_1}$ and $Y_{1:n_2}, \ldots, Y_{r:n_2}$, where the number of observations r_1 in the X-sample is random. In fact, $X_{r_1:n_1}$ is the largest order statistic in the X-sample not exceeding $Y_{r:n_2}$. The sampling situation is depicted in Fig. 21.1. In fact, the Y-sample is used to partition the X-sample into groups by counting X-failures in intervals of adjoint Y-failures. Therefore, prefixing the sample size of the Y-sample as r, the random variables M_i are defined as

$$\begin{aligned} M_1 &= \sum_{j=1}^{n_1} \mathbb{1}_{(-\infty, Y_{1:n_2}]}(X_{j:n_1}), \\ M_i &= \sum_{j=1}^{n_1} \mathbb{1}_{(Y_{i-1:n_2}, Y_{i:n_2}]}(X_{j:n_1}), \quad i = 2, \ldots, r. \end{aligned} \tag{21.2}$$

Notice that $M_{r+1} = n_2 - M_{\bullet r} = n_2 - r_1$ denotes the number of specimens from the X-sample surviving the rth failure in the Y-sample. Obviously, this number is determined by $M_1, \ldots, M_r$.

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_21,

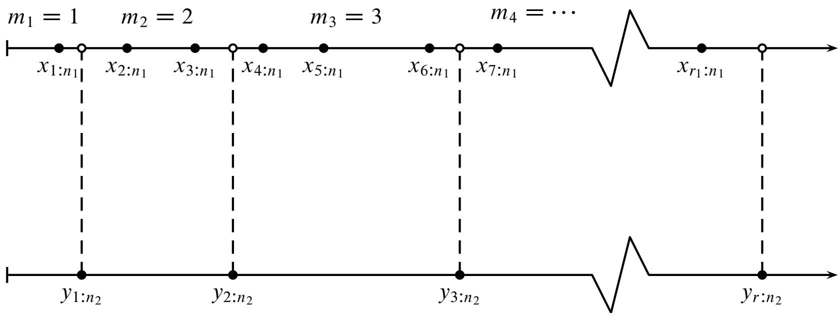

Fig. 21.1 Precedence life-test with Type-II censored data

Precedence-type tests will be particularly useful when the statistical analysis should be based only on early failures. For instance, such a scenario is present when

(i) Expensive units are put on a life test, and not all of them should be destroyed in the experiment, or when
(ii) The experimenter wants to make a quick and reliable decision at an early stage of the life test.

For more details on motivation as well as applications, we refer to Nelson [672,673], Ng and Balakrishnan [685], and the monograph by Balakrishnan and Ng [114].

21.1.1 Precedence-Type Nonparametric Tests with Progressive Censoring

Precedence-type testing with progressively Type-II censored data has been proposed in Ng and Balakrishnan [684] (see also Balakrishnan and Ng [114, Chap. 7]). While the X-sample remains Type-II right censored, the Y-sample is allowed to be progressively Type-II censored according to a censoring scheme $\mathscr{R} = (R_1, \ldots, R_r)$ with $\gamma_1(\mathscr{R}) = n_2$. Therefore, the data is given by independent samples

(i) $X_{1:n_1} \leq \cdots \leq X_{r_1:n_1}$ (lifetimes of standard product),
(ii) $Y_{1:r:n_2} \leq Y_{2:r:n_2} \leq \ldots \leq Y_{r:r:n_2}$ (lifetimes of new product).

The random variables M_i, $i = 1, \ldots, r$, are defined as in (21.2) with $Y_{i:n_2}$ replaced by $Y_{i:r:n_2}$. The sampling situation is illustrated in Fig. 21.2.

Remark 21.1.1. As an alternative to these hypothesis tests, Maturi et al. [642] discussed a method called nonparametric predictive inference (NPI) in the framework of progressive censoring. In order to compare the two groups, lower and upper probabilities for the event that a single future Y-observation exceeds

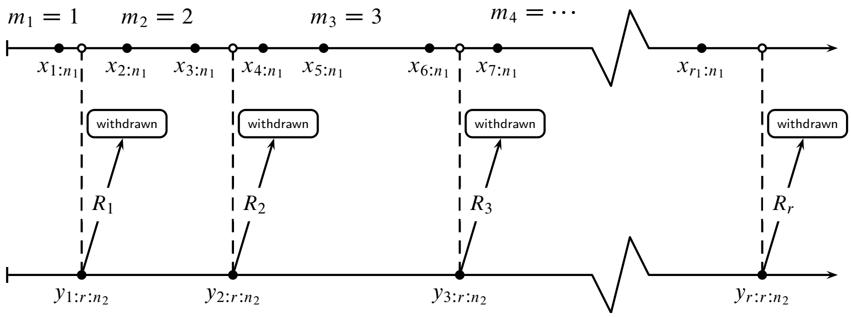

Fig. 21.2 Precedence life-test with progressive Type-II censoring

a future X-observation are calculated. Details on the methods as well as results for progressively censored data can be found in Maturi et al. [642].

Extending procedures for Type-II right censored data, Ng and Balakrishnan [684] proposed the weighted precedence test statistic P_r^* and the weighted maximal precedence test statistic $M_{(r)}^*$ as

$$P_{(r)}^* = \sum_{i=1}^{r} \gamma_i(\mathscr{R}) M_i,$$

$$M_{(r)}^* = \max_{1 \le i \le r} \{\gamma_i(\mathscr{R}) M_i\}.$$

Remark 21.1.2. As an alternative to P_r^* and $M_{(r)}^*$, one may consider Wilcoxon rank-sum precedence statistics as introduced in the case of two Type-II right censored data by Ng and Balakrishnan [683]. Such a maximal Wilcoxon rank-sum statistic for progressively censored data is defined by

$$W_{\max,r}^* = \sum_{i=1}^{r+1} M_i \Big(M_{\bullet i-1} + n_2 - \gamma_i(\mathscr{R}) + \frac{M_i + 1}{2} \Big).$$

As pointed out in Ng and Balakrishnan [685, p. 155], the identity

$$W_{\max,r}^* = \frac{n_1(n_1 + 2n_2 + 1)}{2} - P_{(r)}^*$$

holds so that $W_{\max,r}^*$ and $P_{(r)}^*$ lead to equivalent tests.

In order to establish critical values for the tests, it is necessary to derive the distribution of these test statistics under H_0 as given in (21.1). In the first step, the joint probability mass function of $\mathbf{M}_r = (M_1, \ldots, M_r)$ is established. Notice that, given H_0, these probabilities are distribution-free. In order to get compact expressions, we use the notation $m_{\bullet r} = \sum_{i=1}^{r} m_i$, $j_{\bullet r} = \sum_{i=1}^{r} j_i$ and $R_{\bullet r} = \sum_{i=1}^{r} R_i$ to denote the partial sums of the first r terms.

Theorem 21.1.3. Under the null hypothesis $H_0 : F_1 = F_2$, the probability mass function of $P^{\mathbf{M}_r}$ is given by

$$P_{F_1=F_2}(\mathbf{M}_r = \mathbf{m}_r) = C \sum_{(j_1,\ldots,j_{r-1})\in\mathfrak{A}} \left\{ \prod_{l=1}^{r-1} \binom{R_{\bullet l} - j_{\bullet l}}{j_l} \Gamma(m_l + j_{l+1} + 1) \right\} \times \frac{\Gamma\left(n_1 + n_2 - r - m_{\bullet r} - j_{\bullet r-1} + 1\right)}{\Gamma(n_1 + n_2 + 1)}, \quad (21.3)$$

where

$$C = \binom{n_1}{m_1, \ldots, m_r, n_1 - m_{\bullet r}} \prod_{j=1}^{r} \gamma_j(\mathscr{R}),$$

$$\mathfrak{A} = \Big\{ (i_1, \ldots, i_{r-1}) | 0 \le i_1 \le R_1, 0 \le i_l \le R_{\bullet l} - i_{\bullet l-1}, l = 2, \ldots, r-1 \Big\}.$$

This result has been derived by Ng and Balakrishnan [684] as a special case by assuming a Lehmann alternative specified by

$$H_1 : F_1^{\lambda} = F_2, \quad \lambda > 1. \quad (21.4)$$

Notice that the alternative implies a stochastic ordering of the distributions F_1, F_2, i.e., $F_1 > F_2$. Then, the following theorem holds.

Theorem 21.1.4. Under the Lehmann alternative $H_0 : F_1^{\lambda} = F_2$, the probability mass function of $P^{\mathbf{M}_r}$ is given by

$$P_{F_1^{\lambda}=F_2}(\mathbf{M}_r = \mathbf{m}_r) = C\lambda^r \sum_{(j_1,\ldots,j_r)\in\mathfrak{A}^*} (-1)^{j_{\bullet r}} \prod_{i=1}^{r} \binom{R_i}{j_i} \left\{ \prod_{k=1}^{r-1} \mathrm{B}\left(m_{\bullet k} + \lambda j_{\bullet k} + k\lambda, m_{k+1} + 1\right) \right\} \times \mathrm{B}\left(m_{\bullet r} + \lambda j_{\bullet r} + r\lambda, n_1 - m_{\bullet r} + 1\right),$$

where C is as given in Theorem 21.1.3, $\mathfrak{A}^* = \times_{i=1}^{r} \{1, \ldots, R_i\}$, and $\mathrm{B}(\cdot, \cdot)$ denotes the beta function.

Distribution of Test Statistics

Given n_1, n_2, r, the censoring scheme $\mathscr{R}$, and the total number of observed X-failures $P^*_{(r),\mathrm{obs}} = \sum_{i=1}^{r} \gamma_i(\mathscr{R}) m_i$, the p-values α_{obs} of the tests can be obtained from the expression given in Theorem 21.1.3. For the weighted precedence statistic $P^*_{(r)} = \sum_{i=1}^{r} \gamma_i(\mathscr{R})M_i$, this is given by

$$\alpha_{\text{obs}} = \sum_{\substack{(m_1,\ldots,m_r)\in\mathfrak{A}^\diamond \\ P^*_{(r),\text{obs}} \le P^*_{(r)} \le n_1 n_2}}^{n_1} P_{F_1=F_2}(\mathbf{M}_r = \mathbf{m}_r).$$

where $\mathfrak{A}^\diamond = \{1,\ldots,n_1\}^r$. Therefore, given a level of significance α, the null hypothesis is rejected by the test when $\alpha_{\text{obs}} > \alpha$.

Replacing $P^*_{(r),\text{obs}}$ and $P^*_{(r)}$ by the weighted maximal precedence statistic $M^*_{(r),\text{obs}} = \max_{1\le i\le r}\{\gamma_i(\mathscr{R})m_i\}$ and $M^*_{(r)} = \max_{1\le i\le r}\{\gamma_i(\mathscr{R})M_i\}$, the corresponding p-values can be computed similarly.

Alternatively, critical lower limits s may be computed for a given level of significance α. Near 5 %-critical values s are tabulated in Ng and Balakrishnan [684] and Ng and Balakrishnan [685, Tables 7.1, 7.2].

Under the Lehmann alternative (21.4), a similar expression for the power function has been established. In particular, for the weighted precedence statistic, the power function is given by

$$\alpha(\lambda, s) = \sum_{\substack{(m_1,\ldots,m_r)\in\mathfrak{A}^\diamond \\ s \le P^*_{(r)} \le n_1 n_2}}^{n_1} P_{F_1^\lambda=F_2}(\mathbf{M}_r = \mathbf{m}_r), \quad \lambda \ge 1,$$

where $P_{F_1^\lambda=F_2}(\mathbf{M}_r = \mathbf{m}_r)$ is taken from Theorem 21.1.4. An analogous expression is available for the weighted maximal precedence statistic. Power values under Lehmann alternative are computed by Ng and Balakrishnan [684] and Ng and Balakrishnan [685, Table 7.3] for various scenarios and $\lambda \in \{1,\ldots,6\}$.

Remark 21.1.5. As an alternative to Lehmann alternative, Ng and Balakrishnan [684] and Ng and Balakrishnan [685] have discussed a location-shift alternative, i.e.,

$$H_1 : F_1 = F_2(\cdot + \theta). \tag{21.5}$$

An extensive Monte Carlo simulation has been carried out for finding power values of the weighted (maximal) precedence tests with $\theta \in \{0.5, 1\}$ and various distributions including normal, exponential, gamma, log-normal, and extreme value distributions.

Finally, we reproduce an example presented in Ng and Balakrishnan [685, p. 172/3].

Example 21.1.6. Ng and Balakrishnan [684] applied the weighted (maximal) precedence tests to data generated from Nelson's insulating fluid data (see Nelson

Sample (no. 3) $x_{j:10}$	0.49	0.64	0.82	0.93	1.08	1.99	2.06	2.15	2.57	4.75
Sample (no. 6) $y^{\mathscr{R}}_{j:r:10}$	1.34	1.49	1.56	2.12	5.13					

Table 21.1 Two subsamples of times to insulating fluid breakdown data as reported in Ng and Balakrishnan [684]

[676, p. 462]). They chose subsamples no. 3 and no. 6 where progressive censoring according to the censoring plan $\mathscr{R} = (3, 0^{*3}, 2)$ was employed on sample no. 6. The resulting data are presented in Table 21.1. Applying the weighted precedence tests to the data, they obtained $P^*_{(5),\text{obs}} = 67$ and $M^*_{(5),\text{obs}} = 50$ with corresponding p-values 0.009 and 0.006. These small p-values suggest that the null hypothesis has to be rejected in favor of the alternative. Therefore, we may conclude that there is strong evidence that the lifetime distributions are different.

In order to compare the results with that in a parametric model, Ng and Balakrishnan [685] presented a corresponding analysis assuming that the X-sample is generated from an Exp(ϑ_1)-population, whereas the baseline distribution of the Y-sample is Exp(ϑ_2). Using the MLEs of ϑ_j given in (12.4), the resulting estimates are $\widehat{\vartheta}^*_{1,\text{MLE}} = 1.748$ and $\widehat{\vartheta}^*_{2,\text{MLE}} = 5.184$. From the independence of the samples and the properties of the MLEs (see Theorem 12.1.1), we conclude that the ratio $n_1\widehat{\vartheta}^*_{1,\text{MLE}}/(r\widehat{\vartheta}^*_{2,\text{MLE}})$ has an F(20,10)-distribution under $H_0 : \vartheta_1 = \vartheta_2$. Since the corresponding p-value of the two-sided test is given by 0.0185, the null hypothesis is rejected in this case as well.

Precedence-Type Test Based on Kaplan–Meier Estimator of Cumulative Distribution Function

As an alternative to the preceding test statistics, Ng and Balakrishnan [686] proposed a test statistic based on a Kaplan–Meier-type estimator of the cumulative distribution function F_2. As above, the Y-sample is given by $Y_{1:r:n_2} \le \dots \le Y_{r:r:n_2}$. Then, following Kaplan and Meier [507], a nonparametric estimator of $F_2(Y_{j:r:n_2})$ is given by

$$\widehat{F}_2(Y_{j:r:n_2}) = 1 - \prod_{i=1}^{j}\Big(1 - \frac{1}{\gamma_i(\mathscr{R})}\Big), \quad j = 1, \dots, r. \tag{21.6}$$

The Kaplan–Meier estimator of the cumulative distribution function F_1 is given by $\widehat{F}_1(X_{j:n_1}) = \frac{j}{n_1}$, $j = 1, \dots, n_1$. In order to construct the test statistic, we consider M_i, $i = 1, \dots, r$, as defined in (21.2). Suppose $0 \le Q_j \le M_j$ counts the number of X-failures $x_{l:n_1}$ in the interval $(y_{j-1:r:n_2}, y_{j:r:n_2}]$ satisfying the inequality $\widehat{F}_1(x_{l:n_1}) > \widehat{F}_2(y_{j:r:n_2})$, $j = 1, \dots r$. Then, define

$$Q_{(r)}(\mathbf{M}_r) = \max(0, M_1 - 1) + \sum_{j=2}^{r} \sum_{i=M_{\bullet j-1}+1}^{M_{\bullet j}} \mathbb{1}_{(\widehat{F}_2(Y_{j:r:n_2}),\infty)}\big(\widehat{F}_1(X_{i:n_1})\big),$$

where $M_{\bullet j} = \sum_{k=1}^{j} M_k$. Here, the information of X-observations exceeding $Y_{j:r:n_2}$ is ignored. Another statistic $Q^*_{(r)}(\mathbf{M}_r)$ can be defined by assuming that all remaining X-failures occur before the $(r+1)$th unobserved Y-failure. This latent failure $Y_{r+1:r+1:n_2}$ is defined in the sense that no progressive censoring is employed at the termination time $Y_{r:r:n_2}$ and the process is monitored up to the next failure. Therefore, it has to be assumed that at least one item is left in the experiment at the final censoring time, i.e., $R_r \geq 1$. This corresponds to progressive censoring according to the censoring scheme $(R_1, \ldots, R_{r-1}, 0, R_r - 1)$ with $R_r \geq 1$. Then, $Q^*_{(r)}(\mathbf{M}_r)$ is defined by

$$Q^*_{(r)}(\mathbf{M}_r) = Q_{(r)}(\mathbf{M}_r) + \sum_{i=M_{\bullet r}+1}^{n_1} \mathbb{1}_{(\widehat{F}_2(Y_{r+1:r+1:n_2}),\infty)}\big(\widehat{F}_1(X_{i:n_1})\big).$$

Ng and Balakrishnan [686] proposed to use the average of these two statistics

$$\begin{aligned}\overline{Q}^*_{(r)}(\mathbf{M}_r) &= \frac{1}{2}\big(Q_{(r)}(\mathbf{M}_r) + Q^*_{(r)}(\mathbf{M}_r)\big)\\ &= Q_{(r)}(\mathbf{M}_r) + \frac{1}{2}\sum_{i=M_{\bullet r}+1}^{n_1} \mathbb{1}_{(\widehat{F}_2(Y_{r+1:r+1:n_2}),\infty)}\big(\widehat{F}_1(X_{i:n_1})\big),\end{aligned}$$

Large values of $\overline{Q}^*_{(r)}(\mathbf{M}_r)$ lead to a rejection of H_0. The critical values of the corresponding test can be obtained by using the probability mass function of $P^{\mathbf{M}_r}$ as given in (21.3), i.e.,

$$P_{F_1=F_2}(\overline{Q}^*_{(r)}(\mathbf{M}_r) = q) = \sum_{\substack{(m_1,\ldots,m_r)\in\mathfrak{A}^\diamond\\ \overline{Q}^*_{(r)}(\mathbf{m}_r)=q}} P_{F_1=F_2}(\mathbf{M}_r = \mathbf{m}_r).$$

where $\mathfrak{A}^\diamond = \{1, \ldots, n_1\}^r$. Similarly, results can be obtained for the power function under the Lehmann alternative. Power comparisons are available in Ng and Balakrishnan [686]. Their results show that the statistic $\overline{Q}^*_{(r)}(\mathbf{M}_r)$ is more powerful than the weighted maximal precedence statistic. It is slightly less powerful than $\overline{P}^*_{(r)}$. Similar simulation results are available under location-shift alternative as well.

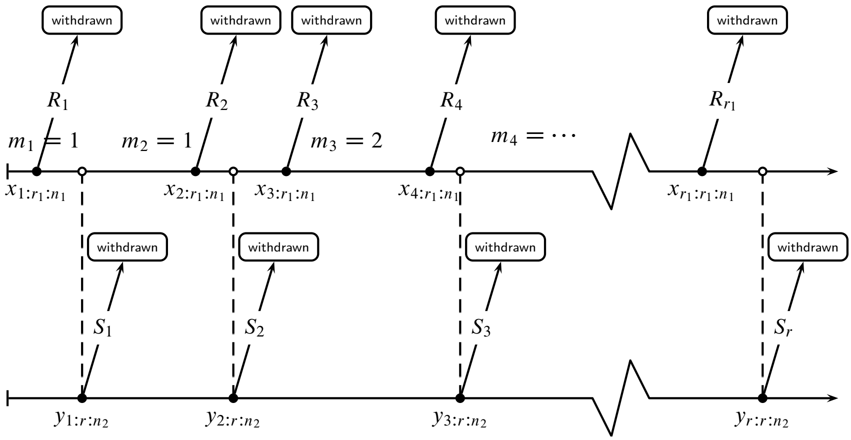

Fig. 21.3 Precedence life-test with progressive Type-II censoring employed to both samples

Two Progressively Censored Samples

In the preceding derivations, it was assumed that only the Y-sample is progressively censored. It is natural to consider also a progressively Type-II censored sample for the X-sample as has been done in Balakrishnan et al. [147]. They considered censoring schemes $\mathscr{R}$ and $\mathscr{S}$ for the X- and Y-samples, respectively, so that the data are given by independent samples

(i) $X^{\mathscr{R}}_{1:r_1:n_1} \leq \cdots \leq X^{\mathscr{R}}_{r_1:r_1:n_1}$,
(ii) $Y^{\mathscr{S}}_{1:r:n_2} \leq \ldots \leq Y^{\mathscr{S}}_{r:r:n_2}$.

This scenario is depicted in Fig. 21.3. Notice that r_1 is random as before. Again, the random variables M_i, $i = 1, \ldots, r$, are as defined in (21.2) with obvious changes.

Balakrishnan et al. [147] proposed two statistics to construct precedence test in this model. First, they constructed a Wilcoxon-type rank-sum test which is a modification of the test proposed in Ng and Balakrishnan [682]. Suppose $Y^{\mathscr{S}}_{i-1:r:n_2} < X^{\mathscr{R}}_{j:r_1:n_1} < Y^{\mathscr{S}}_{i:r:n_2}$. Then, the Wilcoxon rank-sum statistic is computed under the assumption that the R_j units progressively censored in the jth step of the X-sample would have been failed in the interval $(X^{\mathscr{R}}_{j:r_1:n_1}, Y^{\mathscr{S}}_{i:r:n_2})$. Then, the rank sum of X-failures in the pooled sample is given by

$$T_{W,r} = \frac{1}{2}\sum_{k=1}^{r} W_k(W_k + 1) + \sum_{k=2}^{r+1} W_k V_{k-1},$$

where $W_l = \sum_{i=M_{\bullet l-1}+1}^{M_{\bullet l}}(R_i+1)$, $V_l = \sum_{k=1}^{l}(W_k+S_k+1)$, $1 \le l \le r$. H_0 is rejected for small values of $T_{W,r}$. As above, the distribution of $T_{W,r}$ under the null hypothesis can be obtained as

$$P_{F_1=F_2}(T_{W,r}=w) = \sum_{\substack{(m_1,\ldots,m_r)\in\mathfrak{A}^\diamond \\ T_{W,r}=w}} P_{F_1=F_2}(\mathbf{M}_r=\mathbf{m}_r). \tag{21.7}$$

However, the probability mass function of $P^{\mathbf{M}_r}$ is more complicated. Under H_0, it can be calculated from the probability mass function of the placement statistics $\mathbf{M}_r$ discussed in Balakrishnan et al. [141]. In the following, we derive a more compact form of the probability mass function (see (21.11)).

Let $m_{\bullet 0}=0$, $m_{\bullet r+1}=0$ $Y_{0:r:n_2}=-\infty$, $Y_{r+1:r:n_2}=\infty$. Then,

$$P(\mathbf{M}_r=\mathbf{m}_r) = P\Big(\bigcap_{j=1}^{r+1}\{Y_{j-1:r:n_2} < X_{m_{\bullet j-1}+1:r_1:n_1} < X_{m_{\bullet j}:r_1:n_1} < Y_{j:r:n_2}\}\Big).$$

Conditioning on $\mathbf{Y}^{\mathscr{S}}$ and using the independence of the samples, we arrive at

$$\int P\Big(\bigcap_{j=1}^{r+1}\{y_{j-1} < X_{m_{\bullet j-1}+1:r_1:n_1} < X_{m_{\bullet j}:r_1:n_1} < y_j\}\Big) f_2^{\mathbf{Y}^{\mathscr{S}}}(\mathbf{y}_r)d\mathbf{y}_r. \tag{21.8}$$

The probability $P(\bigcap_{j=1}^{r+1}\{y_{j-1} < X_{m_{\bullet j-1}+1:r_1:n_1} < X_{m_{\bullet j}:r_1:n_1} < y_j\})$ can be written as ($y_0=-\infty$, $y_{r+1}=\infty$)

$$\prod_{j=1}^{r+1}\int_{y_{j-1}}^{y_j}\cdots\int_{y_{j-1}}^{x_2}\prod_{i=m_{\bullet j-1}+1}^{m_{\bullet j}}\Big[\gamma_i(\mathscr{R})f_1(x_i)\overline{F}_1^{R_i}(x_i)\Big]dx_{m_{\bullet j-1}+1}\ldots dx_{m_{\bullet j}}. \tag{21.9}$$

Now, with $\mathfrak{A}^* = \{(l_1,\ldots,l_r)|m_{\bullet i-1}+1 \le l_i \le m_{\bullet i}, i=1,\ldots,r\}$, $r_1=m_{\bullet r+1}$, and using the integral identity (2.31) given in Lemma 2.4.8, (21.9) can be written as

$$\begin{aligned}
&\prod_{j=1}^{r_1}\gamma_j(\mathscr{R})\sum_{(j_1,\ldots,j_r)\in\mathfrak{A}^*}\Big(\prod_{i=1}^{r} a_{j_i,m_{\bullet i}+1}^{(m_{\bullet i-1})}\overline{F}_1(y_{i-1})^{\gamma_{m_{\bullet i-1}+1}-\gamma_{j_i}}\overline{F}_1(y_i)^{\gamma_{j_i}-\gamma_{m_{\bullet i}+1}}\Big)\\
&\qquad\qquad\times a_{m_{\bullet r+1}+1,m_{\bullet r+1}+1}^{(m_{\bullet r})}\overline{F}_1(y_r)^{\gamma_{m_{\bullet r}+1}-\gamma_{m_{\bullet r+1}+1}}\\
&=\prod_{j=1}^{r_1}\gamma_j(\mathscr{R})\sum_{\substack{(j_1,\ldots,j_r)\in\mathfrak{A}^*\\ j_{r+1}=m_{\bullet r+1}+1}}\Big(\prod_{i=1}^{r+1} a_{j_i,m_{\bullet i}+1}^{(m_{\bullet i-1})}\Big)\Big(\prod_{j=1}^{r}\overline{F}_1(y_i)^{\gamma_{j_i}-\gamma_{j_i+1}}\Big).
\end{aligned}$$

For brevity, let $\gamma_i = \gamma_i(\mathscr{R})$, $i = 1, \ldots, r_1$, and $\eta_i = \gamma_i(\mathscr{S})$, $i = 1, \ldots, r$. From (21.8), we find

$$P(\mathbf{M}_r = \mathbf{m}_r) = \prod_{j=1}^{r_1} \gamma_j \prod_{j=1}^{r} \eta_j \sum_{\substack{(j_1,\ldots,j_r)\in\mathfrak{A}^* \\ j_{r+1}=m_{\bullet r+1}+1}} \Big(\prod_{i=1}^{r+1} a_{j_i, m_{\bullet i}+1}^{(m_{\bullet i}-1)}\Big) \times \int \underbrace{\prod_{i=1}^{r} \Big[f_2(y_i) \overline{F}_1(y_i)^{\gamma_{j_i}-\gamma_{j_{i+1}}} \overline{F}_2^{S_i}(y_i) \Big]}_{=h(\mathbf{y}_r)} d\mathbf{y}_r. \quad (21.10)$$

For $F_1 = F_2$, the integrand h is proportional to a density function of progressively Type-II censored order statistics with censoring scheme $\mathscr{T}$ defined by $T_i = S_i + \gamma_{j_i} - \gamma_{j_{i+1}}$, $i = 1, \ldots, r$. Hence, the probability mass function is given by

$$P_{F_1=F_2}(\mathbf{M}_r = \mathbf{m}_r) = \prod_{j=1}^{r_1} \gamma_j \sum_{\substack{(j_1,\ldots,j_r)\in\mathfrak{A}^* \\ j_{r+1}=m_{\bullet r+1}+1}} \prod_{i=1}^{r+1} a_{j_i, m_{\bullet i}+1}^{(m_{\bullet i}-1)} \prod_{l=1}^{r} \frac{\eta_l}{\eta_l + \gamma_{j_l} - \gamma_{j_{r+1}}}. \quad (21.11)$$

A similar expression has been used by Balakrishnan et al. [147] to compute critical values for the precedence statistics in the two-sample progressive censoring scenario. This applies to both the Wilcoxon-type statistic (see (21.7)) and to the Kaplan–Meier-type statistic as introduced above. Moreover, an explicit expression can be established for the power function under Lehmann alternative (see (21.4)). For $H_1 : F_1^{\lambda} = F_2$, $\lambda > 1$, $h(\mathbf{y}_r)$ reads

$$h(\mathbf{y}_r) = \lambda^r \prod_{i=1}^{r} \Big[f_1(y_i) F_1^{\lambda-1}(y_i) \overline{F}_1(y_i)^{\gamma_{j_i}-\gamma_{j_{i+1}}} (1 - F_1^{\lambda}(y_i))^{S_i} \Big]$$
$$= \lambda^r \prod_{i=1}^{r} \sum_{k=1}^{S_i} \sum_{l=1}^{\gamma_{j_i}-\gamma_{j_{i+1}}} (-1)^{k+l} \binom{S_i}{k} \binom{\gamma_{j_i} - \gamma_{j_{i+1}}}{l} \Big[f_1(y_i) F_1^{(k+1)\lambda+l-1}(y_i) \Big].$$

Clearly, F_1 can be substituted so that the power function is distribution-free. Moreover, an explicit multiple-sum expression is directly obtained from the above formula. The sum can be interpreted as a generalized mixture of dual generalized order statistics (see Burkschat et al. [234]).

Additionally, Balakrishnan et al. [147] considered the Lehmann-type alternative

$$H_1^* : \overline{F}_1^{\delta} = \overline{F}_2, \quad \delta \in (0, 1).$$

This implies $F > G$ (see Gibbons and Chakraborti [402, Chap. 6.1]). In this setting, the power function can also be derived in an explicit form. With $\overline{F}_1 = \overline{F}_2^{1/\delta}$, we get

$$h(\mathbf{y}_r) = \prod_{i=1}^{r} \Big[f_2(y_i)\overline{F}_2(y_i)^{S_i + (\gamma_{j_i} - \gamma_{j_i+1})/\delta} \Big].$$

Thus, $h(\mathbf{y}_r)$ is proportional to the joint density function of generalized order statistics based on F_2 so that the corresponding integral in (21.10) can be evaluated similar to (21.11). Therefore, the power function is given by

$$P_{\overline{F}_1^\delta = \overline{F}_2}(\mathbf{M}_r = \mathbf{m}_r) = \prod_{j=1}^{r_1} \gamma_j \sum_{\substack{(j_1,\ldots,j_r)\in\mathfrak{A}^* \\ j_{r+1} = m_{\bullet r+1}+1}} \prod_{i=1}^{r+1} a_{j_i, m_{\bullet i}+1}^{(m_{\bullet i}-1)} \prod_{l=1}^{r} \frac{\eta_l \delta}{\eta_l \delta + \gamma_{j_l} - \gamma_{j_{r+1}}}.$$

Balakrishnan et al. [147] computed critical values for the mentioned tests. Moreover, they presented extensive power comparisons for the mentioned Lehmann alternatives as well as for a location-shift alternative (see (21.5)). An application to two examples has also been illustrated by these authors.

21.1.2 Tests for Hazard Rate Ordering

Suppose $X_{1:m:n}, \ldots, X_{m:m:n}$ is a progressively Type-II censored sample from population cumulative distribution function F with censoring plan $\mathscr{R}$ and $Y_1, \ldots, Y_k$ is an IID sample from cumulative distribution function G. The samples are assumed to be independent. Sharafi et al. [802] considered testing the hypothesis

$$H_0 : \lambda_F = \lambda_G$$

versus the alternative

$$H_1 : \lambda_F \leq \lambda_G,$$

where λ_F and λ_G denote the hazard rates of the cumulative distribution functions.

First, Sharafi et al. [802] adapted an approach of Kochar [537] who considered the measure

$$\mathcal{G}(F, G) = E\big[\eta(X, Y) | X \geq Y\big], \tag{21.12}$$

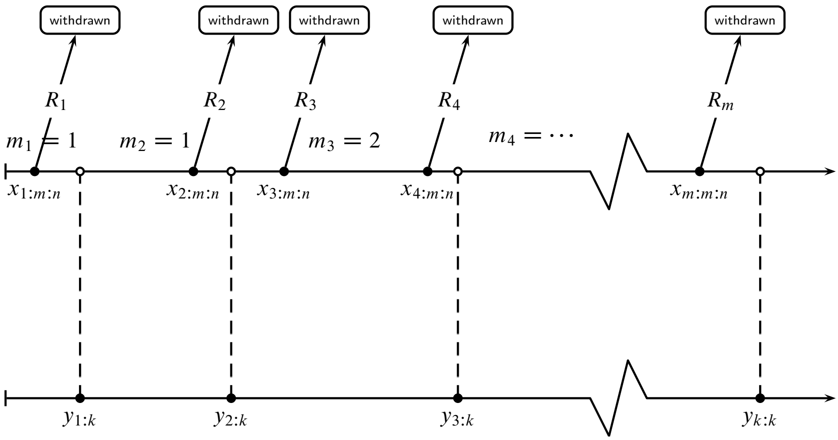

Fig. 21.4 Data setting for testing hazard rate ordering when $I \leq k$

where $\eta(x, y) = \frac{\overline{F}(x)}{\overline{G}(y)} - \frac{\overline{G}(x)}{\overline{F}(y)}$, $x \geq y \geq 0$. This measure has the representation

$$\mathcal{G}(F, G) = \frac{1}{2} + \int_0^\infty F(x)\Big\{\frac{1}{2} + \log \overline{G}(x)\Big\} dG(x).$$

Notice that $\mathcal{G}(F, F) = 0$, so that $\mathcal{G}(F, G) = 0$ under the null hypothesis, whereas it is positive under the alternative.

As for the precedence-type tests, the Y-sample is used to partition the X-sample into groups by counting X-failures in intervals of adjoint Y-failures. Thus, as in (21.2), random counters M_i are defined as

$$M_1 = \sum_{j=1}^{m} \mathbb{1}_{(-\infty, Y_{1:k}]}(X_{j:m:n}),$$

$$M_i = \sum_{j=1}^{m} \mathbb{1}_{(Y_{i-1:k}, Y_{i:k}]}(X_{j:m:n}), \quad i = 2, \ldots, k.$$

Further, $M_{k+1} = \sum_{j=1}^{m} \mathbb{1}_{[Y_{k:k}, \infty)}(X_{j:m:n}) = n - M_{\bullet k}$ denotes the number of progressively Type-II censored order statistics exceeding $Y_{k:k}$. The random variable $I = k - M_{k+1}$ denotes the numbers of Y's which do not exceed $X_{m:m:n}$. The situation is depicted in Fig. 21.4.

In the present setting, Kochar's measure (21.12) leads Sharafi et al. [802] to the test statistics $S_{\min}(I)$, $S_{\max}(I)$, and $S_A(I)$, respectively. The calculation of $S_{\min}(I)$ is based on the assumption that all units progressively censored in the ℓth censoring step fail before either the next progressively Type-II censored order statistic $X_{\ell+1:m:n}$ or the next order statistic $Y_{i:k}$. Moreover, all X-failures occur before $Y_{k:k}$. In fact, these assumptions imply that the combined increasing arrangement of the X- and Y-samples is of the form

$$\begin{aligned}X_{1:n} &\le \cdots \le X_{R_1+1:n} \le Y_{1:k} \le X_{n-\gamma_2+1:n} \le \cdots \le X_{n-\gamma_4:n}\\ &\le Y_{3:k} \le Y_{4:k} \le X_{n-\gamma_4+1:n} \le \cdots \le X_{n:n} < Y_{I+1:k}.\end{aligned}$$

Applying Kochar's construction to this arrangement, a simplified version of the test statistic presented in Sharafi et al. [802] is given by

$$S_{\min}(I) = \frac{1}{nk}\sum_{j=1}^{I}\big(n-\gamma_{m_{\bullet j}+1}\big)a_j + \frac{1}{k}\sum_{j=I+1}^{k} a_j = \frac{1}{k}\sum_{j=1}^{k} a_j - \frac{1}{nk}\sum_{j=1}^{I}\gamma_{m_{\bullet j}+1}a_j,$$

where $a_j = \frac{1}{2} + \log\big(1 - j/(k+1)\big)$, $j = 1,\ldots,k$. The calculation of the statistic $S_{\max}(I)$ is based on the idea that the lifetimes of all progressively censored units exceed $Y_{k:k}$ which means that $X_{j:m:n} = X_{j:n}$, $j = 1,\ldots,m$. Thus, the corresponding (simplified) expression of Kochar's statistic is given by

$$S_{\max}(I) = \frac{1}{nk}\sum_{j=1}^{I} m_{\bullet j}a_j + \frac{m}{nk}\sum_{j=I+1}^{k} a_j,$$

Finally, $S_A(I)$ is defined as the mean of $S_{\min}(I)$ and $S_{\max}(I)$, i.e., $S_A(I) = \frac{1}{2}\big(S_{\min}(I) + S_{\max}(I)\big)$.

As an alternative to these test statistics, Sharafi et al. [802] used results of Balakrishnan and Bordes [89] and proposed test statistics based on a nonparametric estimator of the hazard rate function as given in Remark 20.1.2 (for a complete sample, see Watson and Leadbetter [892, 893]). Using a uniform kernel function and bandwidth $1/2$, the corresponding estimates of the hazard rate are given by

$$\widehat{\lambda}_F(X_{j:m:n}) = \frac{1}{\gamma_j}, 1 \le j \le m, \quad \widehat{\lambda}_G(Y_{j:k}) = \frac{1}{k-j+1}, 1 \le j \le k.$$

Further, as detailed earlier (see p. 457), a latent failure time $X_{m+1:m+1:n}$ is introduced in the sense that no progressive censoring is employed at $X_{m:m:n}$ and the process is monitored up to the next failure. Thus, $R_m \ge 1$ has to be assumed leading to a modified censoring plan $(R_1,\ldots,R_{m-1},0,R_m-1)$. Then,

$$\widehat{\lambda}_F(X_{m+1:m+1:n}) = \frac{1}{R_m} > \frac{1}{R_m+1} = \widehat{\lambda}_F(X_{m:m:n}).$$

Let

$$N_1 = \sum_{j=1}^{k} \mathbb{1}_{(-\infty, X_{1:m:n}]}(Y_{j:k}),$$

$$N_i = \sum_{j=1}^{k} \mathbb{1}_{(X_{i-1:m:n}, X_{i:m:n}]}(Y_{j:k}), \quad i = 2, \ldots, m,$$

$$N_{m+1} = \sum_{j=1}^{k} \mathbb{1}_{(X_{m:m:n}, X_{m+1:m+1:n}]}(Y_{j:k}).$$

This approach leads to the test statistics

$$Q_1^* = \sum_{j=1}^{m} \sum_{\nu=N_{\bullet j-1}+1}^{N_{\bullet j}} \mathbb{1}_{[\widehat{\lambda}_F(X_{j:m:n}),\infty)}(\widehat{\lambda}_G(Y_{\nu:k})),$$

$$Q_2^* = Q_1^* + \sum_{\nu=N_{\bullet m}+1}^{k} \mathbb{1}_{[\widehat{\lambda}_F(X_{m+1:m+1:n}),\infty)}(\widehat{\lambda}_G(Y_{\nu:k})),$$

$$Q_A^* = Q_1^* + \frac{1}{2} \sum_{\nu=N_{\bullet m}+1}^{k} \mathbb{1}_{[\widehat{\lambda}_F(X_{m+1:m+1:n}),\infty)}(\widehat{\lambda}_G(Y_{\nu:k})).$$

Sharafi et al. [802] derived the exact null distributions of all these test statistics. They presented critical values for selected censoring schemes for $S_{\min}(I)$, $S_{\max}(I)$, $S_A(I)$, and Q_A^*. Moreover, an empirical power study, including exponential, gamma, Weibull, and Makeham distributions, has been carried out by these authors.

Part III
Applications in Survival Analysis and Reliability

Chapter 22
Acceptance Sampling Plans

Acceptance sampling plans are used to decide whether a lot of product is acceptable or must be rejected. The topic has been extensively discussed in statistical quality control where various assumptions are made w.r.t. distributions and data. For lifetime data, related references are Hosono and Kase [451], Kocherlakota and Balakrishnan [542], Balasooriya [157], Fertig and Mann [369], Schneider [784], Balasooriya et al. [163], Balakrishnan and Aggarwala [86, Chap. 11], and Fernández [366]. A detailed introduction may be found in, e.g., Montgomery [656].

Using progressively Type-II censored data $\mathbf{X}^{\mathscr{R}} = (X_{1:m:n}, \ldots, X_{m:m:n})$ for the lifetimes of units, a lot is accepted if

(i) $\phi(\mathbf{X}^{\mathscr{R}}) \geq b_l$, or
(ii) $\phi(\mathbf{X}^{\mathscr{R}}) \leq b_u$, or
(iii) $b_l \leq \phi(\mathbf{X}^{\mathscr{R}}) \leq b_u$,

where the decision rule ϕ and the bounds b_l and b_u are determined such that the probability requirements are satisfied.

Cases (i) and (ii) lead to one-sided sampling plans, whereas (iii) yields two-sided plans. For location-scale families as given in (11.1), the Lieberman–Resnikoff procedure is employed to construct the decision rule ϕ (see Lieberman and Resnikoff [592]). For some value k, the decision rule ϕ_k is constructed via estimates $\widehat{\mu}$ and $\widehat{\vartheta}$ of the location and scale parameters μ and ϑ. In particular, ϕ_k exhibits the form

$$\phi_k(\mathbf{X}^{\mathscr{R}}) = \widehat{\mu} - k\widehat{\vartheta},$$

where k denotes the acceptance constant. Therefore, for a lower specification limit b_l, a lot is accepted when $\phi_k(\mathbf{X}^{\mathscr{R}}) = \widehat{\mu} - k\widehat{\vartheta} \geq b_l$. Furthermore, a quantity of interest in our analysis is the proportion of censored items given by

$$\frac{\sum_{j=1}^{m} R_j}{n} = 1 - \frac{m}{n}.$$

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_22,

In order to construct acceptance sampling plans for progressively Type-II censored data, the censoring scheme $\mathscr{R}$ has to be incorporated in the sampling plan. Thus, the acceptance sampling plan is given by the original sample size n, the censoring scheme $\mathscr{R}$, and the acceptance constant k and is denoted by $(n, \mathscr{R}, k)$. Sometimes, the censoring proportions $q_i = R_i/n$, $i = 1, \ldots, m$, replace the censoring scheme. With $\mathscr{Q} = (q_1, \ldots, q_m)$, the acceptance sampling plan is defined as $(n, \mathscr{Q}, k)$. For an exponential baseline distribution, the distribution of the applied estimators is independent of the censoring scheme $\mathscr{R}$. Therefore, the censoring plan does not affect the acceptance sampling plan. In this framework, the triple (n, m, k) is called an acceptance sampling plan. Notice that $q = 1 - m/n$ denotes the fraction of censored units.

The acceptance sampling plan is assessed by the operating characteristic (OC) curve defined by

$$L(p) \equiv L(p; n, \mathscr{R}, k) = P(\widehat{\mu} - k\widehat{\vartheta} > \xi_p), \quad p \in (0, 1),$$

where ξ_p is the pth quantile of the baseline cumulative distribution function F.

Finally, we would like to note that acceptance sampling plans are also called reliability sampling plans when the acceptance sampling procedure is based on lifetime data. Since we are discussing only lifetime data, these terms will be used synonymously throughout this chapter.

22.1 Exponential Distribution

Reliability sampling plans based on exponential progressively Type-II censored order statistics have been considered by Balasooriya and Saw [161], Balakrishnan and Aggarwala [86], and Fernández [366]. For a progressively censored sample $\mathbf{Z}^{\mathscr{R}} = (Z_{1:m:n}, \ldots, Z_{m:m:n})$ from an $\mathrm{Exp}(\mu, \vartheta)$-distribution, the MLEs of the parameters are used to estimate the parameters and, thus, to construct the decision rule. In this scenario, the quantile ξ_p is given by $\xi_p = \mu - \vartheta \log(1 - p)$. Since the distributions of the MLEs do not depend on the censoring scheme $\mathscr{R}$ (see Theorems 12.1.1 and 12.1.4), the resulting sampling plans coincide with those for Type-II right censoring.

22.1.1 Acceptance Sampling Plans Without Consumer Risk

Let $\alpha \in (0, 1)$ be the producer risk. Then, given the OC curve L and the acceptance quality level (AQL) p_α, the sampling plan (n, m, k) has to satisfy the inequality

$$L(p_\alpha; n, m, k) \geq 1 - \alpha.$$

In order to determine the sampling plan, we have to calculate the OC curve. The following presentation is based on Chap. 11 in Balakrishnan and Aggarwala [86] and Kocherlakota and Balakrishnan [542].

First, we consider one-sided sampling plans.

One-Sided Sampling Plans

Depending on the model assumption, we have to consider three different cases.

μ Unknown, ϑ Known

The MLE of μ is given by $\widehat{\mu} = Z_{1:n}$. From Theorem 12.1.4, we get $T = \frac{n(\widehat{\mu}-\mu)}{\vartheta} \sim \mathsf{Exp}(1)$. Therefore, the OC curve is given by

$$\begin{aligned} L_l(p) &= P(Z_{1:n} - k\vartheta > \xi_p) = P\big(T > n(k - \log(1-p))\big) \\ &= \exp\Big\{-n(k - \log(1-p))\Big\}. \end{aligned} \tag{22.1}$$

For a lower acceptance sampling plan, the lot is accepted when $\widehat{\mu} + k_1\vartheta \geq b_l$. Now, given AQL p_α and producer risk α, the condition $L_l(p_\alpha) = 1 - \alpha$ can equivalently be written as

$$\exp\Big\{n(k_1 - \log(1-p_\alpha))\Big\} = 1-\alpha \iff k_1 = -\frac{\log(1-\alpha)}{n} + \log(1-p_\alpha). \tag{22.2}$$

Therefore, given $(p_\alpha, 1-\alpha)$, the acceptance sampling plan is given by (n, m, k_1) with k_1 as in (22.2). The lot is accepted iff

$$\widehat{\mu} + \Big[-\frac{\log(1-\alpha)}{n} + \log(1-p_\alpha)\Big]\vartheta \geq b_l$$

with a given lower limit b_l.

For an upper limit, the decision rule has the form $\widehat{\mu} + k_2\vartheta \leq b_u$. Hence, the OC curve has the form

$$L_l^*(p) = 1 - L_l(p) = P(Z_{1:n} - k\vartheta \leq \xi_{1-p}) = 1 - \exp\Big\{-n(k - \log(p))\Big\}.$$

Proceeding as in (22.1), the acceptance constant k_2 is obtained from the equation

$$1 - \exp\Big\{-n(k_2 - \log(p_\alpha))\Big\} = 1 - \alpha$$

with $(p_\alpha, 1-\alpha)$ as above. This yields the acceptance constant $k_2 = -\frac{\log\alpha}{n} + \log p_\alpha$.

μ Known, ϑ Unknown

The scale parameter is estimated by $\widehat{\vartheta} = \frac{1}{m}\sum_{j=1}^{m}(R_j+1)(Z_{j:m:n}-\mu)$. Using that $2m\widehat{\vartheta}/\vartheta \sim \chi^2(2m)$, the OC curve is given by

$$L_s(p) = P(\mu - k\widehat{\vartheta} > \xi_p) = 1 - F_{\chi^2(2m)}\Big(\frac{2m\log(1-p)}{k}\Big). \tag{22.3}$$

Hence, for $(p_\alpha, 1-\alpha)$ and a lower limit, we get the condition

$$k_1 = \frac{2m\log(1-p_\alpha)}{\chi^2_\alpha(2m)}.$$

Proceeding as above, we find for the upper limit the acceptance constant $k_2 = 2m\log p_\alpha/\chi^2_{1-\alpha}(2m)$.

μ Unknown, ϑ Unknown

Using the MLEs from (12.6) and the results from Theorem 12.1.4, the OC curve can be written as

$$L_{ls}(p) = P(\widehat{\mu} - k\widehat{\vartheta} > \xi_p) = P\Big(V - \frac{n}{2m}kW > -n\log(1-p)\Big), \tag{22.4}$$

where $V \sim \mathsf{Exp}(1)$ and $W \sim \chi^2(2m-2)$ are independent random variables. For $k \geq 0$, Fernández [366] showed that

$$L_{ls}(p) = \frac{(1-p)^n}{(1+nk/m)^{m-1}}, \quad p \in (0,1).$$

For $k < 0$, one has to evaluate the integral

$$\begin{aligned}
&\int_0^\infty f_{\chi^2(2m-2)}(t)P\Big(V > \frac{n}{2m}kt - n\log(1-p)\Big)dt \\
&= (1-p)^n\int_0^{2m\log(1-p)/k} f_{\chi^2(2m-2)}(t)\mathsf{e}^{-nkt/(2m)}dt \\
&\quad + \int_{2m\log(1-p)/k}^\infty f_{\chi^2(2m-2)}(t)dt \\
&= 1 - F_{\chi^2(2m-2)}\big(2m\log(1-p)/k\big) \\
&\quad + \frac{(1-p)^n}{2^{m-1}\Gamma(m-1)}\int_0^{2m\log(1-p)/k} t^{m-2}\mathsf{e}^{-(1+nk/m)t/2}dt.
\end{aligned}$$

Following Fernández [366], the OC curve can be written as

$$L_{ls}(p) = \begin{cases} \frac{(1-p)^n}{(1+nk/m)^{m-1}}, & k \geq 0 \\ 1 - F_{\chi^2(2m-2)}\big(-2n\log(1-p)\big) \\ \qquad + \frac{(1-p)^n[-n\log(1-p)]^{m-1}}{(m-1)!}, & k = -m/n\ . \\ 1 - F_{\chi^2(2m-2)}\big(2m\log(1-p)/k\big) \\ \quad + \frac{(1-p)^n}{(1+nk/m)^{m-1}} F_{\chi^2(2m-2)}\big(2(\frac{m}{k}+n)\log(1-p)\big), & \text{otherwise} \end{cases} \tag{22.5}$$

In order to determine the acceptance constant k_1 for a lower limit decision rule, one has to solve the equation $L_{ls}(p_\alpha) = 1-\alpha$ for given $(p_\alpha, 1-\alpha)$ computationally.

Remark 22.1.1. As pointed out in Balakrishnan and Aggarwala [86], one-sided acceptance sampling plans are closely related to one-sided tolerance limits. Such problems have been investigated for complete samples from an exponential population by Bain and Antle [72], Guenther et al. [416], and Engelhardt and Bain [351]. For further details, we refer to Yeh and Balakrishnan [934].

Two-Sided Sampling Plans

For two-sided sampling plans $b_l \leq \phi(\mathbf{X}^{\mathscr{R}}) \leq b_u$, we can utilize the OC curves established in the one-sided case. According to the Lieberman–Resnikoff procedure, the decision rules have the form $b_l \leq \widehat{\mu} - k\widehat{\vartheta} \leq b_u$, where $\widehat{\mu}$ and $\widehat{\vartheta}$ are suitable estimators for μ and ϑ. If a parameter is known, the corresponding estimator is chosen to be constant. Hence, we can write the OC curve as

$$L(p_1, p_2) = P(\xi_{p_1} \leq \widehat{\mu} - k\widehat{\vartheta} \leq \xi_{p_2}) = P(\widehat{\mu} - k\widehat{\vartheta} > \xi_{p_1}) - P(\widehat{\mu} - k\widehat{\vartheta} > \xi_{p_2}),$$

where $0 < p_1 < p_2 < 1$. Thus, we can use the OC curves for the one-sided plans to find the acceptance constants.

If ϑ is known, we get from (22.1)

$$\begin{aligned} L_{l2}(p_1, p_2) &= L_l(p_1) - L_l(p_2) \\ &= \mathrm{e}^{-nk}\Big\{(1-p_1)^n - (1-p_2)^n\Big\}. \end{aligned}$$

Hence, solving $L_{l2}(p_1, p_2) = 1-\alpha$ for k yields the solution

$$k_1 = -\frac{1}{n}\log\frac{1-\alpha}{(1-p_1)^n - (1-p_2)^n}.$$

For μ known, we arrive at the OC curve [see (22.3)]

$$L_{s2}(p_1, p_2) = F_{\chi^2(2m)}\Big(\frac{2m\log(1-p_2)}{k}\Big) - F_{\chi^2(2m)}\Big(\frac{2m\log(1-p_1)}{k}\Big).$$

Solving the equation $L_{s2}(p_1, p_2) = 1-\alpha$ for k can be achieved only computationally. The same comment applies to the case when both parameters are unknown. Here, one has to use the OC curve given in (22.5) instead of L_s.

22.1.2 Acceptance Sampling Plans with Consumer Risk

In practice, acceptance sampling plans are designed according to an agreement between producer and consumer. This means that the sample size and the acceptance constants have to be fixed such that the acceptance sampling plan accepts lots with

(i) a low proportion defective p_α with a high probability of at least $1-\alpha$,
(ii) a high proportion defective p_β with a low probability of at most β.

As mentioned above, p_α is called AQL and p_β is called rejectable quality level (RQL). The probabilities α and β measure the producer's and consumer's risks. This construction leads to the inequalities

$$L(p_\alpha; n, m, k) \geq 1-\alpha, \qquad L(p_\beta; n, m, k) \leq \beta, \tag{22.6}$$

which have to be satisfied by the sampling plan (n, m, k). In general, it is not possible that the inequalities become equations.

For illustration, we consider the case of a known scale parameter and a lower limit. From (22.1) and (22.2), the OC curve satisfying the first restriction is given as

$$\begin{aligned} L_l(p_\beta) &= P(Z_{1:n} - k\vartheta > \xi_{p_\beta}) = \exp\Big\{-n(k_1 - \log(1-p_\beta))\Big\} \\ &= (1-\alpha)\Big(\frac{1-p_\beta}{1-p_\alpha}\Big)^n. \end{aligned}$$

Hence, the condition $L_l(p_\beta) \leq \beta$ is equivalent to $n \geq \log\frac{\beta}{1-\alpha}/\log\frac{1-p_\beta}{1-p_\alpha}$. Therefore, the minimal sample size leading to the desired properties is given by $n^* = \big\lceil \log\frac{\beta}{1-\alpha}/\log\frac{1-p_\alpha}{1-p_\beta}\big\rceil$. So, each sampling plan (n^*, m, k_1) satisfies the inequalities (22.6). Notice that the observed sample size m can be chosen as one.

If μ is known, the inequality $L_s(p_\beta; n, m, k) \leq \beta$ is equivalent to

$$\frac{\log(1-p_\beta)}{\log(1-p_\alpha)} \leq \frac{\chi^2_{1-\beta}(2m)}{\chi^2_\alpha(2m)} \tag{22.7}$$

(see Pérez-González and Fernández [717]). Sarkar [779] has shown that the quantile $\chi^2_\alpha(r)$ is TP_2 in (r, α). This implies that the ratio $\chi^2_{1-\beta}(2m)/\chi^2_\alpha(2m)$ is increasing in

m provided $\alpha < 1-\beta$. Hence, a minimum integer m^* exists with (22.7). Therefore, an admissible acceptance sampling plan is given by (n, m^*, k_1). n can be chosen as m. Alternatively, k may be chosen in the interval $[k_\alpha, k_\beta]$, where

$$k_\alpha = \frac{2m\log(1-p_\alpha)}{\chi^2_\alpha(2m)}, \quad k_\beta = \frac{2m\log(1-p_\beta)}{\chi^2_{1-\beta}(2m)}.$$

As an alternative, Pérez-González and Fernández [717] proposed an approximate sampling plan using the Wilson–Hilferty approximation of the χ^2-cumulative distribution function. In particular, for $W \sim \chi^2(n)$,

$$3\sqrt{n}\,\frac{(W/n)^{1/3} - 1 + \frac{2}{9n}}{2}$$

is approximately standard normal (see Wilson and Hilferty [897]). Therefore, for large m, the OC curve L_s can be approximated as

$$L_s^*(p) = 1 - \Phi\Big(3\sqrt{m}\big[\{\log(1-p)/k\}^{1/3} - 1 + \tfrac{1}{9m}\big]\Big).$$

Let $u = \frac{\log(1-p_\alpha)}{\log(1-p_\beta)}$, z_α be the α-quantile of the standard normal distribution, and $\kappa = (u^{1/3}z_{1-\beta} - z_\alpha)/(1-u^{1/3})$. Then, the optimal m is determined by $\lceil(\kappa + \sqrt{\kappa^2+4})^2/36\rceil$. Moreover, every acceptance constant $k \in [k_\alpha, k_\beta]$ with

$$k_\alpha = \frac{\log(1-p_\alpha)}{(1-1/(9m) + z_\alpha/\sqrt{9m})^3}, \quad k_\beta = \frac{\log(1-p_\beta)}{(1-1/(9m) + z_{1-\beta}/\sqrt{9m})^3}$$

is possible. A comparison of exact and approximate sampling plans is provided by Pérez-González and Fernández [717].

If both parameters are assumed to be unknown, Fernández [366] has established an algorithm to compute an optimal sampling plan. The OC curve is computed from expression (22.5).

Algorithm 22.1.2 (Fernández [366]).

① Choose $(p_\alpha, 1-\alpha)$ and (p_β, β) and the censoring proportion $q = 1 - m/n$;
② Let $m_0 = n_0(1-q)$ and

$$r_\alpha = \frac{\log(1-\alpha)}{\log(1-p_\alpha)}, \quad r_\beta = \frac{\log(\beta)}{\log(1-p_\beta)}.$$

Solve the equations $L(p_\alpha; n_0, m_0, k_0) = 1-\alpha$ and $L(p_\beta; n_0, m_0, k_0) = \beta$ for (n_0, m_0).

(1) Case $r_\beta = r_\alpha$: Then, $k_0 = 0$, and $n_0 = r_\alpha = r_\beta$;
(2) Case $r_\beta < r_\alpha$: Then, $k_0 > 0$,

$$n_0 = \frac{\log(1-\alpha) - \log\beta}{\log(1-p_\alpha) - \log(1-p_\beta)}$$

and

$$k_0 = k_0(m_0, n_0) = \frac{1}{1-q}\Big[\Big(\frac{(1-p_\alpha)^{n_0}}{1-\alpha}\Big)^{1/(m_0-1)} - 1\Big];$$

(3) Case $r_\beta > r_\alpha$: Then, $k_0 < 0$. The solution (n_0, k_0) has to be computed numerically;

③ $n = \lceil n_0 \rceil$;
④ $m = \lfloor m_0 \rfloor$ if (n, m) satisfies (22.6). Otherwise, let $m = \lceil m_0 \rceil$;
⑤ As acceptance constant k any value from the interval $[k_\alpha, k_\beta]$ with $L(p_\alpha; n, m, k_\alpha) = 1 - \alpha$ and $L(p_\beta; n, m, k_\beta) = \beta$ may be chosen. A reasonable choice may be the center of the interval $k = (k_\alpha + k_\beta)/2$.

For comments on this algorithm as well as more details, we refer to Fernández [366]. Using weighted χ^2-approximations (see Patnaik [714]) and the Wilson–Hilferty approximation, Pérez-González and Fernández [717] established approximate acceptance sampling plans in the two-parameter exponential case. Moreover, they provided an algorithm to compute the optimal sampling plan and a comparison between exact and approximate sampling plans.

22.1.3 Bayesian Variable Sampling Plans with Progressive Hybrid Censoring

Lin et al. [608, 611] addressed the problem of finding exact Bayesian variable sampling plans with progressive hybrid censoring when the lifetimes are $\mathsf{Exp}(\vartheta)$-distributed. Using results of Childs et al. [260] (see Sect. 14.1.1), they obtained the Bayes risk for Type-I and Type-II progressive hybrid censored data for a decision function

$$\delta(\mathbf{X}^{\mathscr{R}}) = \begin{cases} 1, & \widehat{\vartheta} > \xi \\ 0, & \text{otherwise} \end{cases},$$

where $\widehat{\vartheta}$ is the MLE of the scale parameter ϑ in the present scenario. Given $\lambda = 1/\vartheta$ and a gamma prior $h(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}\mathrm{e}^{-\beta\lambda}$, $\lambda > 0$, with $\alpha, \beta > 0$, they considered the cost function

$$\ell(\delta(\mathbf{X}^{\mathscr{R}}), \lambda, n) = nC_s + \delta(\mathbf{X}^{\mathscr{R}})(a_0 + a_1\lambda + a_2\lambda^2) + (1 - \delta(\mathbf{X}^{\mathscr{R}}))C_r,$$

where C_s represents the cost of inspecting an item and C_r and $a_0 + a_1\lambda + a_2\lambda^2 > 0$ measure the loss of rejecting and accepting an item, respectively. Let $D = \sum_{i=1}^{m} \mathbb{1}_{(-\infty,T]}(X_{i:m:n})$ denote the total number of observed failures (see Sect. 5.1). Then, Lin et al. [608] showed that, for fixed $n, m, \mathscr{R}, T, \xi$, the Bayes risk $\mathsf{R}(n, m, \mathscr{R}, T, \xi)$ has an explicit representation in both progressive hybrid censoring settings. Using a simulated annealing algorithm developed by Corana et al. [280], they computed the minimum Bayes risk as well as the optimal sampling plan $(n^\star, m^\star, \mathscr{R}^\star, T^\star, \xi^\star)$ for given $\alpha, \beta, a_0, a_1, a_2, C_s, C_r$. Comments on the accuracy of the presented results can be found in Liang [589] (see also Lin et al. [611]).

22.2 Weibull Distribution

Balasooriya et al. [163] addressed reliability sampling plans for a $\mathsf{Weibull}(\vartheta^\beta, \beta)$-distribution employing the Lieberman–Resnikoff procedure for a lower limit. Using the AMLEs given in (12.49), they determined sample size n and acceptance constant k using asymptotic normality of $\widehat{\mu} - k\widehat{\vartheta}$ (see pp. 307) with mean $\mu - k\vartheta$ and variance $\frac{\vartheta^2}{n}[\gamma_{11}(n) - 2k\gamma_{12}(n) + k^2\gamma_{22}(n)]$. The quantities γ_{ij} are obtained from the approximate variance–covariance matrix

$$\frac{\vartheta^2}{n}\begin{pmatrix} \gamma_{11}(n) & \gamma_{12}(n) \\ \gamma_{21}(n & \gamma_{22}(n) \end{pmatrix}$$

of $(\widehat{\mu}, \widehat{\vartheta})$. Proceeding as in Schneider [784], Balasooriya et al. [163] presented the acceptance constant

$$k = \frac{y_{p_\alpha} z_{1-\beta} - y_{p_\beta} z_\alpha}{z_\alpha - z_{1-\beta}}, \tag{22.8}$$

where z_α denotes the α-quantile of the standard normal distribution and $y_\alpha = \log(-\log(1-\alpha))$, $\alpha \in (0, 1)$. The sample size is computed from the equation

$$n = \left[\frac{z_\alpha - z_{1-\beta}}{y_{p_\alpha} - y_{p_\beta}}\right]^2 [\gamma_{11}(n) - 2k\gamma_{12}(n) + k^2\gamma_{22}(n)] \tag{22.9}$$

with k taken from (22.8). Balasooriya et al. [163] also studied the accuracy of the approximations in determining acceptance sampling plans for progressively Type-II censored data. Moreover, they computed tables of reliability sampling plans for $1 - \alpha = 0.95$, $\beta \in \{0.05, 0.10\}$ and various choices of p_α, p_β matching with MIL-STD-105D, and censoring schemes.

Remark 22.2.1. Ng et al. [689] reconsidered the above problem using the MLEs of the distribution parameters. They found that the sample sizes obtained by their procedure tend to be slightly larger than those computed by Balasooriya et al.

[163]. They suspected that this may be due to the fact that the approximate MLEs tend to be less precise with increased censoring. Their claim has been supported by simulation results.

Jun et al. [489] considered plans from a Weibull distribution with known shape parameter. However, in this scenario, the data can be transformed to exponential progressively Type-II censored order statistics with unknown scale parameter and location parameter $\mu = 0$. Therefore, the results for the exponential distribution can be applied (see pp. 472). Notice that the decision of acceptance and rejection is based on the MLE of the scale parameter given in (12.12).

It should be noted that Fertig and Mann [369] and Schneider [784] addressed reliability sampling plans for Type-II right censored data from Weibull distribution. For progressively Type-II censored order statistics with binomial removals, the problem has been considered by Tse and Yang [858].

Remark 22.2.2. Sampling plans under progressive Type-I censoring (see Sect. 4) are discussed in Balasooriya and Low [159] (see also Balasooriya and Low [160]) for single Weibull failure modes. After transforming the data to the extreme value distribution, they used the corresponding estimators of location and scale and end up with a procedure similar to the one described above for progressively Type-II censored data. Moreover, they extended this approach to the case of multiple failure modes when the Weibull distributions have equal shape parameters.

22.3 Log-Normal Distribution

For a log-normal distribution, Balasooriya and Balakrishnan [158] applied the Lieberman–Resnikoff approach using approximate BLUEs $\widehat{\mu}^*$ and $\widehat{\vartheta}^*$ for the location and scale parameters (see Sect. 11.1.3). Proceeding as described in Sect. 22.2, they established a reliability sampling plan with the same ingredients as in the Weibull case. Hence, the acceptance constant is given by (22.8). The sample size is determined according to (22.9). Balasooriya and Balakrishnan [158] have tabulated sampling plans for some censoring schemes and with (p_α, p_β) matching with MIL-STD-105D.

It is clear that this approach can be adopted to other location-scale families of distributions as well.

22.4 Reliability Sampling Plans for Interval Censored Data

Using the results of Aggarwala [11] (see Sect. 18.1), Huang and Wu [462] discussed reliability sampling plans (n, k, τ) when the data are progressively Type-I interval censored with constant inspection intervals and the population distribution is exponential (see Fig. 1.8). Here, n denotes the total number of units involved in the life test, k is the number of inspections, and τ is the length of the inspection interval. They applied a normal approximation to determine the sample size and the lower specification limit. Moreover, they incorporated costs caused by the implementation of a progressively interval censored life-test sampling plan. In particular, costs for installation (C_a), sampling (nC_s), inspection (kC_I), and operation ($k\tau C_o$) are taken into account. This leads to a cost function

$$\mathsf{TC}(n, k, \tau) = C_a + nC_s + kC_I + k\tau C_o \tag{22.10}$$

which has to be minimized w.r.t. some constraints ensuring the level α and the desired power $1 - \beta$ of the reliability sampling plan. The problem is also discussed in Wu and Huang [911]. Results for two-parameter Weibull distributions in the same direction are established in Wu et al. [921]. Using a linear cost function (22.10) as for the exponential distribution, an algorithm is provided to determine the optimal design.

Lin et al. [608] discussed reliability sampling plans for progressively Type-I grouped censored data with $\mathsf{Weibull}(\vartheta^\beta, \beta)$ lifetimes. Tsai and Lin [855] presented a general method which is based on a likelihood ratio approach. Their results are illustrated by Weibull lifetimes with fixed shape parameter.

22.5 Capability Indices

Capability indices (also known as lifetime performance indices) are frequently used to measure the capability of manufacturing processes (see, e.g., Montgomery [656]). In particular, one-sided versions are widely used in industry and have received attention in the statistical literature (see Kane [505], Vännman [868], and Albing [32]). One-sided capability indices have been introduced by Kane [505, pp. 45] using the upper and lower specification limits USL and LSL as

$$C_{\mathsf{PU}} = \frac{\mathsf{USL} - \mu}{3\sigma}, \quad C_{\mathsf{PL}} = \frac{\mu - \mathsf{LSL}}{3\sigma}$$

where μ and σ denote the mean and the standard deviation of the population distribution (see also Montgomery [656, pp. 335], and Ryan [762, pp. 231]).

22.5.1 Exponential Progressively Type-II Censored Order Statistics

Generalizing the work of Tong et al. [852] and Lee et al. [580], process capability analysis has been addressed for progressively Type-II censored samples from an exponential distribution $\mathsf{Exp}(1/\vartheta)$ by Lee et al. [578], where C_{PL} is defined as

$C_{\mathsf{PL}} = \frac{\mu - \mathsf{LSL}}{\sigma}$. They discussed the problem in terms of lifetime data and interpreted C_{PL} as a lifetime performance index. In the following, we present the corresponding results for $\mathsf{Exp}(\vartheta)$-distribution. In this setting, C_{PL} is

$$C_{\mathsf{PL}} = 1 - \frac{\mathsf{LSL}}{\vartheta}.$$

A product is called a conforming product if its lifetime exceeds the lower specification limit LSL. The probability that the lifetime X exceeds LSL is defined as the conforming rate. For an exponential lifetime X, it is given by

$$p_{\mathsf{PL}} = P(X \geq \mathsf{LSL}) = \mathrm{e}^{C_{\mathsf{PL}} - 1}.$$

In practice, this connection is used as follows: If a conforming rate exceeding a given proportion p is desirable, then the condition $p_{\mathsf{PL}} \geq p$ can be written in terms of the capability index, i.e., $C_{\mathsf{PL}} \geq 1 + \log p$.

Lee et al. [578] discussed maximum likelihood estimation of C_{PL} based on exponential progressively Type-II censored order statistics $Z_{1:m:n}, \ldots, Z_{m:m:n}$. From (12.4), it follows directly that

$$\widehat{C}_{\mathsf{PL}} = 1 - \frac{\mathsf{LSL}}{\widehat{\vartheta}^*_{\mathsf{MLE}}} = 1 - m \cdot \mathsf{LSL}\Big(\sum_{j=1}^{m}(R_j + 1)Z_{j:m:n}\Big)^{-1}.$$

Since the MLE $\widehat{\vartheta}^*_{\mathsf{MLE}}$ has a $\Gamma(\vartheta/m, m)$-distribution, the reciprocal $\big(\widehat{\vartheta}^*_{\mathsf{MLE}}\big)^{-1}$ is inverted gamma distributed with mean $\frac{m}{(m-1)\vartheta}$. Thus, $\widehat{C}_{\mathsf{PL}}$ is biased. However, the estimator

$$\widehat{C}^*_{\mathsf{PL}} = 1 - \frac{(m-1)\mathsf{LSL}}{m\widehat{\vartheta}^*_{\mathsf{MLE}}}$$

is unbiased and, thus, by Theorem 12.1.1, $\widehat{C}^*_{\mathsf{PL}}$ is the UMVUE of C_{PL}. Moreover, statistical tests with level $\alpha \in (0, 1)$ for the test problem (c fixed)

$$H_0 : C_{\mathsf{PL}} \leq c \quad \text{versus} \quad H_1 : C_{\mathsf{PL}} > c \tag{22.11}$$

can directly be constructed from these results since $2m\widehat{\vartheta}^*_{\mathsf{MLE}}/\vartheta \sim \chi^2(2m)$. Hence, defining c_α as critical value and using $\widehat{C}_{\mathsf{PL}}$ as test statistic, the power function is given by

$$p(C_{\mathsf{PL}}) = P_{C_{\mathsf{PL}}}(\widehat{C}_{\mathsf{PL}} > c_\alpha) = 1 - F_{\chi^2(2m)}\Big(\frac{2m(1 - C_{\mathsf{PL}})}{1 - c_\alpha}\Big).$$

Then, under H_0, the inequality $p(C_{\mathsf{PL}}) \leq p(c)$ holds. Solving the equation $p(c) = \alpha$ for c_α leads to the critical value

$$c_\alpha = 1 - \frac{2m(1 - c)}{\chi^2_{1-\alpha}(2m)}.$$

Similarly, a one-sided confidence interval for C_{PL} can be established as

$$\mathcal{K} = \Big[1 - \frac{(1-\widehat{C}_{\mathsf{PL}})\chi^2_{1-\alpha}(2m)}{2m}, \infty\Big).$$

Example 22.5.1. Lee et al. [578] applied their results to Nelson's progressively Type-II censored insulating fluid data 1.1.5 (see also Example 12.1.6) with the lower specification limit LSL $= 1.04$ and the level of significance 5 %. Assuming that the conforming rate exceeds 80 %, the lifetime performance index C_L must exceed $c = 0.777$. Hence, the test problem reads $H_0 : C_{\mathsf{PL}} \le 0.777$ versus $H_1 : C_{\mathsf{PL}} > 0.777$. Then, the critical value $c_{0.05}$ is given by 0.876. On the other hand, from Example 12.1.6, $\widehat{C}_{\mathsf{PL}} = 1 - \frac{1.04}{9.086} = 0.886$ so that the null hypothesis is rejected. Therefore, the lifetime performance index satisfies the required level.

The results can be extended to progressively Type-II censored lifetime data from two-parameter exponential distribution as has been done in Lee et al. [581] for Type-II right censored data. Now, C_{PL} can be written as

$$C_{\mathsf{PL}} = 1 - \frac{\mathsf{LSL} - \mu}{\vartheta}.$$

Then, using the UMVUEs of the parameters given in (12.7), we define the estimator

$$\widehat{C}_{\mathsf{PL}} = 1 - \frac{1}{n} + \frac{m-2}{m-1}\frac{Z_{1:m:n} - \mathsf{LSL}}{\widehat{\vartheta}_{\mathsf{UMVUE}}}.$$

Furthermore, we have $EZ_{1:m:n} = \mu + \vartheta/n$ and $E(1/\widehat{\vartheta}_{\mathsf{UMVUE}}) = (m-1)/[(m-2)\vartheta]$. By the independence of $Z_{1:m:n}$ and $\widehat{\vartheta}_{\mathsf{UMVUE}}$, we get $E\widehat{C}_{\mathsf{PL}} = C_{\mathsf{PL}}$. Then, Theorem 12.1.4 shows that $\widehat{C}_{\mathsf{PL}}$ is the UMVUE of C_{PL}.

Clearly, an α-level test for the test problem (22.11) can also be constructed in this setting using the UMVUE $\widehat{C}_{\mathsf{PL}}$. Defining $Z^* = n(Z_{1:m:n} - \mu)/\vartheta \sim \mathsf{Exp}(1)$ and $W^* = (m-1)\widehat{\vartheta}_{\mathsf{UMVUE}}/\vartheta \sim \Gamma(1, m-1)$, the ratio $\frac{Z_{1:m:n} - \mathsf{LSL}}{\widehat{\vartheta}_{\mathsf{UMVUE}}}$ can be written as

$$\frac{Z_{1:m:n} - \mathsf{LSL}}{\widehat{\vartheta}_{\mathsf{UMVUE}}} = \frac{m-1}{n}\cdot\frac{Z^* - n(1-C_{\mathsf{PL}})}{W^*} = \frac{m-1}{n}T_{C_{\mathsf{PL}}}, \quad \text{say}.$$

It can be shown that $T_{C_{\mathsf{PL}}}$ has the density function

$$f^{T_{C_{\mathsf{PL}}}}(t) = \mathsf{e}^{-n(1-C_{\mathsf{PL}})}\frac{m-1}{(1+t)^m}\Big[\mathbb{1}_{[0,\infty)}(t) + \mathbb{1}_{(-\infty,0)}(t)\mathsf{IG}\Big(-\frac{n(1-C_{\mathsf{PL}})(t+1)}{t}; m-2\Big)\Big].$$

Distribution	Data transformation	Reference
Rayleigh	x^2	Lee et al. [579]
Weibull(ϑ, β)	x^β (β known)	Ahmadi et al. [18] (Progressive first failure censoring)
Burr XII	$\log(1+x^c)$, c known	Lee et al. [577]
Pareto(ϑ)	$\log(x)$	Hong et al. [447, 448] (Type-II right censored data)

Table 22.1 Lifetime performance index for non-exponential distributions which can be transformed to the exponential case

Remark 22.5.2. The distribution of $T_{C_{\mathsf{PL}}}$ can also be determined via the survival function given by

$$P(T_{C_{\mathsf{PL}}} > t) = P\big(Z^* - tW^* > n(1 - C_{\mathsf{PL}})\big).$$

This function equals the OC curve given in (22.4). Therefore, the survival function exhibits the expression given in (22.5).

Noticing that $T_{C_{\mathsf{PL}}}$ is stochastically increasing in C_{PL}, we get

$$T_{C_{\mathsf{PL}}} \leq_{\mathsf{st}} T_c \quad \text{for } C_{\mathsf{PL}} \leq c.$$

Therefore, the critical value for the test in (22.11) can be obtained from the equation

$$P(T_c \geq c_\alpha) = \alpha.$$

22.5.2 Other Distributions

Process capability analysis for non-exponential distributions with progressively censored data has been considered in the literature, too. However, most of the cases result by a transformation of the data. For the sake of completeness, we present the addressed distributions, the transformation of the data, and, if available, the corresponding reference in Table 22.1. The results for the Pareto distribution have been established for Type-II right censored data but can easily be rewritten for progressively Type-II censored samples. For progressive first failure censoring, we refer to the comments given in Sect. 25.4.1. Further details are provided by Laumen and Cramer [568]. Weibull(ϑ, β)-distributions with both parameters unknown are discussed in Hong et al. [449]. Two-parameter gamma distributions have recently been discussed in Laumen and Cramer [568].

Chapter 23
Accelerated Life Testing

Products that are tested in industrial experiments are often extremely reliable leading to large mean times to failure under normal operating conditions. Therefore, adequate information about the lifetime distributions and the associated parameters may be quite difficult to obtain using conventional life-testing experiments. For this reason, the units under test are subjected to higher operational demands than under normal operating conditions. Such methods are widely used and are known as accelerated life testing (ALT). Several approaches are adapted in accelerated life testing, e.g., constant high stress level, progressive stress, or stepwise increasing stress levels. Such a procedure will enable the reliability experimenter to assess the effects of stress factors such as load, pressure, temperature, and voltage on the lifetimes of experimental units. This kind of accelerated life test usually will reduce the time to failure of specimens, thus resulting in more failures than under normal operating conditions. Data collected from an accelerated life test then needs to be extrapolated to estimate the parameters of the lifetime distribution under normal operating conditions. In the literature, several models have been proposed to connect the stress levels in the ALT environment to the parameters of the original distribution. A popular model that will be assumed in the following is the cumulative exposure model introduced by Sedyakin [788] and discussed further by Bagdonavičius [67] and Nelson [675, 677]. Detailed accounts to accelerated life testing are provided by Nelson and Meeker [678], Meeker and Hahn [646], Nelson [677], Meeker and Escobar [645], and Bagdonavičius and Nikulin [68]. Results for the exponential distribution are summarized in Basu [182].

23.1 Step-Stress Models

A special type of accelerated life testing is step-stress testing. In such a testing environment, the experimenter can choose different conditions at various intermediate stages of the life test as follows. n identical units are placed on a life test at an

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_23,

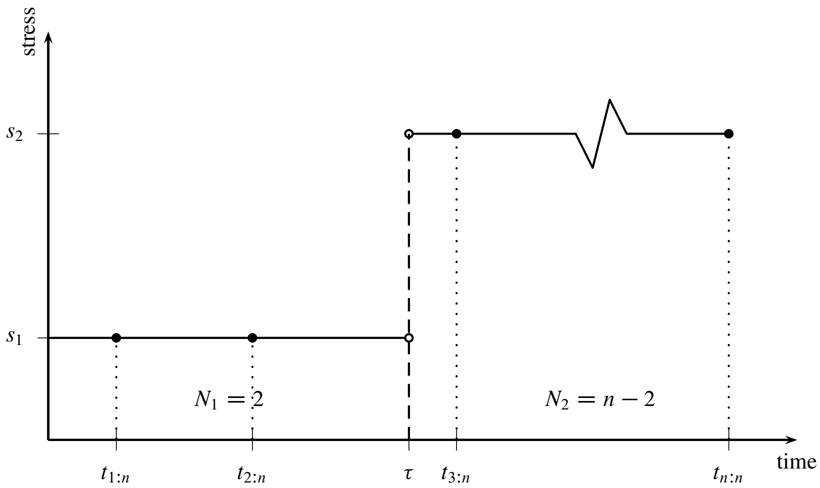

Fig. 23.1 Simple step-stress model with change time τ

initial stress level of s_1. Then, at prefixed times $\tau_1 < \cdots < \tau_k$, the stress levels are changed to $s_2, \ldots, s_{k+1}$, respectively. The special case of two stress levels s_1 and s_2, wherein the stress change occurs at a pre-fixed time τ, is called simple step-stress experiment. This situation is depicted in Fig. 23.1 for observed failure times $t_{1:n} \leq \cdots \leq t_{n:n}$. Further details can be found in Nelson [677, Chap. 10].

The simple step-stress model has been discussed extensively in the literature. Several models have been proposed to model the change of stress imposed on the units under test. Sedyakin [788], Bagdonavičius [67], and Nelson [675] studied the cumulative exposure model and associated inference. Miller and Nelson [649] and Bai et al. [71] discussed the determination of optimal time τ at which to change the stress level from s_1 to s_2. Bhattacharyya and Soejoeti [200] proposed the tampered failure rate model introduced by DeGroot and Goel [329] which assumes that the effect of change of stress is to multiply the initial failure rate function by a factor subsequent to the stress change time. This model has been generalized by Madi [625] to the multiple step-stress case. Xiong [928] and Xiong and Milliken [930] discussed inference for the exponential step-stress model by assuming that the mean lifetime of the experimental units at the ith stress level s_i has a log-linear form. But, Watkins [891] argued that it would be better to work with the original parameters of the exponential model even though a log-linear link function provides a simpler model for inferential purposes. Gouno et al. [408] and Han et al. [430] discussed inferential methods for step-stress models under the exponential distribution when the available data are progressively Type-I censored. Xiong and Ji [929] discussed the analysis of step-stress life tests based on grouped and censored data. While these discussions all focused on inference for exponential step-stress models, Khamis and Higgins [524] and Kateri and Balakrishnan [511] examined inferential methods for a cumulative exposure model with Weibull distributed lifetimes.

Comprehensive reviews of work on step-stress models have been provided by Gouno and Balakrishnan [407], Tang [837], and more recently by Balakrishnan [85]. It is important to mention that inference for specific step-stress models has also been discussed in the general framework of accelerated life testing (see, for example, Shaked and Singpurwalla [800], McNichols and Padgett [644], Lu and Storer [619], Van Dorp and Mazzuchi [866], and Bagdonavičius et al. [69]). Balakrishnan [85] presented an extensive survey on recent work on exact inferential procedures for exponential step-stress models for different types of censored data.

In the following, results on step-stress models are presented when the data are progressively censored. Moreover, we assume throughout that the stress changes are modelled by a cumulative exposure model. For details on this approach, we refer to Balakrishnan [85, Sect. 2].

23.1.1 Inference for Simple Step-Stress Model Under Progressive Type-II Censoring

Progressively Type-II Censored Step-Stress Data

Suppose a sample of n experimental units with lifetimes $T_1, \dots, T_n$ are placed on a simple step-stress test at an initial stress level s_1. At a prefixed time τ, the stress level is to be increased to level $s_2 > s_1$. Moreover, progressive Type-II censoring with censoring scheme $\mathscr{R} = (R_1, \dots, R_r)$ is employed in this experimental setting leading to the data $\mathbf{T}^{\mathscr{R}} = (T_{1:r:n}, \dots, T_{r:r:n})$ with

$$T_{1:r:n} < \cdots < T_{N_1:r:n} \le \tau < T_{N_1+1:r:n} < \cdots < T_{r:r:n}. \tag{23.1}$$

N_1 denotes the numbers of failures observed before time τ, whereas $N_2 = r - N_1$ counts the failures exceeding τ. The generation process is depicted in Fig. 23.2 (for $\mathscr{R} = (0^{*m})$, see also Fig. 23.1). Formally, the random variable N_1 is defined by

$$N_1 = \sum_{j=1}^{r} \mathbb{1}_{(-\infty,\tau]}(T_{j:r:n}). \tag{23.2}$$

Obviously, the probability mass function of N_1 with support $\{0, \dots, r\}$ can be obtained from the relation

$$P(N_1 = i) = P(T_{i:r:n} \le \tau < T_{i+1:r:n}), \quad i = 0, \dots, r,$$

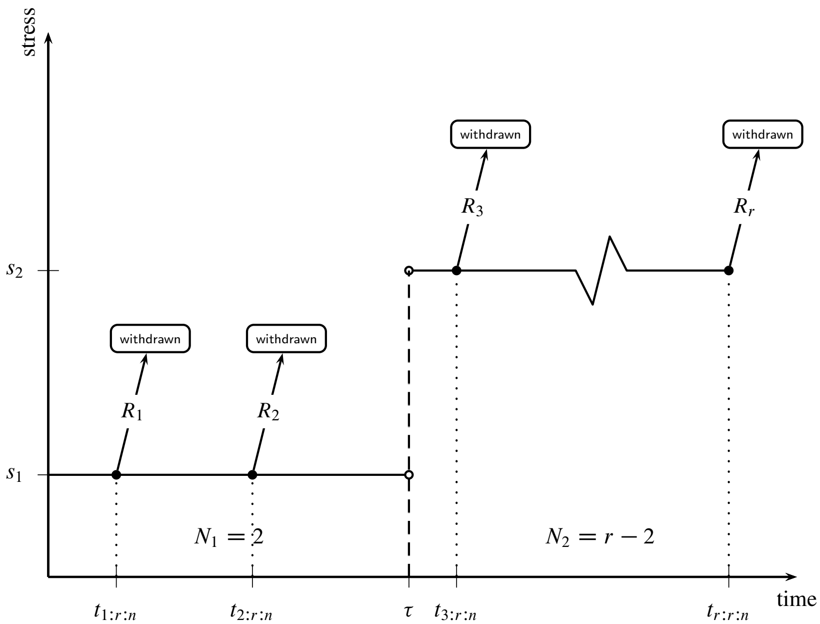

Fig. 23.2 Generation process of progressively Type-II censored stress level data with censoring scheme $\mathscr{R} = (R_1, \dots, R_r)$

where $T_{0:r:n} = -\infty$ and $T_{r+1:r:n} = \infty$. Then, we get from Lemma 2.5.4 and (2.41)

(i) $i = 0$: $P(N_1 = 0) = P(\tau < T_{1:r:n}) = 1 - F_{1:r:n}(\tau) = (1 - F(\tau))^n$;
(ii) $i \in \{1, \dots, r-1\}$: Applying Corollary 2.4.7, we arrive at

$$P(N_1 = i) = F_{i:r:n}(\tau) - F_{i+1:r:n}(\tau) = \frac{1}{\gamma_{i+1}}(1 - F(\tau)) f^{U_{i+1:r:n}}(F(\tau))$$
$$= \Big(\prod_{j=1}^{i} \gamma_j\Big) \sum_{j=1}^{i+1} a_{j,i+1}(1 - F(\tau))^{\gamma_j};$$

(iii) $i = r$: $P(N_1 = r) = P(T_{r:r:n} \leq \tau) = F_{r:r:n}(\tau)$.

These expressions will be useful in the derivation of the exact (conditional) distributions of the MLEs in the exponential case. The distribution of N_2 is directly obtained via the identity $N_2 = r - N_1$.

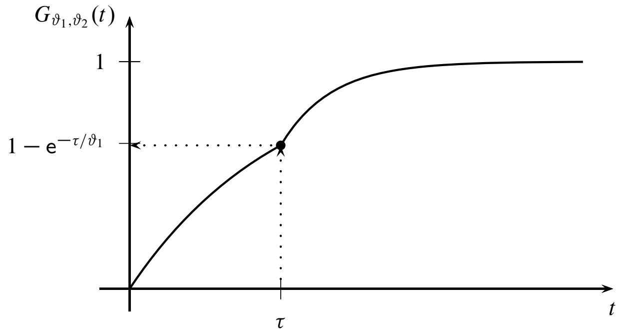

Fig. 23.3 Plot of $G_{\vartheta_1,\vartheta_2}$ in a cumulative exposure model with $\mathsf{Exp}(\vartheta_1)$- and $\mathsf{Exp}(\vartheta_2)$-distributions

Likelihood Function and Maximum Likelihood Estimation

The likelihood inference for the parameters of a simple step-stress model based on a progressively Type-II censored exponential sample has been developed by Xie et al. [927]. They imposed a cumulative exposure model with $\mathsf{Exp}(\vartheta_1)$- and $\mathsf{Exp}(\vartheta_2)$-distributions. Then, the cumulative distribution function is given by

$$G_{\vartheta_1,\vartheta_2}(t) = \begin{cases} F_{\vartheta_1}(t) = 1-\mathrm{e}^{-t/\vartheta_1}, & 0 \le t < \tau \\ F_{\vartheta_2}\Big(t-\big(1-\frac{\vartheta_2}{\vartheta_1}\big)\tau\Big) = 1-\mathrm{e}^{-t/\vartheta_2+\big(\frac{1}{\vartheta_2}-\frac{1}{\vartheta_1}\big)\tau}, & \tau \le t \end{cases}. \tag{23.3}$$

A plot of $G_{\vartheta_1,\vartheta_2}$ is depicted in Fig. 23.3. The corresponding density function is given by

$$g_{\vartheta_1,\vartheta_2}(t) = \begin{cases} f_{\vartheta_1}(t) = \frac{1}{\vartheta_1}\mathrm{e}^{-t/\vartheta_1}, & 0 \le t < \tau \\ f_{\vartheta_2}\Big(t-\big(1-\frac{\vartheta_2}{\vartheta_1}\big)\tau\Big) = \frac{1}{\vartheta_2}\mathrm{e}^{-(t-\tau)/\vartheta_2-\tau/\vartheta_1}, & \tau \le t \end{cases}.$$

Depending on the values of N_1, the likelihood function of ϑ_1 and ϑ_2, based on the progressively Type-II censored sample in (23.1), has different forms. Moreover, for $r = n_1 + n_2$ $(2 \le r \le n)$, we define the quantities

$$\begin{aligned} D_1 &= \sum_{k=1}^{n_1}(R_k+1)t_{k:r:n} + \tau\gamma_{n_1+1}, \\ D_2 &= \sum_{k=n_1+1}^{r}(R_k+1)(t_{k:r:n}-\tau), \end{aligned} \tag{23.4}$$

where we use the convention that $\sum_{k=p}^{q} = 0$ whenever $p > q$. Notice that D_1 and D_2 decompose the total time on test $D_1 + D_2$. These random variables can be interpreted as time on tests w.r.t. the stress levels with D_1 representing the time on test at stress level s_1 and D_2 corresponding to the part with level s_2.

With N_1 as in (23.2), the likelihood function can be written as

$$\mathsf{L}(\vartheta_1, \vartheta_2 | \mathbf{t}_r) = \frac{c_{r-1}}{\vartheta_1^{n_1} \vartheta_2^{n_2}} \exp\left\{-\frac{1}{\vartheta_1} D_1 - \frac{1}{\vartheta_2} D_2\right\}. \tag{23.5}$$

Depending on the value of N_1, the likelihood function exhibits a simpler form.

(i) First, let $N_1 = r$ and $N_2 = 0$, i.e., all failures occur before τ so that no information is available on the right of τ. This means that no information is available on ϑ_2. In this case, $t_{1:r:n} < \cdots < t_{r:r:n} \leq \tau$ and the likelihood function reads

$$\mathsf{L}(\vartheta_1, \vartheta_2 | \mathbf{t}_r) = \frac{c_{r-1}}{\vartheta_1^{r}} \exp\left\{-\frac{1}{\vartheta_1} D_1\right\}. \tag{23.6}$$

(ii) For $N_1 = 0$ and $N_2 = r$, the data in (23.1) are given by $\tau < t_{1:r:n} < \cdots < t_{r:r:n} < \infty$ leading to the likelihood

$$\mathsf{L}(\vartheta_1, \vartheta_2 | \mathbf{t}_r) = \frac{c_{r-1}}{\vartheta_2^{r}} \exp\left\{-\frac{1}{\vartheta_2} D_2 - \frac{n\tau}{\vartheta_1}\right\}. \tag{23.7}$$

(iii) For $N_1 = n_1 \in \{1, \ldots, r-1\}$, we have observations below and exceeding τ so that $0 < t_{1:r:n} < \cdots < t_{n_1:r:n} \leq \tau < t_{n_1+1:r:n} < \cdots < t_{r:r:n} < \infty$. Then, the likelihood function is given by (23.5) with $D_1, D_2 > 0$.

The expressions of the likelihood function in (23.6) and (23.7) show that MLEs of the parameters do not exist in any case. The results are summarized in the next theorem.

Theorem 23.1.1. (i) The likelihood in (23.6) does not involve ϑ_2 so that the MLE of ϑ_2 does not exist for $N_1 = r$ (and $N_2 = 0$);
(ii) The likelihood in (23.7) is increasing ϑ_1, so that the MLE of ϑ_1 does not exist for $N_1 = 0$ (and $N_2 = r$);
(iii) Finally, if at least one failure occurs before τ and after τ, the MLEs of ϑ_1 and ϑ_2 exist. They are given by

$$\widehat{\vartheta}_1 = \frac{D_1}{N_1}, \quad \widehat{\vartheta}_2 = \frac{D_2}{N_2}. \tag{23.8}$$

Notice that in the settings (23.6) and (23.7), the existing MLEs take on the same form.

Exact Conditional Distributions of MLEs

Xie et al. [927] established the exact conditional distributions of the MLEs $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$ given in (23.8). In the following, we will present a different approach leading to simpler representations of the density functions. In particular, we utilize the block independence property and a connection to Type-I progressively hybrid censored data. The first part of the sample, i.e., $T_{1:r:n} < \cdots < T_{N_1:r:n} \le \tau$, can be interpreted as a Type-I progressively hybrid censored sample from an Exp(ϑ_1)-distribution. Moreover, the MLE $\widehat{\vartheta}_1$ of ϑ_1 depends only on this subsample. It exists provided that $N_1 > 0$. This observation leads directly to the following result taken from Theorem 5.1.4 and (14.3).

Theorem 23.1.2. Let $1 \le n_1 \le r$. Then, the conditional density function of $\widehat{\vartheta}_1$, given $N_1 = n_1 < r$, is given by

$$f^{\widehat{\vartheta}_1|N_1=n_1}(t) = \frac{\tau^{n_1}\prod_{j=1}^{n_1+1}\gamma_j}{d!\vartheta^{n_1+1}f_{n_1+1:m:n}(\tau)} B_{n_1-1}(t|\gamma_{n_1+1}\tau,\ldots,\gamma_1\tau)\mathrm{e}^{-t/\vartheta},$$
$$0 \le t \le n\tau.$$

For $N_1 = r$, the density function is given by

$$f^{\widehat{\vartheta}_1|N_1=r}(t) = \frac{\tau^r\prod_{j=1}^{r}\gamma_j}{r!\vartheta^r F_{r:r:n}(T)} B_{r-1}(t|\gamma_{r+1}\tau,\ldots,\gamma_1\tau)\mathrm{e}^{-t/\vartheta}, \quad 0 \le t \le n\tau.$$

The conditional density of $\widehat{\vartheta}_1$, given $1 \le N_1$, is

$$f^{\widehat{\vartheta}_1|N_1\ge 1}(t) = \frac{1}{1-\mathrm{e}^{-n\tau/\vartheta}} \sum_{d=1}^{r} \Big[\prod_{j=1}^{d}\gamma_j\Big] \frac{\tau^d}{(d-1)!\vartheta^d}$$
$$\times B_{d-1}(dt|\gamma_{d+1}\tau,\ldots,\gamma_1\tau)\mathrm{e}^{-dt/\vartheta}, \quad t \ge 0.$$

Using the block independence result established in Theorem 2.5.5, we find that, given $N_1 = n_1 \in \{1,\ldots,r-1\}$, the step-stress sample from (23.1) separates into two (conditionally) independent samples

$$T_{1:r:n} < \cdots < T_{n_1:r:n} \le \tau \quad \text{and} \quad \tau < T_{n_1+1:r:n} < \cdots < T_{r:r:n}.$$

Since $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$ are only functions of the first and second subsample, respectively, it follows directly that the estimators are conditionally independent. Moreover, we conclude from this result that the sample $T_{n_1+1:r:n},\ldots,T_{r:r:n}$ forms progressively Type-II censored order statistics $W_{1:r-n_1:\gamma_{n_1}},\ldots,W_{r-n_1:r-n_1:\gamma_{n_1}}$ from a left-truncated cumulative distribution function $G_{\tau;\vartheta_1,\vartheta_2}$ defined by $G_{\tau;\vartheta_1,\vartheta_2}(t) = 1 - \frac{1-G_{\vartheta_1,\vartheta_2}(t)}{1-G_{\vartheta_1,\vartheta_2}(\tau)}$, $t \ge \tau$, with $G_{\vartheta_1,\vartheta_2}$ defined in (23.3). Simplifying this expression shows that

$$G_{\tau;\vartheta_1,\vartheta_2}(t) = 1 - \mathrm{e}^{-(t-\tau)/\vartheta_2}, \quad t \ge \tau,$$

saying that the sample $T_{n_1+1:r:n}, \ldots, T_{r:r:n}$ is distributed as the first $r - n_1$ $\mathrm{Exp}(\tau, \vartheta_2)$-progressively Type-II censored order statistics from a sample of size γ_{n_1} with known location parameter τ and censoring scheme $(R_{n_1+1}, \ldots, R_r)$. Hence, we can apply the results of Theorem 12.1.1 because $\widehat{\vartheta}_2$ is seen to have the same form as the MLE in the one-sample setting considered in Sect. 12.1. Theorem 23.1.3 summarizes these findings.

Theorem 23.1.3. Let $n_1 \in \{0, \ldots, r\}$. Then:

(i) Given $N_1 = n_1 \in \{1, \ldots, r-1\}$, the MLEs $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$ are conditionally independent;
(ii) For $n_1 < r$, $2(r-n_1)\widehat{\vartheta}_2/\vartheta_2 | N_1 = n_1 \sim \chi^2(2r-2n_1)$.
In particular, $E(\widehat{\vartheta}_2 | N_1 = n_1) = \vartheta_2$ and $\mathrm{Var}(\widehat{\vartheta}_2 | N_1 = n_1) = \vartheta_2^2/(r-n_1)$.

The MLE of ϑ_2 exists provided that $N_2 > 0$, i.e., $N_1 < r$. Then, we get the following result established by Xie et al. [927]. It shows that the conditional density function of $\widehat{\vartheta}_2$ is a mixture of scaled χ^2-distributions.

Theorem 23.1.4. Let $N_1 \le r-1$ and $\widehat{\vartheta}_2$ be the MLE of ϑ_2. Then, for $t \in \mathbb{R}$, the conditional density of $\widehat{\vartheta}_2$, given $0 \le N_1 \le r-1$, has the form

$$f^{\widehat{\vartheta}_2 | N_1 < r}(t) = \frac{2}{\vartheta_2 \overline{F}_{r:r:n}(\tau)} \sum_{d=0}^{r-1} (r-d) P(N_1 = d) f_{\chi^2(2r-2d)}\Big(\frac{2(r-d)t}{\vartheta_2}\Big).$$

The conditional densities of $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$ readily imply the following expressions for mean and variance of $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$. The results for $\widehat{\vartheta}_1$ can be taken from (5.15) and (5.16). Notice that the (conditional) distributions of N_1 and N_2, respectively, typically depend on the parameter ϑ_1.

Corollary 23.1.5. The (conditional) mean and variance of $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$ are given by

$$E(\widehat{\vartheta}_1 | N_1 > 0) = \vartheta_1 - \frac{n\tau \mathrm{e}^{-n\tau/\vartheta_1}}{(1-\mathrm{e}^{-\tau/\vartheta_1})^n} + \frac{\vartheta_1 \tau}{(1-\mathrm{e}^{-\tau/\vartheta_1})^n} \sum_{d=2}^{m} \frac{1}{d(d-1)} f_{d:m:n}(\tau), \tag{23.9}$$

$$\mathrm{MSE}(\widehat{\vartheta}_1 | N_1 > 0) = \vartheta_1^2 E\Big(\frac{1}{N_1}\Big| N_1 > 0\Big) + \frac{n\tau^2 \mathrm{e}^{-n\tau/\vartheta_1}}{(1-\mathrm{e}^{-\tau/\vartheta_1})^n} - \frac{\vartheta_1^2 \tau^2}{(1-\mathrm{e}^{-\tau/\vartheta_1})^n} \sum_{d=2}^{m} \frac{2d-1}{d^2(d-1)^2} f'_{d:m:n}(\tau), \tag{23.10}$$

$$E(\widehat{\vartheta}_2|N_1 < r) = \vartheta_2 \tag{23.11}$$

and

$$\begin{aligned}\mathrm{Var}(\widehat{\vartheta}_2|N_1 < r) &= \vartheta_2^2 \sum_{d=0}^{r-1} P(N_1 = d|N_1 < r)\frac{1}{r-d} \\ &= \vartheta_2^2 E\Big(\frac{1}{r-N_1}\Big|N_1 < r\Big) = \vartheta_2^2 E\Big(\frac{1}{N_2}\Big|N_2 > 0\Big).\end{aligned} \tag{23.12}$$

The expressions of the means in (23.9) and (23.11) show that $\widehat{\vartheta}_1$ is a biased estimator of ϑ_1, while $\widehat{\vartheta}_2$ is an unbiased estimator of ϑ_2. The expressions in (23.10) and (23.12) can be readily used to compute standard errors of the MLEs $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$.

Confidence Intervals for ϑ_1 and ϑ_2

Xie et al. [927] proposed several methods of constructing confidence intervals, including an exact, an approximate, and a bootstrap-based approach. Under the assumption that Open Problem 14.1.2 holds, Xie et al. [927] constructed a two-sided exact $100(1-\alpha)\%$ confidence interval for ϑ_1 by solving the equations

$$P_{\vartheta_{1\ell}}(\widehat{\vartheta}_1 > \vartheta_{1\mathrm{obs}}) = \frac{\alpha}{2}, \quad P_{\vartheta_{1u}}(\widehat{\vartheta}_1 > \vartheta_{1\mathrm{obs}}) = 1 - \frac{\alpha}{2} \tag{23.13}$$

for $\vartheta_{1\ell}, \vartheta_{1u}$, where $\vartheta_{1\mathrm{obs}}$ is the observed value of $\widehat{\vartheta}_1$. Notice that the method of pivoting the cumulative distribution function as described on p. 331 is based on the assumption that the equations in (23.13) have solutions. Therefore, the same comments apply as presented on p. 331 (see also Balakrishnan et al. [156]).

Using the same approach, confidence intervals for ϑ_2 can be obtained given a similar assumption on the monotonicity of the (conditional) survival function. But, this property can be proved in the present setting as has been done in Balakrishnan and Iliopoulos [102, Sect. 4.2] for the case of Type-II censored step-stress data.

Applying the observed Fisher information matrix as well as asymptotic properties of maximum likelihood estimators, Xie et al. [927] also proposed $100(1-\alpha)\%$ approximate confidence intervals for ϑ_1 and ϑ_2 using a normal approximation. Finally, they investigated three bootstrap methods: studentized t-interval, percentile interval, and BCA interval. For a detailed discussion as well as comparisons through Monte Carlo simulations, we refer to Xie et al. [927] and Balakrishnan [85].

Optimal Censoring and Optimal Test Plan

For censoring schemes $\mathscr{R} = (R_1, \ldots, R_r)$, Xie et al. [927] discussed optimal progressive censoring in a simple step-stress test. Given $1 \le N_1 \le r-1$, they considered two optimality criteria.

(1) Minimum variance:

$$\psi(\mathscr{R}) = \mathsf{Var}(\widehat{\vartheta}_1 | 1 \le N_1 \le r-1) + \mathsf{Var}(\widehat{\vartheta}_2 | 1 \le N_1 \le r-1) \longrightarrow \min_{\mathscr{R}};$$

(2) Minimum mean squared error:

$$\varphi(\mathscr{R}) = \mathsf{MSE}(\widehat{\vartheta}_1 | 1 \le N_1 \le r-1) + \mathsf{Var}(\widehat{\vartheta}_2 | 1 \le N_1 \le r-1) \longrightarrow \min_{\mathscr{R}}.$$

Assuming $1 \le N_1 \le r-1$ and using the explicit expressions for the variances and mean squared errors given in Corollary 23.1.5, Xie et al. [927, Tables 9.5, 9.6] determined best and worse censoring schemes for the parameter choice $\vartheta_1 = \mathrm{e}^{1.5}$, $\vartheta_2 = \mathrm{e}^{0.5}$ and selected values of τ. Moreover, they discussed the problem of finding an optimal time τ^* for changing the stress level from s_1 to s_2 given a specified choice of n, r, censoring scheme $\mathscr{R} = (R_1, \ldots, R_r)$, and pre-fixed parameter values ϑ_1 and ϑ_2 (see Table 9.7 in Xie et al. [927]).

Open problem 23.1.6. For Type-II censored data, Xiong [928] introduced a link function which assumes the logarithms of the parameters to be a linear function of the stress level

$$\log \vartheta_i = \alpha + \beta s_i, \quad i = 1, 2, \tag{23.14}$$

where α, β are unknown parameters. An extension to progressively censored data is not available so far.

Remark 23.1.7. Yang and Tse [933] discussed a step-stress model for progressive Type-I interval censored data with random removals and a log-linear link function between the stress levels and the parameters. Weibull lifetimes are investigated in Tse et al. [861] (see also Ding et al. [341] and Ding and Tse [340]). Burr-XII distributions with progressively Type-II censored data in the presence of random removals are discussed in Sun et al. [828].

23.1.2 Inference for a Simple Step-Stress Model with Random Change Under Progressive Type-II Censoring

In the previous subsection, the stress has been changed at a pre-fixed time τ. Now, the stress is changed at the r_1th failure time as depicted in Fig. 23.4. The experiment terminates at the rth failure time where $1 \le r_1 < r$. This approach has the advantage

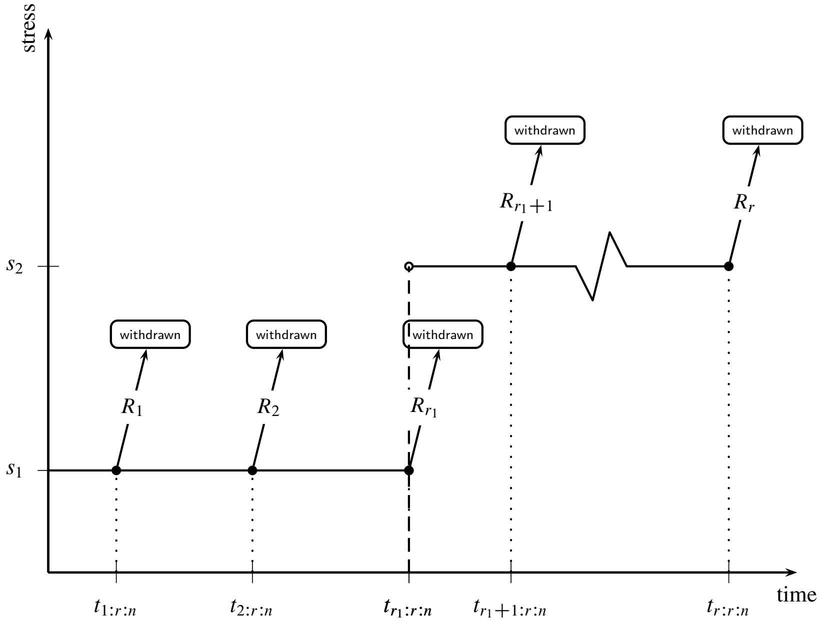

Fig. 23.4 Generation process of progressively Type-II censored stress level data with random change time

that the number of observed failures at each stress level is fixed and no longer random.

This model has been discussed in Wang and Yu [889] (for similar results with Type-II censored data, see Kundu and Balakrishnan [558]). Introducing the quantities (see (23.4))

$$\begin{aligned} D_1 &= \sum_{k=1}^{r_1-1}(R_k+1)t_{k:r:n} + \gamma_{r_1}t_{r_1:r:n}, \\ D_2 &= \sum_{k=r_1+1}^{r-1}(R_k+1)(t_{k:r:n}-t_{r_1:r:n}) + \gamma_r(t_{r:r:n}-t_{r_1:r:n}), \end{aligned} \tag{23.15}$$

Wang and Yu [889] showed that the likelihood function is proportional to

$$\mathsf{L}(\vartheta_1,\vartheta_2) \propto \vartheta_1^{-r_1}\vartheta_2^{r_1-r}\exp\Big\{-\frac{D_1}{\vartheta_1}-\frac{D_2}{\vartheta_2}\Big\}.$$

Obviously, the MLEs are given by

$$\widehat{\vartheta}_1 = \frac{D_1}{r_1}, \quad \widehat{\vartheta}_2 = \frac{D_2}{r - r_1}. \tag{23.16}$$

Connection to Sequential Order Statistics

Combining the ideas of Balakrishnan et al. [152] and Beutner [192], the Wang-Yu step-stress model can also be motivated by an approach based on sequential order statistics. In particular, the data can be seen as sequential order statistics or generalized order statistics. To be more precise, we describe the model using the notation of Cramer and Kamps [301]. Given a population cumulative distribution function F, let the cumulative distribution functions $F_1, \dots, F_r$ be defined as

$$F_j(t) = \begin{cases} 1 - \overline{F}^{\gamma_j((n-j+1)\vartheta_1)}(t), & j = 1, \dots, r_1 \\ 1 - \overline{F}^{\gamma_j((n-j+1)\vartheta_2)}(t), & j = r_1 + 1, \dots, r \end{cases}. \tag{23.17}$$

Then, the joint density function of sequential order statistics $X^*_{(1)}, \dots, X^*_{(r)}$ is given by

$$f^{X^*_{(1)},\dots,X^*_{(r)}}(\mathbf{t}_r) = \vartheta_1^{-r_1}\vartheta_2^{r_1-r} \prod_{j=1}^{r} \gamma_j \Big(\prod_{j=1}^{r-1} \overline{F}^{m_j}(t_j) f(t_j) \Big) \overline{F}^{\gamma_r/\vartheta_2 - 1}(t_r) f(t_r), \tag{23.18}$$

where

$$m_j = \begin{cases} \frac{\gamma_j}{\vartheta_1} - \frac{\gamma_{j+1}}{\vartheta_1} - 1, & j = 1, \dots, r_1 - 1 \\ \frac{\gamma_j}{\vartheta_1} - \frac{\gamma_{j+1}}{\vartheta_2} - 1, & j = r_1 \\ \frac{\gamma_j}{\vartheta_2} - \frac{\gamma_{j+1}}{\vartheta_2} - 1, & j = r_1 + 1, \dots, r - 1 \end{cases}.$$

Obviously, the density function (23.18) is the joint density function of generalized order statistics with parameters $m_1, \dots, m_{r-1}$ and γ_r from the population cumulative distribution function F (see (2.7)).

For exponential lifetimes, the density function (23.18) simplifies since some parts can be combined. In particular, we get

$$\begin{aligned} &\sum_{j=1}^{r-1}(m_j + 1)t_j + \frac{\gamma_r}{\vartheta_2}t_r \\ &= \frac{1}{\vartheta_1}\sum_{j=1}^{r_1-1}(R_j + 1)t_j + \frac{\gamma_{r_1}}{\vartheta_1}t_{r_1} - \frac{\gamma_{r_1+1}}{\vartheta_2}t_{r_1} + \frac{1}{\vartheta_2}\sum_{j=r_1+1}^{r-1}(R_j + 1)t_j + \frac{\gamma_r}{\vartheta_2}t_r \\ &= \frac{D_1}{\vartheta_1} + \frac{D_2}{\vartheta_2} \end{aligned}$$

with D_1 and D_2 given in (23.15). Therefore, the joint density function (23.18) reads

$$f^{T_{1:r:.n},\ldots,T_{r:r:n}}(\mathbf{t}_r) = \Big(\prod_{j=1}^{r}\gamma_j\Big)\vartheta_1^{-r_1}\vartheta_2^{r_1-r}\exp\Big\{-\frac{D_1}{\vartheta_1}-\frac{D_2}{\vartheta_2}\Big\}.$$

This expression for the density function allows us to derive many properties of the corresponding failure times $T_{1:r:.n},\ldots,T_{r:r:n}$. In particular, we can directly apply properties of generalized order statistics from standard exponential distribution and parameters $\gamma_1/\vartheta_1,\ldots,\gamma_{r_1-1}/\vartheta_1,\gamma_{r_1+1}/\vartheta_2,\ldots,\gamma_r/\vartheta_2$ as given in Kamps [498,499] and Cramer and Kamps [301]. For instance, normalized spacings of generalized order statistics are independent exponentially random variables. Therefore, we have

$$\gamma_j^*(T_{j:r:n}-T_{j-1:r:n}),\quad j=1,\ldots,r,$$

where $T_{0:r:n} = 0$ and $\gamma_j^* = \gamma_j/\vartheta_1$, $j = 1,\ldots,r_1-1$, and $\gamma_j^* = \gamma_j/\vartheta_2$, $j=r_1,\ldots,r$, as IID standard exponential random variables. Moreover, the random variables D_1 and D_2 have the representation

$$D_1 = \sum_{j=1}^{r_1}\gamma_j(T_{j:r:n}-T_{j-1:r:n}) \text{ and } D_2 = \sum_{j=r_1+1}^{r}\gamma_j(T_{j:r:n}-T_{j-1:r:n})$$

in terms of the spacings. This yields the following properties which are similar to those presented in Theorem 2 of Balakrishnan et al. [152] for Type-II censored step-stress data.

Theorem 23.1.8. Let $\widehat{\vartheta}_1$ and $\widehat{\vartheta}_2$ be the MLEs of ϑ_1 and ϑ_2 as given in (23.16). Then,

(i) $(\widehat{\vartheta}_1, \widehat{\vartheta}_2)$ is a complete and sufficient statistic for $(\vartheta_1,\vartheta_2)$;
(ii) $\widehat{\vartheta}_1$, $\widehat{\vartheta}_2$ are independent;
(iii) $\widehat{\vartheta}_j$ is an unbiased estimator of ϑ_j, $j=1,2$;
(iv) $\widehat{\vartheta}_j$ is the UMVUE of ϑ_j, $j=1,2$;
(v) The distributions of the MLEs are given by

$$\frac{2r_1\widehat{\vartheta}_1}{\vartheta_1}\sim\chi^2(2r_1),\quad \frac{2(r-r_1)\widehat{\vartheta}_2}{\vartheta_2}\sim\chi^2(2r-2r_1);$$

(vi) The estimators are strongly consistent and asymptotically normal for $r_1\to\infty$ and $r-r_1\to\infty$, respectively.

Parametrization via Log-Linear Link Function

Wang and Yu [889] introduced a log-linear link function for the parameters of the step-stress model as given in (23.14). Then, the MLEs of α and β are obtained as

$$\widehat{\alpha} = \frac{s_1 \log(D_2/(r-r_1)) - s_2 \log(D_1/r_1)}{s_1 - s_2},$$
$$\widehat{\beta} = \frac{\log(D_1/r_1) - \log(D_2/(r-r_1))}{s_1 - s_2}.$$

Moreover, given a stress level s_0, the MLEs of the mean life $\vartheta_0 = \exp(\alpha + \beta s_0)$ and the hazard rate $\lambda_0 = 1/\vartheta_0$ are given by

$$\widehat{\vartheta}_0 = \exp(\widehat{\alpha} + \widehat{\beta} s_0), \quad \widehat{\lambda}_0 = \exp(-\widehat{\alpha} - \widehat{\beta} s_0).$$

Wang and Yu [889] also considered minimum variance unbiased estimation and showed that the UMVUE of ϑ_0 is given by

$$\widetilde{\vartheta}_0 = \frac{\Gamma(r_1)\Gamma(r_2)}{\Gamma(r_1 + k_0 + 1)\Gamma(r_2 - k_0)} D_1^{k_0+1} D_2^{k_0}$$

provided $r_2 > k_0 = \frac{s_0 - s_1}{s_1 - s_2}$. The UMVUE of λ_0 is given by

$$\widetilde{\lambda}_0 = \frac{\Gamma(r_1)\Gamma(r_2)}{\Gamma(r_1 - k_0 - 1)\Gamma(r_2 + k_0)} D_1^{-k_0-1} D_2^{k_0}$$

if $r_1 > k_0 + 1$. Expressions for the variances of the UMVUEs and MLEs are also given in Wang and Yu [889]. Finally, it is shown that the mean squared error of an MLE exceeds that of the corresponding UMVUE. These results are based on the properties given in Theorem 23.1.8.

Wang [884] established confidence intervals for α, β, and ϑ_0. Obviously, $\widehat{\beta}$ has a transformed F-distribution since $\widehat{\beta}$ can be written as

$$\widehat{\beta} = \frac{\log Q}{s_1 - s_2} \quad \text{with } Q = \frac{D_1/r_1}{D_2/(r-r_1)} \text{ and } Q \cdot \frac{\vartheta_2}{\vartheta_1} \sim \mathsf{F}(2r_1, 2(r-r_1)).$$

Hence, the statistical interval

$$\left[\widehat{\beta} - \frac{\log \mathsf{F}_{p/2}(2r_1, 2(r-r_1))}{s_1 - s_2}, \widehat{\beta} - \frac{\log \mathsf{F}_{1-p/2}(2r_1, 2(r-r_1))}{s_1 - s_2}\right]$$

forms a $100(1-p)\%$ exact confidence interval for β. Notice that $\mathrm{e}^{(s_1-s_2)\beta} = \vartheta_1/\vartheta_2$. Expressions for confidence intervals for α and ϑ_0 can be written in terms of quantiles of the cumulative distribution functions

$$F_{W_\alpha}(w) = \int_0^\infty f_{\chi^2(2r_1)}(t) F_{\chi^2(2(r-r_1))}\big((wt^{k_1})^{1/(k_1+1)}\big) dt,$$
$$F_{W_\vartheta}(w) = \int_0^\infty f_{\chi^2(2(r-r_1))}(t) F_{\chi^2(2r_1)}\big((wt^{k_0})^{1/(k_0+1)}\big) dt,$$

where $k_0 = \frac{s_0 - s_1}{s_1 - s_2}$, $k_1 = \frac{s_2}{s_1 - s_2}$, $W_\alpha = 2(r-r_1)^{k_1+1} r_1^{-k_1} \mathrm{e}^{\widehat{\alpha - \alpha}}$, and $W_\vartheta = 2r_1^{k_0+1}(r - r_1)^{-k_0} \widehat{\vartheta}_0/\vartheta_0$.

23.1.3 Inference for Multiple Step-Stress Model Under Progressive Type-I Censoring

Progressively Type-I Censored Step-Stress Data

In progressive Type-I censoring, R_i surviving units are censored at pre-fixed times $0 < \tau_1 < \tau_2 < \cdots < \tau_k$. This setting can be extended to a step-stress model by introducing different stress levels $s_1 < \cdots < s_k$. Initially, the stress level imposed on the n units on test is at a level of s_1. Then, at a pre-fixed time τ_1, the stress level is changed to s_2; next, at time τ_2, the stress level is changed to s_3, etc. Hence, at censoring time τ_j,

(i) A prefixed number R_j of items is withdrawn from the experiment provided that R_j items are still on test. Otherwise, the available items are removed and the experiment is terminated. The effectively applied censoring scheme is denoted by $\mathscr{R}^* = (R_1^*, \ldots, R_k^*)$;
(ii) The stress level is increased from s_j to s_{j+1}.

This scenario is depicted in Fig. 23.5.

Gouno et al. [408] and Han et al. [430] discussed this model with equi-spaced censoring times, i.e., when $\tau_j = j\tau$, $j = 1, \ldots, k$. The following notations are used:

(1) n_j denotes the number of units that failed at stress level s_j, i.e., in the time interval $[(j-1)\tau, j\tau)$;
(2) R_j denotes the number of surviving units progressively censored at time $\tau_j = j\tau$;
(3) $t_{1,j:n} \le \cdots \le t_{n_j,j:n}$ denote the increasingly ordered failure times observed at stress level s_j, $j = 1, \ldots, k$;
(4) $\pi_j = R_j/n$, $j = 1, \ldots, k-1$, denote the proportion of units censored at τ_j;
(5) $N_j = n - \sum_{j=1}^{i-1} n_j - \sum_{j=1}^{i-1} R_j$ denotes the number of units remaining in the experiment at τ_j.

Likelihood Function and MLEs

Gouno et al. [408] and Han et al. [430] considered an exponential lifetime distribution $\mathrm{Exp}(\vartheta_i)$ at stress level s_i, where the mean lifetimes ϑ_i satisfy a log-linear link function of the form (23.14), i.e., $\log \vartheta_i = \alpha + \beta s_i$, $i = 1, \ldots, k$, and α and β are unknown parameters. Then, under the cumulative exposure model, the lifetime density function $g_{\boldsymbol{\vartheta}}$ of a unit is given by (see (23.3))

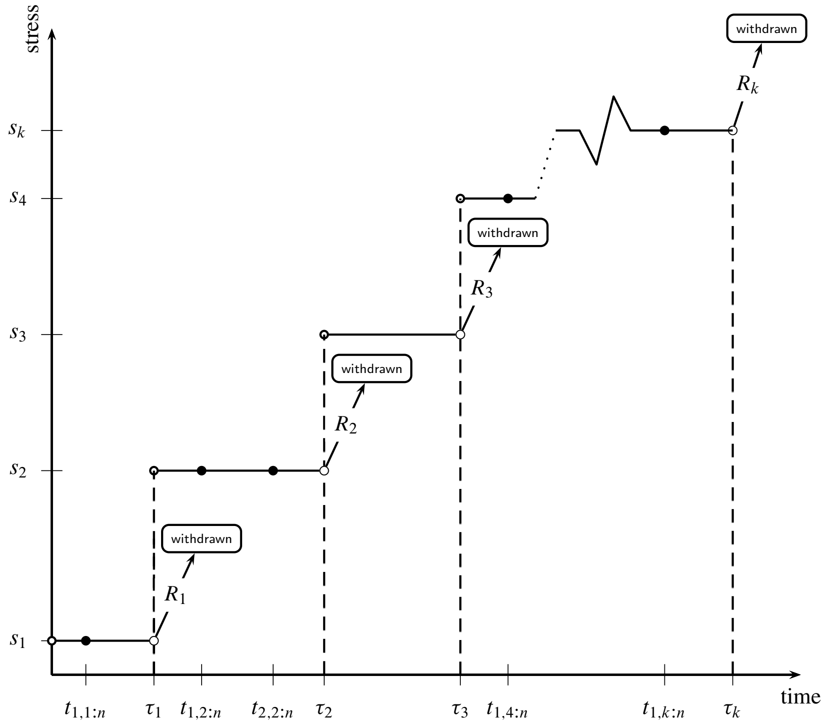

Fig. 23.5 Generation process of progressively Type-I censored order statistics in a step-stress scenario. The observations are denoted by $t_{1,j:n} \le \cdots \le t_{n_j,j:n}$ in each interval (τ_{j-1}, τ_j) provided that the number of observations n_j is positive

$$g_{\boldsymbol{\vartheta}}(t) = \begin{cases} \frac{1}{\vartheta_1} \exp\left\{-\frac{t}{\vartheta_1}\right\}, & 0 \le t \le \tau \\ \frac{1}{\vartheta_2} \exp\left\{-\frac{t-\tau}{\vartheta_2} - \frac{\tau}{\vartheta_1}\right\}, & \tau \le t \le 2\tau \\ \frac{1}{\vartheta_3} \exp\left\{-\frac{t-2\tau}{\vartheta_3} - \frac{\tau}{\vartheta_2} - \frac{\tau}{\vartheta_1}\right\}, & 2\tau \le t \le 3\tau \\ \vdots & \vdots \\ \frac{1}{\vartheta_k} \exp\left\{-\frac{t-(k-1)\tau}{\vartheta_k} - \sum_{j=1}^{k-1} \frac{\tau}{\vartheta_j}\right\}, & (k-1)\tau \le t \le \infty \end{cases} . \tag{23.19}$$

Then, given the data $\mathbf{n} = (n_1, \ldots, n_k)$ and $\mathbf{t} = (\mathbf{t}_1, \ldots, \mathbf{t}_k)$ with $\mathbf{t}_j = (t_{1,j:n}, \ldots, t_{n_j,j:n})$, the likelihood function is given by

$$\mathsf{L}(\vartheta_1, \ldots, \vartheta_k) = \prod_{i=1}^{k} \frac{N_i!}{(N_i - n_i)!} \prod_{i=1}^{k} \vartheta_i^{-n_i} \exp\left(-\sum_{i=1}^{k} \frac{U_i}{\vartheta_i}\right),$$

where

$$U_i = \sum_{j=1}^{n_i} (t_{j,i:n} - (i-1)\tau) + (N_i - n_i)\tau, \quad i = 1, \ldots, k.$$

Then, the MLEs of ϑ_i are given by $\widehat{\vartheta}_i = U_i/n_i$ provided that $n_i > 0$. Using the log-linear link function in (23.14), the corresponding log-likelihood function reads

$$\ell(\alpha, \beta) = \text{const} - \alpha \sum_{i=1}^{k} n_i - \beta \sum_{i=1}^{k} n_i s_i - \mathrm{e}^{-\alpha} \sum_{i=1}^{k} U_i \mathrm{e}^{-\beta s_i}.$$

Assuming that we observe at least one failure, i.e., $\sum_{i=1}^{k} n_i \geq 1$, $\ell(\alpha, \beta)$ can be bounded from above by

$$\ell(\widehat{\alpha}(\beta), \beta) = \text{const} - \log\left(\sum_{i=1}^{k} U_i \mathrm{e}^{-\beta s_i}\right) \sum_{i=1}^{k} n_i - \beta \sum_{i=1}^{k} n_i s_i,$$

where equality holds for $\alpha = \widehat{\alpha}(\beta)$ with

$$\widehat{\alpha}(\beta) = \log\left(\frac{\sum_{i=1}^{k} U_i \mathrm{e}^{-\beta s_i}}{\sum_{i=1}^{k} n_i}\right).$$

Now, it can be shown that $h(\beta) = -\log\left(\sum_{i=1}^{k} U_i \mathrm{e}^{-\beta s_i}\right)$ is a strictly concave function. Furthermore, defining $p_i = U_i \mathrm{e}^{-\beta s_i} / \sum_{j=1}^{k} U_j \mathrm{e}^{-\beta s_j}$, the second derivative of h w.r.t. β can be written as

$$\frac{\partial^2}{\partial \beta^2} h(\beta) = -\sum_{i=1}^{k} p_i s_i^2 + \left(\sum_{i=1}^{k} p_i s_i\right)^2 = -(ES^2 - E^2 S) = -\mathsf{Var}(S) \leq 0,$$

where S denotes a random variable with values in $\{s_1, \ldots, s_k\}$ and probability mass function $p_1, \ldots, p_k$. Therefore, a unique solution $\widehat{\beta}$ exists so that $\widehat{\alpha} = \widehat{\alpha}(\widehat{\beta})$ and $\widehat{\beta}$ are the MLEs of α and β. $\widehat{\beta}$ can be obtained as the solution of the nonlinear equation

$$\left[\sum_{i=1}^{k} n_i\right]\left[\sum_{i=1}^{k} U_i s_i \mathrm{e}^{-\widehat{\beta} s_i}\right] - \sum_{i=1}^{k} n_i s_i \sum_{i=1}^{k} U_i \mathrm{e}^{-\widehat{\beta} s_i} = 0.$$

Gouno et al. [408] based further statistical inference on the assumption that the asymptotic distribution of the MLE $(\widehat{\alpha}, \widehat{\beta})'$ is a bivariate normal distribution with mean $(\alpha, \beta)'$ and covariance matrix $[\mathscr{I}_n(\alpha, \beta)]^{-1}$, where $\mathscr{I}_n(\alpha, \beta)$ is the expected Fisher information matrix. A similar approach has been used in this area by many authors, e.g., by Nelson [675]. Gouno et al. [408] established the following representation of the expected Fisher information matrix:

$$\mathscr{I}_n(\alpha,\beta) = n\begin{pmatrix} \sum\limits_{i=1}^{k} A_i(\tau) & \sum\limits_{i=1}^{k} A_i(\tau)s_i \\ \sum\limits_{i=1}^{k} A_i(\tau)s_i & \sum\limits_{i=1}^{k} A_i(\tau)s_i^2 \end{pmatrix}, \tag{23.20}$$

where

$$A_i(\tau) = \left[1 - \sum_{j=1}^{i-1} \frac{\pi_j}{G_j(\tau)}\right] G_{i-1}(\tau)F_i(\tau), \quad 1 \le i \le k, \tag{23.21}$$

$$F_i(\tau) = 1 - \mathrm{e}^{-\tau/\vartheta_i}, \quad 1 \le i \le k,$$

$$G_j(\tau) = \prod_{i=1}^{j}(1 - F_i(\tau)), \quad 1 \le j \le k, \quad G_0(\tau) = 1.$$

Optimal Step-Stress Test

The choice of the stress change point τ has an impact on the bias and variance of the estimators. Therefore, it stands to reason to choose τ in an optimal way. Using the preceding results, Gouno et al. [408] proposed the following criteria based on the expression of the Fisher information in (23.20) or the variance–covariance matrix of (α,β) given by $(\mathscr{I}_n(\alpha,\beta))^{-1}$.

(i) Variance optimality: The first criterion is based on the idea of minimizing the variance of the mean life estimator $\widehat{\vartheta}_0$, given a stress s_0. As a measure for this quantity, Gouno et al. [408] considered the variance of $\log\widehat{\vartheta}_0$ given by

$$\begin{aligned}\phi(\tau) &= n\,\mathsf{Var}(\log\widehat{\vartheta}_0) = \mathsf{Var}(\widehat{\alpha} + \widehat{\beta}s_0) \\ &= n(1,s_0)\mathscr{I}_n^{-1}(\alpha,\beta)(1,s_0)' \\ &= \frac{2\sum_{i=1}^{k} A_i(\tau)(s_i - s_0)^2}{\sum_{i=1}^{m}\sum_{j=1}^{k} A_i(\tau)A_j(\tau)(s_i - s_j)^2} \longrightarrow \min_{\tau};\end{aligned} \tag{23.22}$$

(ii) D-optimality: Noticing that the volume of the asymptotic joint confidence region of (α,β) is proportional to the determinant of the square root of the inverse Fisher information matrix $|\mathscr{I}_n(\alpha,\beta)|^{-1/2}$, Gouno et al. [408] suggested a minimum volume approach. This yields the criterion

$$
\begin{aligned}
g(\tau) &= n^{-2}|\mathscr{I}_n(\alpha, \beta)| \\
&= \sum_{i=1}^{k} A_i(\tau) \sum_{j=1}^{k} A_j(\tau) s_j^2 - \left[\sum_{i=1}^{k} A_i(\tau) s_i \right]^2 \\
&= \frac{1}{2} \sum_{i=1}^{k} \sum_{j=1}^{k} A_i(\tau) A_j(\tau) (s_i - s_j)^2 \longrightarrow \min_{\tau} . \qquad (23.23)
\end{aligned}
$$

Gouno et al. [408] and Han et al. [430] addressed the existence and uniqueness of optimal τ for the simple step-stress model in these scenarios. Moreover, they provided some computational results on the optimal change time τ for different settings. However, given a prefixed censoring scheme $\mathscr{R} = (R_1, \ldots, R_{k-1})$, it may happen that either the experiment is terminated before all censoring steps are employed or no failures are observed before the termination time $k\tau$. These problems are typical for Type-I censoring. In order to overcome these difficulties, Gouno et al. [408] assumed a large sample size n, small global censoring proportions $\pi_i = R_i/n$, and a small number of stress levels k. These assumptions guided them to restrict the search for optimal τ to the region

$$
\mathcal{C}_\tau = \{\tau : A_i(\tau) > 0 \text{ for } i = 2, \ldots, k\},
$$

with $A_i(\tau)$ as in (23.21). However, Han et al. [430] pointed out that this restriction works only on an average and not for each sample. Moreover, they argued that the assumption of large sample sizes may often be violated in practice.

A Modified Progressive Censoring Scheme and Optimal Step-Stress Test

In order to get rid of the problems described above, Balakrishnan and Han [99] proposed a modified progressive censoring scheme. In contrast to censoring according to a pre-fixed censoring plan $\mathscr{R} = (R_1, \ldots, R_{k-1})$, they introduced a relative censoring scheme which censors a fixed proportion

$$
\boldsymbol{\pi}^* = (\pi_1^*, \cdots, \pi_{k-1}^*)
$$

of the surviving units at the end of each stress level ($0 \le \pi_i^* < 1, i = 1, \ldots, k-1$). This yields a random number of removals at each censoring step of the experiment because the effectively removed number of items depends on the number of surviving specimens at the particular censoring time. Since all the remaining items are withdrawn from the test at the end of stress level s_k, $\pi_k^* = 1$. Therefore, the number of censored items at the end of stress level s_i may be defined as

$$
R_i^* = \Upsilon((N_i - n_i)\pi_i^*) \qquad \text{for } i = 1, \ldots, k-1,
$$

where $\Upsilon(\cdot)$ is a discretizing function such as *round*$(\cdot)$, *floor*$(\cdot)$, *ceiling*$(\cdot)$, and *trunc*$(\cdot)$. This censoring scheme is an adaptive procedure since it takes into account how many failures have been observed before the censoring time. From the condition $0 \le \pi_i^* < 1$, it follows that the number of effectively removed items R_i^* satisfies the inequality $0 \le R_i^* \le N_i - n_i$, $i = 1, \ldots, k-1$. If i^* denotes the minimum of such indices satisfying $R_i^* = N_i - n_i$, the life test terminates at the end of the i^*th stage. Therefore, the life test is allowed to terminate before reaching the last stress level s_k.

As mentioned above, the employed censoring scheme $\mathscr{R}^* = (R_1^*, \ldots, R_{k-1}^*)$ is random. Hence, the proportion of removed units $\boldsymbol{\pi} = \mathscr{R}^*/n = (\pi_1, \ldots, \pi_{k-1})$ is random, too. If progressive censoring is not present in the life test, i.e., if $\boldsymbol{\pi}^* = (0^{*k-1})$, the employed censoring scheme is given by $\mathscr{R}^* = (0^{*k-1})$. Obviously, this scenario corresponds to the case of a k-level step-stress testing under Type-I right censoring. In addition, if $R_k^* > 0$ or $n_k > 0$ (equivalently, $N_k = n_k + R_k^* > 0$), it implies that the life test has proceeded onto the last stress level s_k.

Then, under such a modified progressive censoring scheme, Balakrishnan and Han [99] derived the maximum likelihood estimators of the parameters α and β as well as an explicit expression for the Fisher information matrix. The Fisher information has the same form as (23.20) with, $1 \le i \le k$,

$$A_i(\tau) = F_i(\tau) \prod_{j=1}^{i-1} (1 - F_j(\tau))(1 - \pi_j^*) = F_i(\tau) G_{i-1}(\tau) \prod_{j=1}^{i-1} (1 - \pi_j^*).$$

They used these results to design an optimal step-stress test under this setup by considering the following optimality criteria:

(1) Variance optimality as defined in (23.22);
(2) D-optimality as defined in (23.23);
(3) A-optimality which maximizes the trace of the Fisher information matrix.

With these three criteria, Balakrishnan and Han [99] have presented several numerical results on optimal time duration τ under different settings and have also made some comparisons and empirical observations with regard to the optimal τ under these criteria. In addition, they have also discussed the existence and uniqueness of the optimal time duration τ under all three optimality criteria.

Remark 23.1.9. Finally, we would like to mention that similar considerations have been addressed in the model of progressive censoring with random removals. We refer the interested reader to Ding et al. [341], Shen et al. [803], and Ding and Tse [340].

Step-Stress Test with Link Function Based on Box–Cox Transformation

Assuming exponential lifetimes, the (constant) hazard rate function for the ith stress level is given by $\lambda_i = 1/\vartheta_i$, $i = 1, \ldots, k$. Extending the preceding approach, Fan et al. [360] considered a link function based on a Box–Cox transformation which incorporates the stress in a different way. Given time points $\tau_1 < \cdots < \tau_k$, the stress imposed on the units changes at τ_i. The stress environment in the interval $(\tau_{i-1}, \tau_i]$ is represented by a vector $\mathbf{z}_i = (1, z_{i1}, \ldots, z_{i\ell})'$ which is connected to the hazard rate by

$$\lambda_i^{(\xi)} = \mathbf{z}_i' \boldsymbol{\beta},$$

where the hazard rates satisfy a Box–Cox model of the form

$$\lambda_i^{(\xi)} = \begin{cases} \frac{\lambda_i^{\xi} - 1}{\xi} & \text{if } \xi \neq 0 \\ \log \lambda_i & \text{if } \xi = 0 \end{cases},$$

and $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_\ell)'$ and ξ are unknown parameters. For $\xi = 0$, $\ell = 1$, and $\mathbf{z}_i = (1, s_i)$, this approach reduces to the log-linear link function considered in (23.14). Fan et al. [360] discussed likelihood as well as Bayesian inference for $\boldsymbol{\beta}$ and ξ. Moreover, they presented a comparative study.

Progressively Type-I Interval Censored Exponential Data

Wu et al. [917] considered progressively Type-I censored data when only the number of failures is available for each stress level. Figure 23.6 illustrates the situation based on Fig. 23.5 for arbitrary inspection times $\tau_1 < \cdots < \tau_k$. Assuming equi-spaced inspection times $\tau_j = j\tau$, $j = 0, \ldots, k$, for some $\tau > 0$, a cumulative exposure model as in (23.19) is assumed. Moreover, a log-linear link function $\log \vartheta_i = \alpha + \beta s_i$, $i = 1, \ldots, k$, as in (23.14) relates stress levels to the parameters of the distribution. Then, as pointed out in Wu et al. [917], the likelihood function is proportional to

$$\mathsf{L}(\boldsymbol{\vartheta}) \propto \prod_{j=1}^{k} \left(1 - \mathrm{e}^{-\tau/\vartheta_j}\right)^{D_j} \mathrm{e}^{-(\eta_j - D_j)\tau/\vartheta_j}, \tag{23.24}$$

where $\eta_j = n - D_{\bullet j-1} - R_{\bullet j-1}$, $j = 1, \ldots, k$. The resulting likelihood equations for α and β have to be solved numerically. Wu et al. [917] proposed to use a normal approximation to compute approximate confidence intervals by following the approach of Gouno et al. [408]. They established a representation of the expected Fisher information matrix similar to (23.20). Finally, they discussed the determination of both variance-optimal and D-optimal inspection interval length τ. They applied their results to data given in Table 23.1. For details, we refer to Wu et al. [917].

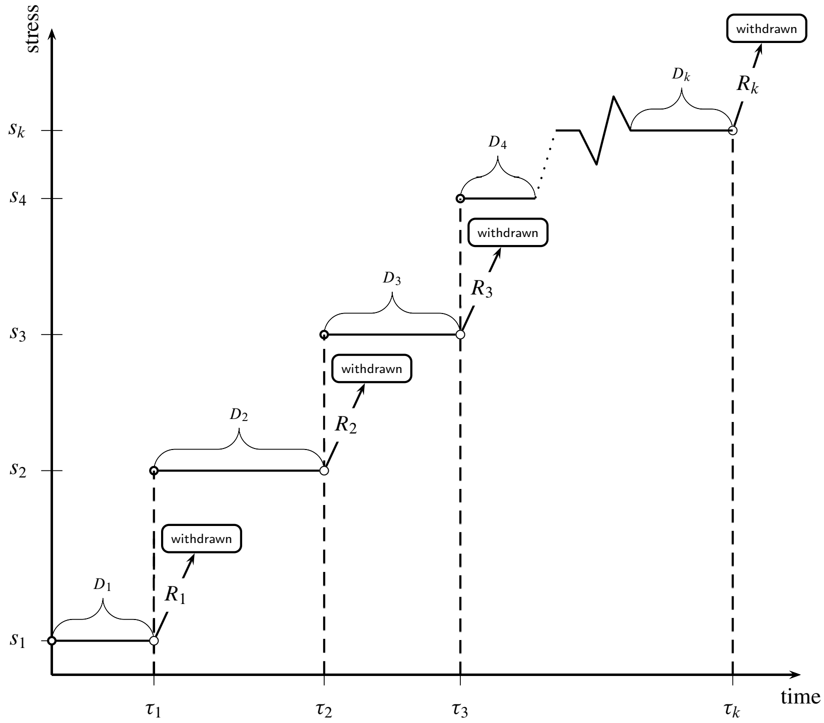

Fig. 23.6 Generation process of progressively interval-censored data in a step-stress scenario

	Stress level s_j				
	5	10	15	20	25
Number of failures D_j	1	1	3	1	1
Number of removals R_j	1	0	1	0	0

Table 23.1 Outcome of progressively Type-I group-censored step-stress experiment as reported in Wu et al. [917]. The data are generated from step-stress data (lifetime of cryogenic cable insulation) reported by Nelson [677, p. 496, Table 2.1]

An extension to arbitrary inspection times $\tau_1, \ldots, \tau_k$ has been worked out by Wu et al. [920] for exponential step-stress data. They discussed likelihood inference for the parameter $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_k)'$ assuming a log-linear link function $\log \vartheta_j = \boldsymbol{\beta}' s$, where $s = (1, s_1, \ldots, s_k)'$ denotes the vector of stress levels. A model involving Weibull distributions has been considered in Yue and Shi [935].

23.1.4 Multiple Step-Stress Model with Progressive Censoring: An Approach Based on Sequential Order Statistics

Following the ideas of Balakrishnan et al. [152], a multiple step-stress model with additional progressive censoring can be introduced. In fact, the resulting model corresponds to that proposed in Wang and Yu [889] and further discussed in Wang [884]. Given that the stress changes at prefixed failures $\rho_1 < \cdots < \rho_k$, we take into account additionally a censoring scheme $\mathscr{R}$. As in (23.17), the cumulative distribution functions in the model of sequential order statistics are given by

$$F_j(t) = 1 - \overline{F}^{\gamma_j/(n-j+1)\vartheta_i}(t), \quad j = \rho_{i-1+1}, \dots, \rho_i, 1 \le i \le k,$$

where $\rho_0 = 0$ and $\gamma_j = \sum_{\ell=j}^{m}(R_\ell + 1)$, $1 \le j \le m = \rho_k$. The likelihood function is given by

$$L(\boldsymbol{\vartheta}) = \prod_{j=1}^{m} \gamma_j \prod_{i=1}^{k} \vartheta_i^{-(\rho_i - \rho_{i-1})} \exp\Big\{-\sum_{i=1}^{k} \frac{D_i}{\vartheta_i}\Big\},$$

where

$$D_i = \sum_{j=1+\rho_{i-1}}^{\rho_i} \gamma_j (X_{(j)} - X_{(j-1)}), \quad i = 1, \dots, k.$$

As before, it follows from properties of sequential order statistics that the spacings of the failure times $X^*_{(1)}, \dots, X^*_{(m)}$ are independent exponential random variables. Hence, the statistics $D_1, \dots, D_k$ are independent random variables with $2D_i/\vartheta_i \sim \chi^2(2(\rho_i - \rho_{i-1}))$, $i = 1, \dots, k$ (see also Wang and Yu [889, Lemma 4]). This shows that the MLE of $\boldsymbol{\vartheta}$ is given by $\widehat{\vartheta}_i = D_i/(\rho_i - \rho_{i-1})$, $i = 1, \dots, k$. Moreover, the properties presented in Theorem 23.1.8 can be easily extended to the present model. Wang and Yu [889] discussed likelihood inference assuming a log-linear link function as given in (23.14). Interval estimation is addressed in Wang [884].

23.2 Progressive Stress Models

Progressive stress accelerated life tests for progressively Type-II censored data have been studied by Abdel-Hamid and AL-Hussaini [4]. They imposed the following conditions:

(1) The lifetimes of n units under test have Weibull(a, b)-distributions;
(2) The data are given by k independent progressively censored samples. The ith sample is drawn from a total of n_i specimens employing a censoring scheme $\mathscr{R}_i = (R_{1,i}, \dots, R_{m_i,i})$ and, thus, forms a sample of size m_i $(n = n_\bullet)$;

(3) The stress function in each sample is given by $s_i(t) = v_i t$, $t \geq 0$, where $0 < v_1 < \cdots < v_k$ are known stress parameters;
(4) The scale parameter of the Weibull distribution depends on the stress function, i.e.,

$$a(t) = \frac{1}{c[s(t)]^d} = \frac{1}{c(v_i t)^d}, \quad t > 0;$$

(5) The linear cumulative exposure model is imposed to model the stress effect.

As pointed out in Abdel-Hamid and AL-Hussaini [4], the cumulative distribution function of the unit lifetime under progressive stress is given by a time transformation, i.e.,

$$G_i(t) = F(\Delta_i(t)),$$

where $\Delta_i(t) = cv_i^d t^{d+1}/(d+1)$, $t \geq 0$, and F denotes the cumulative distribution function of a $\mathsf{Weibull}(1, b)$-distribution. Hence, G_i is the cumulative distribution function of a $\mathsf{Weibull}(\vartheta_i, b(d+1))$-distribution with

$$\vartheta_i = \left(\frac{d+1}{cv_i^d}\right)^{1/(d+1)}. \tag{23.25}$$

Summing up the above assumptions, we have the following data scenario:

(i) The data consists of k independent progressively Type-II censored samples $X_{i,1:m_i:n_i}, \ldots, X_{i,m_i:m_i:n_i}$, $i = 1, \ldots, k$;
(ii) The censoring scheme in the ith sample is given by $\mathscr{R}_i$;
(iii) The population distribution in the ith sample is given by a $\mathsf{Weibull}(\vartheta_i, b(d+1))$-distribution.

Hence, we can interpret the situation as a k-sample problem with independent samples from possibly different Weibull populations. In this sense, the situation is similar to the multi-sample settings discussed in Balakrishnan et al. [130] or Cramer and Kamps [298] in terms of sequential order statistics.

Now, Abdel-Hamid and AL-Hussaini [4] applied a log-transform to the data which leads to similar data from an extreme value distribution with observations $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_k)$, $\mathbf{x}_i = (x_{1,i}, \ldots, x_{m_i,i})$. This leads to the log-likelihood function

$$\ell(b, c, d; \mathbf{x}) = \text{const} + m_\bullet \log[b(d+1)] + \sum_{i=1}^{k} \sum_{j=1}^{m_i} \left[(b(d+1)(x_{j,i} - \log \vartheta_i) - (R_{j,i}+1)\left(\frac{e^{x_{j,i}}}{\vartheta_i}\right)^{b(d+1)} \right],$$

where ϑ_i is given in (23.25). This function has to be maximized w.r.t. b, c, d which has to be done computationally. The likelihood equations can be easily derived (see Abdel-Hamid and AL-Hussaini [4]). Simulation results as well as several computational approaches to construct approximative confidence intervals may also be found in Abdel-Hamid and AL-Hussaini [4]. Moreover, they discussed a graphical procedure to check whether a Weibull model is adequate to the given data. The method proceeds by a Kaplan–Meier-type estimator based on progressively censored data (see (21.6)) (see also Lai et al. [567]).

Similar models have been discussed for Burr Type-XII distributions in Abdel-Hamid [3], for Pareto distributions in Abushal and Soliman [10], and for Rayleigh distributions in Abdel-Hamid and AL-Hussaini [5].

Chapter 24
Stress–Strength Models with Progressively Censored Data

In a reliability context, let a random variable Y describe the strength of a unit subjected to a certain stress represented by the random variable X. The unit fails when the stress X exceeds the strength Y. Thus, the probability

$$R = P(X < Y)$$

may serve as a measure for the reliability of the unit (see, e.g., Tong [849, 850, 851], Beg [184], Constantine et al. [279], and Kotz et al. [547]). Stress–strength models have been widely investigated in the literature. For a detailed account on models, inferential results, and applications, we refer to Kotz et al. [547]. In particular, Chap. 7 of this monograph provides an extensive survey on applications and examples of stress–strength models.

Inference for R has been based on various assumptions regarding the underlying data. For instance, for two independent samples of $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$, Birnbaum [205] showed that the Mann–Whitney statistic

$$U = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \mathbb{1}_{(X_i,\infty)}(Y_j)$$

yields an unbiased nonparametric estimator of R, i.e.,

$$\hat{R} = \frac{1}{n_1 n_2} U.$$

Since this pioneering work, many results have been obtained for various models and assumptions. For an extensive review, we refer to Kotz et al. [547].

In the following, we are interested in the estimation of R based on two independent progressively Type-II censored samples $X^{\mathscr{R}}_{1:m_1:n_1}, \ldots, X^{\mathscr{R}}_{m_1:m_1:n_1}$ and $Y^{\mathscr{S}}_{1:m_2:n_2}, \ldots, Y^{\mathscr{S}}_{m_2:m_2:n_2}$. The sample sizes as well as the censoring schemes may be different.

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_24,

24.1 Exponentially Distributed Stress and Strength

In this section, we assume that the population distribution is exponential, i.e.,

$$X \sim \mathsf{Exp}(\mu, \vartheta_1), \quad Y \sim \mathsf{Exp}(\mu, \vartheta_2)$$

with a common location parameter $\mu \in \mathbb{R}$. In this case, R can be easily calculated as the ratio

$$R = \frac{\vartheta_2}{\vartheta_1 + \vartheta_2}. \tag{24.1}$$

In the inferential part, we distinguish the cases when μ is known or unknown.

24.1.1 Exponentially Distributed Stress and Strength with Known Location Parameter

For convenience, let $\mu = 0$. The following results are taken from Cramer and Kamps [297] and Cramer [283] who considered the problem in the more general scenario of order statistics from independent samples from Weinman multivariate exponential distributions. Some of these results were rediscovered for progressively Type-II censored order statistics by Saraçoğlu et al. [769].

From (12.4), we conclude that the MLEs of ϑ_1 and ϑ_2 are given by

$$\widehat{\vartheta}^*_{1,\mathsf{MLE}} = \frac{1}{m_1}\sum_{j=1}^{m_1}(R_j + 1)X^{\mathscr{R}}_{j:m_1:n_1}, \quad \widehat{\vartheta}^*_{2,\mathsf{MLE}} = \frac{1}{m_2}\sum_{j=1}^{m_2}(S_j + 1)Y^{\mathscr{S}}_{j:m_2:n_2},$$

respectively. As in Theorem 12.1.1, we have the MLEs to be complete sufficient statistics. Both estimators are unbiased and $2m_j\widehat{\vartheta}^*_{j,\mathsf{MLE}}/\vartheta_j \sim \chi^2(2m_j)$ so that $\mathsf{Var}(\widehat{\vartheta}^*_{j,\mathsf{MLE}}) = \frac{\vartheta_j^2}{m_j}$, $j = 1, 2$ (see also Corollary 17.1.1). Now, the MLE $\widehat{R}^*_{\mathsf{MLE}}$ of R is obtained by plugging in the MLEs of the parameters ϑ_j so that

$$\widehat{R}^*_{\mathsf{MLE}} = \frac{\widehat{\vartheta}^*_{2,\mathsf{MLE}}}{\widehat{\vartheta}^*_{1,\mathsf{MLE}} + \widehat{\vartheta}^*_{2,\mathsf{MLE}}}.$$

Using the properties of the MLEs, it is pointed out in Cramer [283] that $\widehat{R}^*_{\mathsf{MLE}}$ has a three-parameter beta distribution $\mathsf{G3B}(m_1, m_2, \lambda)$ with $\lambda = m_2\vartheta_1/(m_1\vartheta_2)$. The density function is given by

$$\psi_{m_1,m_2,\lambda}(t) = \frac{\lambda^{m_2}}{\mathsf{B}(m_1, m_2)}\frac{t^{m_2-1}(1-t)^{m_1-1}}{[1-(1-\lambda)t]^{m_1+m_2}}, \quad t \in (0, 1),$$

(see also Saraçoğlu et al. [769]). Further details about this distribution can be found in Libby and Novick [591], Chen and Novick [253], Pham-Gia and Duong [719], and Johnson et al. [484].

Let $A = (1 - R)/R$ and $\widehat{A} = (1 - \widehat{R}^*_{\mathsf{MLE}})/\widehat{R}^*_{\mathsf{MLE}}$. Utilizing the fact that

$$\frac{\widehat{A}}{A} = \frac{\widehat{\vartheta}_{1,\mathsf{MLE}}/\vartheta_1}{\widehat{\vartheta}_{2,\mathsf{MLE}}/\vartheta_2} \sim \mathsf{F}(m_1, m_2),$$

we have

$$\left[\frac{\mathsf{F}_{\alpha/2}(m_1, m_2)}{\mathsf{F}_{\alpha/2}(m_1, m_2) + \widehat{A}}, \frac{\mathsf{F}_{1-\alpha/2}(m_1, m_2)}{\mathsf{F}_{1-\alpha/2}(m_1, m_2) + \widehat{A}}\right]$$

as a $(1 - \alpha)$ confidence interval for R (see Saraçoğlu et al. [769]).

Cramer and Kamps [297] also obtained the UMVUE of R. The estimator is piecewise defined via a hypergeometric function

$$\widehat{R}^*_{\mathsf{UMVUE}} = \begin{cases} F\Big(1 - m_2, 1; m_1; \frac{m_1\widehat{\vartheta}^*_{1,\mathsf{MLE}}}{m_2\widehat{\vartheta}^*_{2,\mathsf{MLE}}}\Big), & m_1\widehat{\vartheta}^*_{1,\mathsf{MLE}} \le m_2\widehat{\vartheta}^*_{2,\mathsf{MLE}} \\ 1 - F\Big(1 - m_1, 1; m_2; \frac{m_2\widehat{\vartheta}^*_{2,\mathsf{MLE}}}{m_1\widehat{\vartheta}^*_{1,\mathsf{MLE}}}\Big), & m_1\widehat{\vartheta}^*_{1,\mathsf{MLE}} > m_2\widehat{\vartheta}^*_{2,\mathsf{MLE}} \end{cases}.$$

The representation can be expressed as a finite sum using expansions of hypergeometric functions. For instance, in the first case, we get

$$\widehat{R}^*_{\mathsf{UMVUE}} = \sum_{j=0}^{m_2} (-1)^j \frac{\binom{m_2-1}{j}}{\binom{m_1+j-1}{j}} \left(\frac{m_1\widehat{\vartheta}^*_{1,\mathsf{MLE}}}{m_2\widehat{\vartheta}^*_{2,\mathsf{MLE}}}\right)^j.$$

For Type-II censored samples, this result has been obtained by Bartoszewicz [172]. Tong [849, 850, 851] discussed the case of complete samples. Saraçoğlu et al. [769] addressed Bayesian inference for R assuming independent gamma priors. They found a Bayesian estimate in terms of a hypergeometric function and presented an algorithm to construct a credible interval.

24.1.2 Exponentially Distributed Stress and Strength with Common Unknown Location Parameter

In this section, inference for R is addressed for two-parameter exponential distribution with a common unknown location parameter. For complete samples, this problem has been discussed in Bai and Hong [70] (see also Cramer and Kamps [296]). The results for progressively Type-II censored order statistics are taken from Cramer and Kamps [297] and Cramer [283]. Let

$$W_1 = \sum_{j=1}^{m_1} \gamma_j(\mathscr{R})(X^{\mathscr{R}}_{j:m_1:n_1} - X^{\mathscr{R}}_{j-1:m_1:n_1}),$$

$$W_2 = \sum_{j=1}^{m_2} \gamma_j(\mathscr{S})(Y^{\mathscr{S}}_{j:m_2:n_2} - Y^{\mathscr{S}}_{j-1:m_2:n_2}),$$

where $X^{\mathscr{R}}_{0:m_1:n_1} = Y^{\mathscr{S}}_{0:m_2:n_2} = 0$. Then, it is shown in Cramer [283] that the MLEs of $\mu, \vartheta_1, \vartheta_2$ are given by

$$\widehat{\mu}_{\mathsf{MLE}} = \min(X^{\mathscr{R}}_{1:m_1:n_1}, Y^{\mathscr{S}}_{1:m_2:n_2}),$$
$$\widehat{\vartheta}_{1,\mathsf{MLE}} = \frac{1}{m_1}(W_1 - n_1\widehat{\mu}_{\mathsf{MLE}}), \quad \widehat{\vartheta}_{2,\mathsf{MLE}} = \frac{1}{m_2}(W_2 - n_2\widehat{\mu}_{\mathsf{MLE}}).$$

Remark 24.1.1. The MLE $\widehat{\vartheta}_{1,\mathsf{MLE}}$ (and, by analogy, $\widehat{\vartheta}_{2,\mathsf{MLE}}$) can be written as

$$\widehat{\vartheta}_{1,\mathsf{MLE}} = \frac{1}{m_1}\sum_{j=1}^{m_1} \gamma_j(\mathscr{R})(X^{\mathscr{R}}_{j:m_1:n_1} - X^{\mathscr{R}}_{j-1:m_1:n_1})$$
$$= \frac{1}{m_1}\sum_{j=1}^{m_1}(R_j + 1)(X^{\mathscr{R}}_{j:m_1:n_1} - \widehat{\mu}_{\mathsf{MLE}}),$$

where $X^{\mathscr{R}}_{0:m_1:n_1} = \min(X^{\mathscr{R}}_{1:m_1:n_1}, Y^{\mathscr{S}}_{1:m_2:n_2})$, say. Using the above expression, Cramer [283] obtained that $\widehat{\mu}_{\mathsf{MLE}}$ and $(\widehat{\vartheta}_{1,\mathsf{MLE}}, \widehat{\vartheta}_{2,\mathsf{MLE}})$ are independent. $\widehat{\mu}_{\mathsf{MLE}}$ is $\mathsf{Exp}\Big(\mu, \big(n_1/\vartheta_1 + n_2/\vartheta_2\big)^{-1}\Big)$-distributed. The distributions of $2m_j\widehat{\vartheta}_{j,\mathsf{MLE}}$, $j = 1, 2$, are binomial mixtures of χ^2-distributions. In particular, with $\eta = \frac{n_1\vartheta_2}{n_2\vartheta_1 + n_1\vartheta_2}$, the distributions can be written as

$$\begin{aligned} 2m_1\widehat{\vartheta}_{1,\mathsf{MLE}} &\sim \eta\chi^2(2m_1 - 2) + (1-\eta)\chi^2(2m_1), \\ 2m_2\widehat{\vartheta}_{2,\mathsf{MLE}} &\sim (1-\eta)\chi^2(2m_2 - 2) + \eta\chi^2(2m_2). \end{aligned} \tag{24.2}$$

Moreover, $(\widehat{\mu}_{\mathsf{MLE}}, W_1, W_2)$ is a complete sufficient statistic for $(\mu, \vartheta_1, \vartheta_2)$. These results can be used to establish the UMVUE of μ. Introducing the estimators

$$\widehat{\vartheta}_{j,\star} = \frac{m_j}{m_j - 1}\widehat{\vartheta}_{j,\mathsf{MLE}}, \quad j = 1, 2,$$

it is given by

$$\widehat{\mu}_{\mathsf{UMVUE}} = \widehat{\mu}_{\mathsf{MLE}} - \frac{\widehat{\vartheta}_{1,\star}\widehat{\vartheta}_{2,\star}}{n_2\widehat{\vartheta}_{1,\star} + n_1\widehat{\vartheta}_{2,\star}}$$

(see Cramer [283]; for complete samples, see Chiou and Cohen [262] and Ghosh and Razmpour [401]).

Furthermore, the MLE of R is given by

$$\widehat{R}_{\mathsf{MLE}} = \frac{\widehat{\vartheta}_{2,\mathsf{MLE}}}{\widehat{\vartheta}_{1,\mathsf{MLE}} + \widehat{\vartheta}_{2,\mathsf{MLE}}}.$$

The distribution of $\widehat{R}_{\mathsf{MLE}}$ is a binomial mixture of generalized beta distributions. According to Cramer [283],

$$\widehat{R}_{\mathsf{MLE}} \sim \eta \mathsf{G3B}(m_1 - 1, m_2, \lambda) + (1-\eta)\mathsf{G3B}(m_1, m_2 - 1, \lambda), \tag{24.3}$$

where $\eta = \frac{n_1\vartheta_2}{n_2\vartheta_1 + n_1\vartheta_2}$ as in (24.2) and $\lambda = (m_2\vartheta_1)/(m_1\vartheta_2)$. This result can be used to obtain explicit expressions for the moments of $\widehat{R}_{\mathsf{MLE}}$. Following an idea of Bhattacharyya and Johnson [199], this connection can be used to establish an approximate confidence interval for R. Details are presented in Cramer [283].

As in the case of known location parameter, the UMVUE of R can also be constructed. Cramer and Kamps [297] established the following expression. For $(m_1 - 1)\widehat{\vartheta}_{1,\star} \le (m_2 - 1)\widehat{\vartheta}_{2,\star}$,

$$\widehat{R}_{\mathsf{UMVUE}} = \frac{n_1\widehat{\vartheta}_{2,\star} + \widehat{\vartheta}_{1,\star}[n_2 - n_1]F\left(2 - m_2, 1; m_1; \frac{(m_1-1)\widehat{\vartheta}_{1,\star}}{(m_2-1)\widehat{\vartheta}_{2,\star}}\right)}{n_1\widehat{\vartheta}_{2,\star} + n_2\widehat{\vartheta}_{1,\star}},$$

while for $(m_1 - 1)\widehat{\vartheta}_{1,\star} > (m_2 - 1)\widehat{\vartheta}_{2,\star}$,

$$\widehat{R}_{\mathsf{UMVUE}} = 1 - \frac{n_2\widehat{\vartheta}_{1,\star} + \widehat{\vartheta}_{2,\star}[n_1 - n_2]F\left(2 - m_1, 1; m_2; \frac{(m_2-1)\widehat{\vartheta}_{2,\star}}{(m_1-1)\widehat{\vartheta}_{1,\star}}\right)}{n_2\widehat{\vartheta}_{1,\star} + n_1\widehat{\vartheta}_{2,\star}}.$$

Notice that, for equal sample sizes $n_1 = n_2$, the representations simplify considerably:

$$\widehat{R}_{\mathsf{UMVUE}} = \frac{\widehat{\vartheta}_{2,\star}}{\widehat{\vartheta}_{1,\star} + \widehat{\vartheta}_{2,\star}}. \tag{24.4}$$

In this case, Cramer [283] has shown that the distribution of $\widehat{R}_{\mathsf{UMVUE}}$ is a binomial mixture similar to (24.3) as is for the MLE $\widehat{R}_{\mathsf{MLE}}$. However, as pointed out by Cramer and Kamps [297], the UMVUE in (24.4) has a special Gauß hypergeometric distribution (see Armero and Bayarri [48]) with density function

$$f^R(t) = \binom{m_1+m_2-2}{m_1-1} \frac{(m_2-1)^{m_2}}{(m_1-1)^{m_2-1}} \frac{\lambda^{m_2}}{\lambda+1} \times \frac{(1-t)^{m_1-2} t^{m_2-2}}{\left(1+\left(\frac{m_2-1}{m_1-1}\lambda-1\right)t\right)^{m_1+m_2-1}}, \quad t \in (0,1),$$

where $\lambda = \vartheta_1/\vartheta_2, m_1, m_2 \geq 2$.

The multi-sample case is considered in Cramer [283], i.e., s_1 and s_2 independent progressively censored samples are considered. As in the two-sample case, expressions for MLE and UMVUE are established. Moreover, it is shown that, under some regularity conditions, MLE and UMVUE are asymptotically equivalent. Moreover, they are asymptotically normal. This result may be used to construct approximate confidence intervals for R.

24.2 Further Stress–Strength Distributions

Asgharzadeh et al. [64] have discussed stress–strength models with progressively censored data from Weibull lifetime distributions with a common shape parameter, i.e.,

$$X \sim \mathsf{Weibull}(\vartheta_1, a), \quad Y \sim \mathsf{Weibull}(\vartheta_2, a), \quad \vartheta_1, \vartheta_2, a > 0.$$

In this case, the stress–strength reliability R is independent of the shape parameter and given by the expression (24.1), too. Thus, if a is known, the results for the exponential model can be directly applied. If a is unknown, the MLE of R can be obtained from the MLEs

$$\widehat{\vartheta}_{1,\mathsf{MLE}} = \frac{1}{m_1}\sum_{j=1}^{m_1}(R_j+1)(X_{j:m_1:n_1}^{\mathscr{R}})^{\widehat{\alpha}}, \quad \widehat{\vartheta}_{2,\mathsf{MLE}} = \frac{1}{m_2}\sum_{j=1}^{m_2}(S_j+1)(Y_{j:m_2:n_2}^{\mathscr{S}})^{\widehat{\alpha}},$$

where $\widehat{\alpha}$ is the solution of the equation $(m_1+m_2)/g(\alpha) = \alpha$ and

$$g(\alpha) = \frac{m_1\sum_{j=1}^{m_1}(R_j+1)x_j^{\alpha}\log x_j}{\sum_{j=1}^{m_1}(R_j+1)x_j^{\alpha}} + \frac{m_2\sum_{j=1}^{m_2}(S_j+1)y_j^{\alpha}\log y_j}{\sum_{j=1}^{m_2}(S_j+1)y_j^{\alpha}} - \sum_{j=1}^{m_1}\log x_j - \sum_{j=1}^{m_2}\log y_j.$$

Asgharzadeh et al. [64] proposed a simple fix-point-based procedure to get an estimate for α. Moreover, they adopted an approach of Kundu and Gupta [559] and approximated the likelihood equation by transforming the data to a location–scale model from the extreme value distribution. Using the approximated MLEs of ϑ_1

i	1	2	3	4	5	6	7	8	9	10
$x_{i:10:69}$	1.312	1.479	1.552	1.803	1.944	1.858	1.966	2.027	2.055	2.098
R_i	1	0	1	2	0	0	3	0	1	50
$y_{i:10:63}$	1.901	2.132	2.257	2.361	2.396	2.445	2.373	2.525	2.532	2.575
S_i	0	2	1	0	1	1	2	0	0	44

Table 24.1 Progressively censored stress–strength data taken from Asgharzadeh et al. [64, Table 7]

and ϑ_2, this leads to a different estimator for R. Furthermore, they proposed interval estimates by using asymptotic normality of an estimator of R as well as by considering two bootstrap approaches. Finally, Bayesian estimation of R is discussed using inverse gamma priors for ϑ_j and a gamma prior for α.

Example 24.2.1. Asgharzadeh et al. [64] applied their results to progressively Type-II censored data generated from data presented in Kundu and Gupta [559]. The data sets are given in Table 24.1. They obtained the estimates $0.176, 0.179$, and 0.328 for the MLE, the approximate MLE, and the Bayesian estimate of R.

Remark 24.2.2. Valiollahi et al. [865] addressed a Weibull model with common scale but different shape parameters, i.e., $X \sim \text{Weibull}(\vartheta, a_1), Y \sim \text{Weibull}(\vartheta, a_2)$. In this case, the stress–strength probability exhibits an integral representation

$$P(Y < X) = 1 - \frac{a_1}{\vartheta} \int_0^\infty t^{a_1-1} \mathrm{e}^{-(t^{a_1}+t^{a_2})/\vartheta} dt.$$

Likelihood inference, approximate estimation, and Bayesian inference are considered. The results are applied to the data given in Table 24.1.

Lio and Tsai [613] discussed estimation of R when the data are sampled from a Burr-XII distribution with density functions

$$f_i(t) = \beta_i \alpha t^{\alpha-1}(1+t^\alpha)^{-\beta_i-1}, \quad t > 0, \alpha, \beta_i > 0, i = 1, 2,$$

in the presence of progressive first-failure censoring. Then, R is independent of the common shape parameter α and given as the ratio $\beta_1/(\beta_1 + \beta_2)$. They proved that, under weak conditions, the MLEs of the distribution parameters uniquely exist. Plugging this result into R yields the MLE of R. The authors also established confidence intervals based on Fisher information and bootstrap procedures.

Basirat et al. [180] considered a proportional hazards model with different proportionality parameters for the population distributions. It includes exponential, Weibull (with known shape parameter), Pareto, and Burr-XII distributions as special cases. The stress–strength probability has the form given in (24.1). Then, the corresponding results are derived along the lines for the exponential distribution.

Chapter 25
Multi-sample Models

25.1 Competing Risk Models

25.1.1 Model and Notation

In competing risk modeling, it is assumed that a unit may fail due to several causes of failure. For two competing risks, the lifetime of the ith unit is given by

$$X_i = \min\{X_{1i}, X_{2i}\}, \quad i = 1, \ldots, n,$$

where X_{ji} denotes the latent failure time of the ith unit under the jth cause of failure, $j = 1, 2$. We assume that the latent failure times are independent, where $X_{ji} \sim F_j$, $j = 1, 2$, $i = 1, \ldots, n$. This ensures that the distributions of the latent failure times entirely determine the competing risk model (see, e.g., Crowder [316], Tsiatis [862], and Pintilie [721]). Additionally, the sample $X_1, \ldots, X_n$ is progressively Type-II censored and it is assumed that the cause of each failure is known. Therefore, the available data are given by

$$(X_{1:m:n}, C_1), (X_{2:m:n}, C_2), \ldots, (X_{m:m:n}, C_m),$$

where $C_i = 1$ if the ith failure is due to first cause and $C_i = 2$ otherwise. The observed data is denoted by $(x_1, c_1), (x_2, c_2), \ldots, (x_m, c_m)$. Further, we define the indicators

$$\mathbb{1}_{\{j\}}(C_i) = \begin{cases} 1, & C_i = j \\ 0, & \text{else} \end{cases}.$$

Thus, the random variables $m_1 = \sum_{i=1}^m \mathbb{1}_{\{1\}}(C_i)$ and $m_2 = \sum_{i=1}^m \mathbb{1}_{\{2\}}(C_i)$ describe the number of failures due to the first and the second cause of failure, respectively.

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_25,

Observe that both m_1 and m_2 are binomials with sample size m and probability of success $P(X_{11} \le X_{21})$ and $1 - P(X_{11} \le X_{21})$, respectively. For a given censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$, the joint density function (w.r.t. the measure $\lambda^m \otimes \#^m$) is given by (see Kundu et al. [564])

$$f^{\mathbf{X}^{\mathscr{R}},\mathbf{C}}(\mathbf{x}_m, \mathbf{c}_m) = \Big(\prod_{j=1}^m \gamma_j\Big) \prod_{i=1}^m \Big[\big[f_1(x_i)\overline{F}_2(x_i)\big]^{\mathbb{1}_{\{1\}}(c_i)} \big[f_2(x_i)\overline{F}_1(x_i)\big]^{\mathbb{1}_{\{2\}}(c_i)} \times \big[\overline{F}_1(x_i)\overline{F}_2(x_i)\big]^{R_i}\Big].$$

Denoting by λ_j the hazard rate of F_j, $j = 1, 2$, the density function can be written as

$$f^{\mathbf{X}^{\mathscr{R}},\mathbf{C}}(\mathbf{x}_m, \mathbf{c}_m) = \Big(\prod_{j=1}^m \gamma_j\Big) \prod_{i=1}^m [\lambda_1(x_i)]^{\mathbb{1}_{\{1\}}(c_i)} [\lambda_2(x_i)]^{\mathbb{1}_{\{2\}}(c_i)} \big[\overline{F}_1(x_i)\overline{F}_2(x_i)\big]^{R_i+1}, \tag{25.1}$$

for $x_1 < \cdots < x_m$.

25.1.2 Exponential Distribution

Kundu et al. [564] have discussed competing risks for $\mathsf{Exp}(\vartheta_j)$-distributions, $j = 1, 2$. In this case, the joint density function is given by

$$f^{\mathbf{Z}^{\mathscr{R}},\mathbf{C}}(\mathbf{z}_m, \mathbf{c}_m) = \Big(\prod_{j=1}^m \gamma_j\Big) \vartheta_1^{-m_1} \vartheta_2^{-m_2} \exp\Big\{-\Big(\frac{1}{\vartheta_1} + \frac{1}{\vartheta_2}\Big) \sum_{i=1}^m (R_i + 1) z_i\Big\}. \tag{25.2}$$

Denoting by $\mathbf{Z}_*^{\mathscr{R}}$ progressively Type-II censored order statistics from an $\mathsf{Exp}\big(\frac{1}{\vartheta_1} + \frac{1}{\vartheta_2}\big)^{-1}$-distribution and noticing that $m_2 = m - m_1$, the joint density function factorizes as

$$\begin{aligned} f^{\mathbf{Z}^{\mathscr{R}},\mathbf{C}}(\mathbf{z}_m, \mathbf{c}_m) &= f^{\mathbf{Z}_*^{\mathscr{R}}}(\mathbf{z}_m) \times \Big(\frac{\vartheta_1}{\vartheta_1 + \vartheta_2}\Big)^{m-m_1} \Big(1 - \frac{\vartheta_1}{\vartheta_1 + \vartheta_2}\Big)^{m_1} \\ &= f^{\mathbf{Z}_*^{\mathscr{R}}}(\mathbf{z}_m) \times P(\mathbf{C} = \mathbf{c}). \end{aligned}$$

This shows that $\mathbf{Z}^{\mathscr{R}}$ and $\mathbf{C}$ are independent random variables. Therefore, m_1 and the total time on test statistic $\mathsf{TTT} = \sum_{i=1}^m (R_i + 1) Z_{i:m:n}$ are independent, too. Additionally,

$$2\big(\tfrac{1}{\vartheta_1}+\tfrac{1}{\vartheta_2}\big)\mathsf{TTT}\sim\chi^2(2m)\quad\text{and}\quad m_1\sim\mathsf{bin}\big(m,R\big),\tag{25.3}$$

where $R=\frac{\vartheta_2}{\vartheta_1+\vartheta_2}$ denotes the stress–strength probability (see (24.1)). Moreover, representation (25.2) illustrates that (m_1,TTT) is a complete sufficient statistic.

As pointed out by Kundu et al. [564], the MLEs of ϑ_1 and ϑ_2 exist if $0<m_1<m$. In this case, the MLEs are given by

$$\widehat{\vartheta}_j=\frac{1}{m_j}\sum_{i=1}^{m}(R_i+1)Z_{i:m:n}=\frac{1}{m_j}\mathsf{TTT},\quad j=1,2.\tag{25.4}$$

Then, the (conditional) distribution of $\widehat{\vartheta}_j$, given $m_j>0$, has an explicit expression. For instance, as shown by Kundu et al. [564], the cumulative distribution function is given by

$$F^{\widehat{\vartheta}_1|m_1>0}(t)=\frac{1}{1-(1-R)^m}\sum_{i=1}^{m}\binom{m}{i}R^i(1-R)^{m-i}P\Big(\tfrac{1}{i}\mathsf{TTT}\le t\Big).$$

Using the distribution given in (25.3), the cumulative distribution function can be written as

$$F^{\widehat{\vartheta}_1|m_1>0}(t)=\sum_{i=1}^{m}p_i F_{\chi^2(2m)}\Big(2i\big(\tfrac{1}{\vartheta_1}+\tfrac{1}{\vartheta_2}\big)t\Big)$$

with p_i defined appropriately showing that the distribution is a mixture of scaled χ^2-distributions. This result yields directly the first two conditional moments of $\widehat{\vartheta}_1$ as

$$E(\widehat{\vartheta}_1|m_1>0)=m\big(\tfrac{1}{\vartheta_1}+\tfrac{1}{\vartheta_2}\big)^{-1}\sum_{i=1}^{m}\frac{p_i}{i},$$

$$E(\widehat{\vartheta}_1^2|m_1>0)=m(m+1)\big(\tfrac{1}{\vartheta_1}+\tfrac{1}{\vartheta_2}\big)^{-2}\sum_{i=1}^{m}\frac{p_i}{i^2}.$$

Similar expressions hold for $\widehat{\vartheta}_2$, given that $m_2>0$ (or, equivalently, $m_1<m$). These results show that the MLEs are clearly biased.

In order to avoid this problem, Kundu et al. [564] considered a different parametrization of the exponential distributions, i.e., $\lambda_j=1/\vartheta_j$, $j=1,2$. Then, $R=\frac{\lambda_1}{\lambda_1+\lambda_2}$ and the MLEs are given by

$$\widehat{\lambda}_j=\frac{m_j}{\sum_{i=1}^{m}(R_i+1)Z_{i:m:n}}=\frac{m_j}{\mathsf{TTT}},\quad j=1,2.$$

These estimators always exist assuming that $\widehat{\lambda}_j=0$ is a proper value of the parameter leading to a degenerate exponential distribution. Then, using the distribution properties of m_1 and TTT, we see that $1/\big(2(\lambda_1+\lambda_2)\mathsf{TTT}\big)$ has an inverse χ^2-distribution with $2m$ degrees of freedom. For $m>1$ and $m>2$, the mean and variance are given by $1/(2m-2)$ and $1/[(2m-2)^2(m-2)]$, respectively. This yields

$$E\widehat{\lambda}_1 = mR\cdot 2(\lambda_1+\lambda_2)\cdot\frac{1}{2m-2} = \frac{m}{m-1}\lambda_1.$$

Therefore, $\widehat{\lambda}_1^* = \frac{m-1}{m}\widehat{\lambda}_1$ is an unbiased estimator of λ_1. Due to the sufficiency of (m_1, TTT), it is the UMVUE of λ_1. The same arguments lead to the UMVUE of λ_2. According to Kundu et al. [564] the variance–covariance matrix of the UMVUEs is given by

$$\mathsf{Cov}\begin{pmatrix}\widehat{\lambda}_1^*\\ \widehat{\lambda}_2^*\end{pmatrix} = \frac{\lambda_1\lambda_2}{m(m-2)}\begin{pmatrix} m-1+m\lambda_1/\lambda_2 & -1\\ -1 & m-1+m\lambda_2/\lambda_1\end{pmatrix}.$$

Again, it is possible to calculate an explicit expression for the cumulative distribution function of the MLEs. The distributions of the MLEs are binomial mixtures of scaled inverse χ^2-distributions. For instance, the cumulative distribution function of $\widehat{\lambda}_1$ can be written as

$$\begin{aligned} F^{\widehat{\lambda}_1}(t) &= \sum_{i=0}^{m} P(m_1=i)P(\mathsf{TTT}>i/t)\\ &= \sum_{i=0}^{m}\binom{m}{i}R^i(1-R)^{m-i}\overline{F}_{\chi^2(2m)}(2i(\lambda_1+\lambda_2)/t)\\ &= 1-\sum_{i=0}^{m}\binom{m}{i}R^i(1-R)^{m-i}F_{\chi^2(2m)}(2i(\lambda_1+\lambda_2)/t). \end{aligned} \tag{25.5}$$

In order to construct a confidence interval for λ_j, Kundu et al. [564] assumed that, for any λ_2, $\overline{F}^{\widehat{\lambda}_1}$ is a strictly increasing function of λ_1. However, this can be seen directly from representation (25.5) using the monotonicity of $F_{\chi^2(2m)}$, and that

$$R^i(1-R)^{m-i}$$

is increasing in λ_j, $j=1,2$. Then, for fixed λ_2 and $\alpha\in(0,1)$, it follows that a strictly increasing function c_{α,λ_2} exists with $F_{\widehat{\lambda}_1}(c_{\alpha,\lambda_2}(\lambda_1))=\alpha$. Due to the complex structure, the solution has to be computed numerically. Hence, $(0, c_{\alpha,\lambda_2}^{-1}(\widehat{\lambda}_1))$ provides a $100(1-\alpha)\%$ confidence interval for λ_1. However, the upper limit involves the (unknown parameter) λ_2. Therefore, Kundu et al. [564] proposed to replace λ_2 by its MLE or UMVUE. This yields, for example, the approximate two-sided confidence interval

$$\Big(c_{1-\alpha/2,\widehat{\lambda}_2}(\widehat{\lambda}_1)^{-1}), c_{\alpha/2,\widehat{\lambda}_2}(\widehat{\lambda}_1)^{-1}\Big).$$

Alternatively, approximate confidence intervals can be constructed using the Fisher information matrix. For instance,

$$\mathscr{I}(\mathbf{Z}^{\mathscr{R}}, \mathbf{C}; \lambda_1, \lambda_2) = \frac{m}{\lambda_1 + \lambda_2} \begin{pmatrix} 1/\lambda_1 & 0 \\ 0 & m/\lambda_2 \end{pmatrix}$$

and by using a normal approximation. Moreover, Kundu et al. [564] demonstrated that a bootstrap approach can also be utilized to establish confidence intervals.

Kundu et al. [564] also proposed Bayes estimates of λ_j under squared-error loss assuming that λ_1 and λ_2 are independent with gamma priors. Moreover, they showed that the resulting estimates can be used to construct credible intervals. For more details as well as a simulation study, one may refer to their work.

Example 25.1.1. Kundu et al. [564] generated a progressively censored mortality data from a data set originally reported by Hoel [443]. The original data were obtained from a laboratory experiment in which male mice received a radiation dose of 300 roentgens. The cause of death for each mouse was determined by autopsy. Restricting the analysis to two causes of death (reticulum cell sarcoma $\widehat{=}$ cause 1; other causes $\widehat{=}$ cause 2), $n = 77$ observations remain in the analysis. The censoring scheme was given by $m = 25$ and $\mathscr{R} = (2^{*24}, 4)$ leading to the progressively Type-II censored competing risk sample

$$\begin{aligned}&(40,2),(42,2),(62,2),(163,2),(179,2),(206,2),(222,2),(228,2),(252,2),\\&(259,2),(318,1),(385,2),(407,2),(420,2),(462,2),(517,2),(517,2),\\&(524,2),(525,1),(558,1),(536,1),(605,1),(612,1),(620,2),(621,1).\end{aligned}$$

Assuming exponential lifetimes, the estimates for ϑ_1 and ϑ_2 are $\widehat{\vartheta}_1 = 29082/7 \approx 4154.6$ and $\widehat{\vartheta}_2 = 29082/18 \approx 1615.7$. For λ_1 and λ_2, the estimates $\widehat{\lambda}_1 = 0.000241$ and $\widehat{\lambda}_2 = 0.000619$ result. Notice that the values presented in Kundu et al. [564] are in error.

Finally, the MLE of the relative risk R is given by

$$\widehat{R} = \frac{\widehat{\lambda}_1}{\widehat{\lambda}_1 + \widehat{\lambda}_2} = \frac{m_1}{m}.$$

Obviously, $E\widehat{R} = R$ so that $\widehat{R}$ is unbiased. From the sufficiency of m_1 we deduce that $\widehat{R}$ is the UMVUE of R, too.

Remark 25.1.2. Obviously, the preceding results can be directly applied to Weibull distribution with known shape parameter. Given a Rayleigh distribution, the MLE of λ_j is given by

$$\widehat{\lambda}_j = \frac{m_j}{\sum_{i=1}^m (R_i+1)X_{i:m:n}^2}, \quad j = 1, 2.$$

25.1.3 Weibull Distributions

Competing risks with progressively Type-II censored data from Weibull$(1/\lambda_j, \beta)$-distributions have been addressed by Pareek et al. [703]. Considering the joint density function in (25.1), it follows that the log-likelihood function is given by

$$\ell(\lambda_1, \lambda_2, \beta) = \sum_{j=1}^m \log \gamma_j + m \log \beta + m_1 \log \lambda_1 + m_2 \log \lambda_2 \\ + (\beta - 1)\sum_{j=1}^m \log x_j - (\lambda_1 + \lambda_2)\sum_{j=1}^m (R_j+1)x_j^\beta.$$

Defining $\widehat{\lambda}_j(\beta) = m_j / \sum_{j=1}^m x_j^\beta$, $j = 1, 2$, and proceeding as in the proof of Theorem 12.2.1, the log-likelihood is bounded from above by

$$\ell(\widehat{\lambda}_1(\beta), \widehat{\lambda}_2(\beta), \beta) = \sum_{j=1}^m \log \gamma_j + m \log \beta \\ + (\beta - 1)\sum_{j=1}^m \log x_j - m \log \sum_{j=1}^m (R_j+1)x_j^\beta.$$

This bound coincides with that one for progressively Type-II censored order statistics from a Weibull distribution as given in the proof of Theorem 12.2.1. Therefore, the MLE $\widehat{\beta}$ of β is obtained as the unique solution of equation (12.13). Hence, $\widehat{\lambda}_j(\widehat{\beta})$, $j = 1, 2$, are the MLEs of the scale parameters.

Remark 25.1.3. Notice that the density function (25.1) can be written in the form

$$f^{\mathbf{X}^{\mathscr{R}},\mathbf{C}}(\mathbf{x}_m, \mathbf{c}_m) = f^{\mathbf{X}_*^{\mathscr{R}}}(\mathbf{x}_m) \times P(\mathbf{C} = \mathbf{c}),$$

where $\mathbf{X}_*^{\mathscr{R}}$ are progressively Type-II censored order statistics from a Weibull distribution with scale parameter $\lambda_1 + \lambda_2$. This is due to the fact that the hazard rate of a Weibull distribution is given by $\lambda_j(t) = \lambda_j \beta t^{\beta-1}$. Therefore, the part dependent on the data is either included for $c = 1$ or $c = 2$. This illustrates that the quantities m_1 and m_2 are independent of $\mathbf{X}^{\mathscr{R}}$. Hence, the MLEs and (m_1, m_2) are independent. Moreover, $(m_1, m_2, \mathbf{X}^{\mathscr{R}})$ forms a sufficient statistic.

Using a log-transformation, Pareek et al. [703] transformed the model to a location–scale model from an extreme value distribution with the location parameters being different but the scale parameter being the same. Then, they derived approximate MLEs by proceeding as in Sect. 12.9.2.

Example 25.1.4. Pareek et al. [703] analyzed the data presented in Example 25.1.1 assuming Weibull lifetimes. Using an iterative solution, they presented the estimates $\widehat{\beta} = 1.9246$, $\widehat{\lambda}_1 = 8.1102 \cdot 10^{-7}$, and $\widehat{\lambda}_2 = 2.0855 \cdot 10^{-6}$. The AMLEs yield the estimates $\widehat{\beta} = 1.9338$, $\widehat{\lambda}_1 = 7.6603 \cdot 10^{-7}$, and $\widehat{\lambda}_2 = 1.9698 \cdot 10^{-6}$. Using the observed Fisher information, they also provided approximate confidence intervals.

The above estimates of β suggest Rayleigh lifetimes. Assuming a Rayleigh distribution, the estimates of λ_j as given in Remark 25.1.2 lead to the values $\widehat{\lambda}_1 = 5.07678 \times 10^{-7}$ and $\widehat{\lambda}_2 = 1.30546 \times 10^{-6}$.

The expected Fisher information can also be obtained in this scenario. This proceeds analogous to the one-sample case. In particular, the log-likelihood function can be written as

$$\ell(\lambda_1, \lambda_2, \beta) = \text{const} + \ell_m(\lambda_1, \beta) + \ell_m(\lambda_2, \beta) - m_1 \log \lambda_2 - m_2 \log \lambda_1,$$

where $\ell_m(\cdot, \cdot)$ denotes the log-likelihood in the two-parameter setting with sample size m (cf. (12.14) with $\lambda = 1/\vartheta$). Therefore, the Fisher information matrix in the competing risk scenario has the following entries:

$$\begin{aligned}
\mathscr{I}_{11}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) &= \frac{m}{\lambda_1^2} \cdot E m_1 = \frac{m}{\lambda_1(\lambda_1 + \lambda_2)}, \\
\mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) &= \frac{m}{\lambda_2^2} \cdot E m_2 = \frac{m}{\lambda_2(\lambda_1 + \lambda_2)}, \\
\mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) &= \mathscr{I}_{21}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) = 0, \\
\mathscr{I}_{33}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) &= \mathscr{I}_{22}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \beta) + \mathscr{I}_{33}(\mathbf{X}^{\mathscr{R}}; \lambda_2, \beta), \\
\mathscr{I}_{13}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) &= \mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \beta), \\
\mathscr{I}_{23}(\mathbf{X}^{\mathscr{R}}; \lambda_1, \lambda_2, \beta) &= \mathscr{I}_{12}(\mathbf{X}^{\mathscr{R}}; \lambda_2, \beta).
\end{aligned}$$

The expressions for the Fisher information with two parameters are the Fisher information in a progressively censored sample from a Weibull distribution with these parameters and sample size m.

Remark 25.1.5. Bayesian inference for competing risks from Weibull distributions has been addressed by Kundu and Pradhan [563]. They assumed that the shape parameter has a log-concave prior density function and that for given shape parameter, the scale parameters have Beta–Dirichlet priors.

25.1.4 Lomax Distribution

Cramer and Schmiedt [305] discussed the competing risk model for scaled Lomax lifetimes as an alternative to exponential and Weibull distributions (see also Bryson [223]) with cumulative distribution functions and hazard rates

$$F_j(t) = 1 - (1 + \vartheta t)^{-1/(\vartheta\beta_j)}\,, \quad h_j(t) = \frac{1}{\beta_j}\,(1 + \vartheta x)^{-1}\,, \quad t > 0, j = 1, 2\,.$$

For $\vartheta \to 0$, exponential distributions $\mathsf{Exp}(\beta_j)$ result in the limit.

Using the independence of the latent failure times X_{1i} and X_{2i}, $i = 1, \ldots, n$, the stress–strength probability is given by

$$R = P(X_{1i} \le X_{2i}) = \frac{\beta_2}{\beta_1 + \beta_2}.$$

Applying the identity $f_j = h_j \cdot (1 - F_j)$ for $j = 1, 2$, this yields the likelihood

$$\mathsf{L}\,(\vartheta, \beta_1, \beta_2) = \prod_{j=1}^{m} \gamma_j \beta_1^{-m_1}\,\beta_2^{-m_2}\,\prod_{i=1}^{m}(1 + \vartheta x_i)^{-[(R_i+1)(1/\beta_1+1/\beta_2)/\vartheta+1]}\,,$$

where $x_1 < \cdots < x_m$. Using the same argument as in Remark 25.1.3, (m_1, m_2) and $\mathbf{X}^{\mathscr{R}}$ are independent. Now, Cramer and Schmiedt [305] showed that for given $\vartheta > 0$ and for $m_1, m_2 > 0$, the MLEs of β_1 and β_2 exist and can be written as

$$\widehat{\beta}_j = \widehat{\beta}_j(\vartheta) = \frac{1}{m_j\vartheta}\sum_{i=1}^{m}(R_i + 1)\log(1 + \vartheta\, X_{i:m:n}), \quad j = 1, 2. \tag{25.6}$$

For $\vartheta \to 0$, the expressions lead to those for the exponential distribution given in (25.4).

If ϑ is unknown, Cramer and Schmiedt [305] showed that the MLE of ϑ is obtained as the value that maximizes the profile log-likelihood

$$p(\vartheta) = m \log \vartheta - m \log \Big(\sum_{j=1}^{m} (R_j + 1) \log(1 + \vartheta\, x_j) \Big) - \sum_{j=1}^{m} \log(1 + \vartheta\, x_j)\,.$$

It can be shown that p is bounded but the maximum may be attained at zero yielding the estimate $\widehat{\vartheta} = 0$. This is interpreted in the sense that the underlying distributions are exponentials. This illustrates that $\widehat{\vartheta}$ can be used to construct a test for testing exponential distribution against Lomax alternatives. For uncensored data, such an approach has been discussed in Kozubowski et al. [548]. In this case, the MLEs of β_1 and β_2 have the representation (25.4).

For $\widehat{\vartheta} > 0$, the MLE of β_j is $\widehat{\beta}_j(\widehat{\vartheta})$, $j = 1, 2$ (see (25.6)).

Example 25.1.6. The data from Example 25.1.1 have also been discussed in Cramer and Schmiedt [305]. They found the estimate $\widehat{\vartheta} = 0$ showing that exponential models yield the best fit. Hence, the estimates of β_1 and β_2 are taken from Example 25.1.1.

25.2 Joint Progressive Censoring

In joint progressive censoring, the sample is based on two baseline samples $X_1, \ldots, X_{n_1}$ (product/Type A) and $Y_1, \ldots, Y_{n_2}$ (product/Type B) of independent random variables, where

$$X_i \sim F_1, \qquad Y_j \sim F_2.$$

The progressive censoring is applied to the pooled sample

$$\{X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}\}$$

given a prefixed censoring scheme $\mathscr{R} \in \mathscr{C}^m_{m,n_1+n_2}$. Moreover, it is assumed that the type of the failed unit and the types of withdrawn units are known. Therefore, the sample is given by $(\mathbf{C}, \mathbf{W}^{\mathscr{R}}, \mathscr{S})$, where

$$\begin{aligned} \mathbf{C} &= (C_1, \ldots, C_m) \in \{0, 1\}^m, \\ \mathbf{W}^{\mathscr{R}} &= (W_{1:m:n_1+n_2}, \ldots, W_{m:m:n_1+n_2}), \\ \mathscr{S} &= (S_1, \ldots, S_m). \end{aligned}$$

The indicators C_j have the value 1 if the failed unit is of Type A, and otherwise $C_j = 0$. $W_{j:m:n}$ denotes the jth failure time in the progressively censored experiment. Finally, $\mathscr{S}$ denotes a random censoring scheme. S_j is the number of removed units of Type A in the jth withdrawal. Thus, $R_j - S_j$ denotes the numbers of withdrawn units of Type B at the jth censoring step.

Assuming that F_1 and F_2 are absolutely continuous with density functions f_1 and f_2, respectively, the joint density function is given by

$$f^{(\mathbf{C},\mathbf{W}^{\mathscr{R}},\mathscr{S})}(\mathbf{c}_m,\mathbf{w}_m,\mathbf{s}_m) = D(\mathbf{c}_m,\mathbf{s}_m)\prod_{j=1}^{m}\Big[f_1^{c_i}(w_i)f_2^{1-c_i}(w_i)\overline{F}_1^{s_i}(w_i)\overline{F}_2^{R_i-s_i}(w_i)\Big],$$

where $w_1 \le \cdots \le w_m$ and D is a normalizing constant. The normalizing constant $D(\mathbf{c}_m,\mathbf{s}_m)$ is given by the product $D_1(\mathbf{c}_m,\mathbf{s}_m)D_2(\mathbf{c}_m,\mathbf{s}_m)$. Introducing the notation $\gamma_j^{(A)} = n_1 - \sum_{i=1}^{j-1}(c_i + s_i)$, $\gamma_j^{(B)} = n_2 - \sum_{i=1}^{j-1}(1 - c_i + R_i - s_i)$, and $\gamma_j^{(AB)} = n_1 + n_2 - \sum_{i=1}^{j-1}(R_i + 1)$ for the number of units of Type A, B, and A/B, respectively, remaining in the experiment before the jth failure, the factors are defined by

$$D_1(\mathbf{c}_m,\mathbf{s}_m) = \prod_{j=1}^{m}[c_j\gamma_j^{(A)} + (1-c_j)\gamma_j^{(B)}],$$

$$D_2(\mathbf{c}_m,\mathbf{s}_m) = \prod_{j=1}^{m-1}\frac{\binom{\gamma_j^{(A)}-c_j}{s_j}\binom{\gamma_j^{(B)}-1+c_j}{R_j-s_j}}{\binom{\gamma_j^{(AB)}-1}{R_j}}.$$

Notice that $\gamma_j^{(AB)} = \gamma_j^{(A)} + \gamma_j^{(B)}$.

Rasouli and Balakrishnan [747] considered inference for joint progressively censored data from $\mathsf{Exp}(\vartheta_1)$- and $\mathsf{Exp}(\vartheta_2)$-distributions extending results for Type-II right censored data established in Balakrishnan and Rasouli [118]. In this case, the likelihood function $\ell(\vartheta_1,\vartheta_2)$ is proportional to

$$\frac{1}{\vartheta_1^{k_1}\vartheta_2^{m-k_1}}\exp\Big\{-\frac{u_1}{\vartheta_1}-\frac{u_2}{\vartheta_2}\Big\},$$

where k_1 and $k_2 = m - k_1$ denote the number of observed failures from the X- and Y-samples, respectively, and

$$u_1 = \sum_{j=1}^{m}(c_j + s_j)w_j, \qquad u_2 = \sum_{j=1}^{m}(1 - c_j + R_j - s_j)w_j.$$

Clearly, the MLE of ϑ_j is given by $\widehat{\vartheta}_j = \frac{1}{k_j}U_j$ provided that $k_j > 0$, $j = 1, 2$. Rasouli and Balakrishnan [747] obtained explicit but complicated expressions for the (conditional) moment generating function, the density function, and the survival function as well as for the mean, second moment, and covariance of the MLEs $\widehat{\vartheta}_j$, $j = 1, 2$. Assuming independent gamma priors, they found Bayesian estimates for

Plane	Ordered failure times
7914	3 5 5 13 14 15 22 22 23 30 36 39 44 46 50 72 79 88 97 102 139 188 197 210
7913	1 4 11 16 18 18 18 24 31 39 46 51 54 63 68 77 80 82 97 106 111 141 142 163 191 206 216

Table 25.1 Failure times of air-conditioning systems in airplanes 7913 and 7914 taken from Proschan [730]

$\mathbf{w}_{10}$	1	3	4	5	5	13	14	15	18	18
$\mathbf{c}_{10}$	0	1	0	1	0	1	1	1	0	0
$\mathbf{s}_{10}$	1	0	1	2	1	1	1	1	2	0

Table 25.2 Jointly progressive Type-II censored sample that resulted from the data in Table 25.1 with $m = 10$ and $\mathscr{R} = (2^{*9}, 23)$

ϑ_j and credible intervals. Moreover, different methods of constructing confidence intervals are also discussed. They also analyzed data generated from Proschan's Boeing data as reported in Table 25.1.

Example 25.2.1. Assuming exponential lifetimes, Rasouli and Balakrishnan [747] analyzed the joint progressively censored data in Table 25.2. They obtained the estimates $\widehat{\vartheta}_1 = 62.612$ and $\widehat{\vartheta}_2 = 74.23$. They also computed the exact 95 % confidence intervals $(25.61, 169.43)$ and $(35.11, 225.23)$, respectively. Alternative confidence intervals based on other methods and for other censoring schemes are also presented.

Parsi et al. [713] considered Weibull$(\vartheta_1^{\beta_1}, \beta_1)$- and Weibull$(\vartheta_2^{\beta_2}, \beta_2)$-distributions and presented the likelihood equations and provided computational solutions for the MLEs as well as an extensive simulation study.

25.3 Concomitants

For bivariate IID data $(X_1, Y_1), \dots, (X_n, Y_n)$, David [322] introduced the notion of concomitants of order statistics. The data are ordered w.r.t. the first component leading to order statistics $X_{1:n} \le \cdots \le X_{n:n}$, while the second component is accompanying the first one. The value of the second component associated with $X_{r:n}$ is called the rth concomitant $Y_{[r:n]}$. Many results have been established under various distributional assumptions. For comprehensive reviews, we refer to David and Nagaraja [326, 327].

In a similar manner, concomitants for progressively Type-II censored order statistics can be introduced. Suppose (X_j, Y_j), $1 \le j \le n$, are independent and have a bivariate density function f. Then, the joint density function of the concomitants $\mathbf{Y}_{[m]}^{\mathscr{R}} = (Y_{[1:m:n]}, \dots, Y_{[m:m:n]})'$ can be written as

$$f^{\mathbf{Y}_{[m]}^{\mathscr{R}}}(\mathbf{y}_m) = \int f^{\mathbf{Y}_{[m]}^{\mathscr{R}}|\mathbf{X}^{\mathscr{R}}}(\mathbf{y}_m|\mathbf{x}_m) f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}_m) d\mathbf{x}_m.$$

Following the same arguments as in Bhattacharya [197], the subsequent conditional independence result is true (see also Bhattacharya [196] and Yang [932]).

Lemma 25.3.1. Given $\mathbf{X}^{\mathscr{R}} = \mathbf{x}_m$, the concomitants $Y_{[1:m:n]}, \ldots, Y_{[m:m:n]}$ are independent and identically distributed. The (conditional) joint density function is given by

$$f^{\mathbf{Y}_{[m]}^{\mathscr{R}}|\mathbf{X}^{\mathscr{R}}}(\mathbf{y}_m|\mathbf{x}_m) = \prod_{i=1}^{m} f(y_i|x_i). \tag{25.7}$$

The joint density function $f^{\mathbf{X}^{\mathscr{R}},\mathbf{Y}_{[m]}^{\mathscr{R}}}$ is given by

$$f^{\mathbf{X}^{\mathscr{R}},\mathbf{Y}_{[m]}^{\mathscr{R}}}(\mathbf{x}_m, \mathbf{y}_m) = \Big[\prod_{j=1}^{m} \gamma_j\Big]\Big[\prod_{i=1}^{m} f(y_i|x_i) f(x_i) \overline{F}^{R_i}(x_i)\Big].$$

The marginal density function of the concomitants $\mathbf{Y}_{[m]}^{\mathscr{R}}$ has the representation

$$f^{\mathbf{Y}_{[m]}^{\mathscr{R}}}(\mathbf{y}_m) = \int \prod_{i=1}^{m} f(y_i|x_i) f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{x}_m) d\mathbf{x}_m. \tag{25.8}$$

Obviously, marginal density functions can be obtained directly from (25.8) by integrating out the corresponding variables. For instance, let $1 \le r_1 < \cdots < r_k \le m$. Then,

$$f^{Y_{[r_1:m:n]},\ldots,Y_{[r_k:m:n]}}(\mathbf{y}_k) = \int \prod_{i=1}^{k} f(y_i|x_i) f^{X_{r_1:m:n},\ldots,X_{r_k:m:n}}(\mathbf{x}_k) d\mathbf{x}_k$$

which is given in Izadi and Khaledi [473]. Using (2.28), Bairamov and Eryılmaz [78] presented the marginal density functions in terms of density functions of concomitants of minima as

$$f^{Y_{[r:m:n]}}(y) = \int f^{X_{r:m:n}}(x) f(y|x) dx = \Big(\prod_{i=1}^{r} \gamma_i\Big) \sum_{j=1}^{r} \frac{a_{j,r}}{\gamma_j} f^{Y_{[1:\gamma_j]}}(y).$$

This formula can be easily applied to compute moments. For instance, Bairamov and Eryılmaz [78] considered the bivariate Farlie–Gumbel–Morgenstern distribution with density function

$$f(x, y) = 1 + \alpha(1 - 2x)(1 - 2y), \quad x, y \in (0, 1), \alpha \in (-1, 1).$$

A direct calculation shows that

$$f^{Y_{[r:m:n]}}(y) = 1 + \alpha(1-2y)\Big(\prod_{i=1}^{r}\gamma_i\Big)\sum_{j=1}^{r}\frac{a_{j,r}}{\gamma_j}\frac{\gamma_j-1}{\gamma_j+1}$$
$$= 1 + \alpha(1-2y)\Big(1-2\prod_{i=1}^{r}\frac{\gamma_i}{\gamma_i+1}\Big), \quad y \in (0,1).$$

For related results in terms of generalized order statistics based on the Farlie–Gumbel–Morgenstern distribution, we refer to Beg and Ahsanullah [185].

Izadi and Khaledi [473] studied stochastic orderings of concomitants from progressively Type-II censored order statistics. Given the assumptions of Theorem 3.2.6 and Remark 3.2.7, they found the properties w.r.t. likelihood ratio order, (reversed) hazard rate order, stochastic order, and mean residual life order. For instance, let $n_1 = \gamma_1(\mathscr{S}), n_2 = \gamma_1(\mathscr{R})$. Then, for $j \le i$,

(i) If $f(x,y)$ is TP_2 in (x,y), then $Y^{\mathscr{S}}_{[j:m:n_1]} \le_{\mathsf{lr}} Y^{\mathscr{R}}_{[i:m:n_2]}$;
(ii) If the conditional hazard rate $\lambda(y|x) = \frac{f(y|x)}{1-F(y|x)}$ is decreasing in x, then $Y^{\mathscr{S}}_{[j:m:n_1]} \le_{\mathsf{hr}} Y^{\mathscr{R}}_{[i:m:n_2]}$.

25.3.1 Missing Information Principle and EM-Algorithm

Progressively Type-II censored concomitant data from bivariate normal distributions with parameter $\theta = (\mu_X, \sigma_X, \mu_Y, \sigma_Y, \rho)$ have been considered by Balakrishnan and Kim [106]. They applied a missing information principle like the one described in Sect. 9.1.2 to the concomitant data. Let $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_m)$ be the vector of progressively censored random variables, where $\mathbf{W}_j = (W_{j1}, \ldots, W_{jR_j})$ denotes those random variables corresponding to units withdrawn in the jth step of the progressive censoring procedure. Moreover, let $\mathbf{V}_{[\cdot]} = (\mathbf{V}_{[1]}, \ldots, \mathbf{V}_{[m]})$ be the vector of corresponding concomitants, where $\mathbf{V}_{[j]} = (V_{[j1]}, \ldots, V_{[jR_j]})$. Then, a result similar to Theorem 9.1.8 has been established by Balakrishnan and Kim [106].

Theorem 25.3.2. Given $(\mathbf{X}^{\mathscr{R}}_j, \mathbf{Y}^{\mathscr{R}}_{[j]}) = (\mathbf{x}_j, \mathbf{y}_j)$, the conditional density function of $(W_{jk}, V_{[jk]})$, $k \in \{1, \ldots, R_j\}$, is given by

$$f^{W_{jk},V_{[jk]}|\mathbf{X}^{\mathscr{R}}_j,\mathbf{Y}^{\mathscr{R}}_{[j]}}(w,v|\mathbf{x}_j,\mathbf{y}_j) = f^{W_{jk},V_{[jk]}|X_{j:m:n}}(w,v|x_j) = \frac{f(w,v)}{1-F(x_j)},$$
$$w > x_j, v \in \mathbb{R}, \quad (25.9)$$

and $(W_{jk}, V_{[jk]})$ and $(W_{j\ell}, V_{[j\ell]})$, $k \ne \ell$, are conditionally independent, given $X_{j:m:n} = x_j$.

Representation (25.9) yields expressions for the (conditional) moments of the missing concomitants. In particular, for the bivariate normal distribution, Balakrishnan and Kim [106] obtained

$$\begin{aligned}
E[V_{[jk]}|X_{j:m:n}=x) &= \mu_Y+\rho\sigma_Y h(z),\\
E[V_{[jk]}^2|X_{j:m:n}=x) &= \sigma_Y^2(1+\rho^2 zh(z))+2\rho\sigma_Y\mu_Y h(z)+\mu_Y^2,\\
E[W_{jk}V_{[jk]}|X_{j:m:n}=x) &= \mu_Y(\mu_X+\sigma_X h(z))\\
&\quad+\rho\sigma_Y(\mu_X h(z)+\sigma_X+\sigma_X zh(z)),
\end{aligned}$$

where $z=z(x)=\frac{x-\mu_X}{\sigma_X}$ and $h(x)=\frac{\varphi(x)}{1-\Phi(x)}$ denote the hazard rate of the standard normal distribution. These expressions have been used to develop an EM-algorithm to obtain maximum likelihood estimates of the parameters of the bivariate normal distribution. For details, we refer to Balakrishnan and Kim [106].

25.4 Progressively Censored Systems Data

Originally, inference for progressively Type-II censored data addresses the population cumulative distribution function F of the original sample $X_1,\ldots,X_n$, i.e., $\mathbf{X}^{\mathscr{R}}$ is a progressively censored sample from the cumulative distribution function F. However, it suggests itself that the original data $X_1,\ldots,X_n$ may correspond to observations of lifetimes of systems. In this case, the system lifetime depends in some way on the component lifetime cumulative distribution function F.

Example 25.4.1 (First-failure progressively censored data). Wu and Kuş [913] introduced progressive censoring in the presence of first-failure censoring. Here, the original data can be interpreted as observations from equally structured series systems with k components. In this case, one has observations of n series systems with k IID components each having a lifetime distribution F. Therefore, the system lifetime is given by $X_j=Y_{1:k}^{(j)}=\min\{Y_{j1},\ldots,Y_{jk}\}$, where $Y_{ji}\sim F$ denotes the lifetime of the ith component in the jth series system. Hence, $X_{1:m:n}^{\mathscr{R}},\ldots,X_{m:m:n}^{\mathscr{R}}$ is a progressively Type-II censored sample with censoring scheme $\mathscr{R}$, but with lifetime distribution $F_k=1-(1-F)^k$. However, inference is carried out for F and not for F_k.

Extending Example 25.4.1, it is natural to consider lifetimes $X_1,\ldots,X_n$ of n independent systems whose component lifetime cumulative distribution function is given by F. As an example, we may consider $(n-r+1)$-out-of-k systems so that $X_j=Y_{r:k}^{(j)}$ is a sample of order statistics. In general, one may consider monotone systems with structure function ϕ so that the system lifetime Y is given by

$$Y=\phi(Y_1,\ldots Y_k)$$

provided that the system has some (IID) component lifetimes $Y_1,\ldots,Y_k$.

25.4.1 Progressive First-Failure Censoring: Series Systems

As mentioned in Example 25.4.1, progressively Type-II censored order statistics generated by first-failure censoring can be interpreted as progressively Type-II censored order statistics from a transformed cumulative distribution function. In particular, the quantile function of F_k is given by

$$F_k^{\leftarrow}(t) = F^{\leftarrow}\big(1-(1-t)^{1/k}\big), \quad t \in (0,1).$$

Using the quantile representation (2.3.6), we obtain

$$X_{r:m:n} \stackrel{\mathsf{d}}{=} F_k^{\leftarrow}\left(1-\prod_{j=1}^{r} U_j^{1/\gamma_j}\right) \stackrel{\mathsf{d}}{=} F^{\leftarrow}\left(1-\prod_{j=1}^{r} U_j^{1/(k\gamma_j)}\right), \quad 1 \le r \le m.$$

Hence, we can interpret the data as progressively Type-II censored order statistics from population cumulative distribution function F with parameters $k\gamma_1, \dots, k\gamma_m$. The corresponding censoring scheme is given by

$$\mathscr{S} = k\mathscr{R} + (k-1)(1^{*m}), \text{ i.e., } S_j = kR_j + k - 1, j = 1, \dots, m.$$

This connection illustrates that all results obtained for progressively Type-II censored order statistics can be easily adapted to progressive first-failure censoring.

Therefore, the results of Wu and Kuş [913] for the Weibull distribution in the presence of first-failure censoring parallel those for the Weibull distribution for standard progressive censoring. The same comment applies to the work on Gompertz distribution by Soliman et al. [817] (see Wu et al. [914]) and to Burr Type-XII distribution by Soliman et al. [815,819]. Further papers in this direction are Soliman et al. [814], Hong et al. [449], Wang and Shi [888], Wu and Huang [912], Lio and Tsai [613], Ahmadi et al. [18], and Mahmoud et al. [630].

Stress–strength estimation based on the first-failure censored data is considered in Lio and Tsai [613].

25.4.2 Parallel Systems

For failure data on parallel systems, the structure function is given by $\phi(\mathbf{x}_k) = \max_{1\le j\le k} x_j$. Hence, we consider progressively Type-II censored order statistics with censoring scheme $\mathscr{R}$ and cumulative distribution function $F_k = F^k$. Such a situation has been considered in Pradhan [726] and Potdar and Shirke [725]. For instance, Pradhan [726] assumed exponential lifetimes leading to the cumulative distribution function

$$F_k(t;\lambda) = \left(1 - \mathrm{e}^{-\lambda t}\right)^k, \quad t \geq 0,$$

with some $\lambda > 0$ (see (12.41) but, here, k is known). He then discussed likelihood inference for λ resulting in a nonlinear likelihood equation which has to be solved numerically. He proposed a Newton–Raphson procedure to compute the estimate. Moreover, confidence and tolerance intervals are discussed. Potdar and Shirke [725] generalized this setting by considering a scale family of distributions with standard member G, i.e.,

$$F_k(t;\lambda) = G^k(t/\lambda), \quad t \geq 0.$$

To obtain the MLE, they proposed a procedure based on the EM-algorithm as well as a Newton–Raphson procedure. As an example, they discussed the half-logistic distribution.

It is obvious from the model assumptions that the related inference can be seen as that of progressively Type-II censored order statistics from an exponentiated distribution, where the shape parameter k is supposed to be known. In this setting, Asgharzadeh [62] discussed approximate MLEs for the exponentiated exponential distribution so that his results can be applied to parallel systems.

Chapter 26
Optimal Experimental Designs

The problem of finding optimal elements of $\mathscr{C}_{m,n}^m$ has been discussed in many setups since the problem was first addressed in Balakrishnan and Aggarwala [86] (see also Sect. 10 in Balakrishnan [84]). Many optimality criteria have been proposed, and many results on optimal censoring designs have been established. In particular, given some optimality criterion, Burkschat [228] has established general results using stochastic orders of the underlying progressively Type-II censored samples. Balakrishnan and Aggarwala [86] and Burkschat et al. [235, 237] obtained optimal censoring designs in terms of minimum variance of best linear unbiased estimators. Ng et al. [689] considered minimum variance criteria for maximum likelihood estimates for Weibull distributions. Precision of quantile estimates is considered in Balakrishnan and Han [98], Kundu and Pradhan [562], Pradhan and Kundu [727], and Pareek et al. [703]. Fisher information in a progressively Type-II censored sample is addressed in Balakrishnan et al. [140], Abo-Eleneen [6], Pareek et al. [703], and Cramer and Schmiedt [305]. Further approaches on optimal experimental design in progressive censoring are discussed in Hofmann et al. [444] and Burkschat [227].

Here, it is assumed that progressive Type-II censoring is carried out by design. The experimenter may fix an appropriate censoring plan prior to the start of the experiment. Furthermore, we assume that $n \geq m$ units are available for test and that m failure times have to be observed so that $n-m = \sum_{i=1}^{m} R_i$ (surviving) units need to be withdrawn from the experiment. The crucial question is

How should the experimenter choose the censoring numbers $R_1, \ldots, R_m$?

A general answer to this problem cannot be expected. In particular, it turns out that the optimal designs depend heavily on both the optimality criterion to be used and the distributional assumption made.

Given an optimality criterion, it is clear that an optimal censoring scheme exists in any case because the optimization is based on a finite number of schemes only. But, since the set of admissible schemes $\mathscr{C}_{m,n}^m$ has $\binom{n-1}{m-1}$ elements (see (1.5)), it follows by the approximation $\binom{n-1}{m-1} \overset{n\to\infty}{\sim} \frac{n^{m-1}}{(m-1)!}$ that enumeration will not work

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7_26,

even for moderate values of n and m. In particular, a sample size of $n = 100$ and $m = 20$ observed failure times will lead to more than 10^{20} possible schemes. This illustrates the need for some mathematical theory to tackle this optimization problem.

Although the number of schemes grows fast, the structure of $\mathscr{C}_{m,n}^m$ is quite nice. The convex hull of $\mathscr{C}_{m,n}^m$ forms a simplex of dimension $m-1$ in $\mathbb{R}^m$ whose vertices are given by the OSPs $\mathscr{O}_1, \ldots, \mathscr{O}_m$. This shows that the one-step censoring plans are extremal points of the set of admissible schemes which are of particular interest in optimization with regard to this set. In particular, this observation explains why one-step plans turn out to be optimal in many situations. Before going into further details, we will discuss a general approach of finding optimal censoring designs as it can be found in Burkschat [228].

26.1 Preliminaries

First, we formulate the problem as a mathematical optimization problem.

Definition 26.1.1. A mapping $\psi : \mathscr{C}_{m,n}^m \longrightarrow \mathbb{R}$ is said to be an optimality criterion. A censoring scheme $\mathscr{S}$ is said to be ψ-optimal if

$$\psi(\mathscr{S}) = \min_{\mathscr{R} \in \mathscr{C}_{m,n}^m} \psi(\mathscr{R}). \tag{26.1}$$

It is clear from the above definition that we can consider maximization instead of minimization. This problem is included by choosing $-\psi$ instead of ψ so that the corresponding solution of the maximization is $-\psi$-optimal. Therefore, we can restrict ourselves to the minimization problem, having in mind obvious changes.

A first, but very effective property, in determining optimal solutions results from a partial ordering on $\mathscr{C}_{m,n}^m$ introduced in Definition 7.1.9: For $\mathscr{R}, \mathscr{S} \in \mathscr{C}_{m,n}^m$,

$$\mathscr{R} \preccurlyeq \mathscr{S} \iff \sum_{i=1}^k R_i \le \sum_{i=1}^k S_i, \; k = 1, \ldots, m-1.$$

Recalling that

$$\mathscr{O}_m = (0^{*m-1}, n-m) \preccurlyeq \mathscr{R} \preccurlyeq (n-m, 0^{*m-1}) = \mathscr{O}_1 \tag{26.2}$$

(see Lemma 7.1.11), we get the following result.

Theorem 26.1.2. If ψ has some monotonicity properties w.r.t. $\preccurlyeq$, then $\mathscr{O}_1$ or $\mathscr{O}_m$ are ψ- or $-\psi$-optimal.

This result illustrates that first-step- or right censoring will be optimal provided that the objective function exhibits some monotonicity properties w.r.t. the partial ordering $\preccurlyeq$. As can be seen subsequently, this result is a powerful tool to find ψ-optimal experimental designs.

Another important tool in tackling the optimization problem is the following reformulation. The optimization problem may be rewritten by replacing $\mathscr{C}^m_{m,n}$ by the set (see (1.4))

$$\mathscr{G}_{m,n} = \{(\gamma_2, \dots, \gamma_m) \in \mathbb{N}^{m-1} \,|\, n > \gamma_2 > \cdots > \gamma_m \geq 1\}.$$

Denoting the number of objects remaining in the experiment before the censoring steps by $\gamma_k(\mathscr{R}) = \sum_{j=k}^m (R_j + 1)$, $k = 1, \dots, m$ (see (1.2)), and the corresponding vector by $\boldsymbol{\gamma}(\mathscr{R})$, we have a one-to-one correspondence $\boldsymbol{\gamma}(\mathscr{R}) \leftrightarrows \mathscr{R}$. Namely, we have the linear transformation

$$\boldsymbol{\gamma}(\mathscr{R}) = A\Big(\mathscr{R} + \mathbb{1}_m\Big) \quad \text{with } A = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

Moreover, the image of this mapping is given by

$$\mathscr{G}^*_{m,n} = \{(n, \gamma_2, \dots, \gamma_m) \in \mathbb{N}^m \,|\, n > \gamma_2 > \cdots > \gamma_m \geq 1\}.$$

Notice that the first component must equal n. Since A is a regular matrix, we define κ by the inverse relation

$$\mathscr{R} = \kappa(\boldsymbol{\gamma}(\mathscr{R})) = A^{-1}\boldsymbol{\gamma}(\mathscr{R}) - \mathbb{1}_m.$$

Hence, the optimization problem (26.1) is equivalent to the following problem, where ψ is replaced by $\overline{\psi} = \psi \circ \kappa$:

$$\overline{\psi}(\boldsymbol{\gamma}(\mathscr{O})) = \min_{\mathscr{R} \in \mathscr{C}^m_{m,n}} \overline{\psi}(\boldsymbol{\gamma}(\mathscr{R})) = \min_{\boldsymbol{\gamma}(\mathscr{R}) \in \mathscr{G}^*_{m,n}} \overline{\psi}(\boldsymbol{\gamma}(\mathscr{R})).$$

Obviously, the optimization process depends only on $\gamma_2, \dots, \gamma_m$. In many cases, it is easier to solve the minimization problem

$$\overline{\psi}(\boldsymbol{\gamma}^*) = \min_{\boldsymbol{\gamma} \in \mathscr{G}^*_{m,n}} \overline{\psi}(\boldsymbol{\gamma}) = \min_{(\gamma_2, \dots, \gamma_m) \in \mathscr{G}_{m,n}} \overline{\psi}((n, \gamma_2, \dots, \gamma_m))$$

than (26.1). Then, the optimal plan is computed from the relation $\mathscr{O} = \kappa(\boldsymbol{\gamma}^*)$.

Introducing the componentwise ordering

$$\boldsymbol{\gamma}(\mathscr{R}) \geq \boldsymbol{\gamma}(\mathscr{S}) \iff \gamma_k(\mathscr{R}) \geq \gamma_k(\mathscr{S}),\ k = 1, \dots, m,$$

as a partial ordering on $\mathscr{G}^*_{m,n}$, this shows

$$\mathscr{R} \preccurlyeq \mathscr{S} \iff \boldsymbol{\gamma}(\mathscr{R}) \geq \boldsymbol{\gamma}(\mathscr{S}). \tag{26.3}$$

26.2 Probabilistic Criteria

Burkschat [228] considered several criteria based on the experimental time $X^{\mathscr{R}}_{m:m:n}$ of a progressively Type-II censored experiment and on the total time on test statistic $T^{\mathscr{R}} = \sum_{j=1}^{m}(R_j + 1)X^{\mathscr{R}}_{j:m:n}$. The results are based on stochastic orderings of the involved random variables. The crucial result used for experimental time $X^{\mathscr{R}}_{m:m:n}$ is based on the following stochastic ordering result due to Burkschat [228] (see also Balakrishnan et al. [140] and Dahmen et al. [320]).

Lemma 26.2.1. Let $1 \le j \le m$ and $\mathscr{S} \preccurlyeq \mathscr{R}$. Then,

$$X^{\mathscr{R}}_{j:m:n} \le_{\mathrm{st}} X^{\mathscr{S}}_{j:m:n}. \tag{26.4}$$

(26.4) implies with $j = m$

$$X^{\mathscr{R}}_{m:m:n} \le_{\mathrm{st}} X^{\mathscr{S}}_{m:m:n} \quad \text{provided that } \mathscr{S} \preccurlyeq \mathscr{R}.$$

Hence, we readily get the ordering (cf. Lemma 7.1.11)

$$X^{\mathscr{O}_m}_{m:m:n} \le_{\mathrm{st}} X^{\mathscr{R}}_{m:m:n} \le_{\mathrm{st}} X^{\mathscr{O}_1}_{m:m:n} \tag{26.5}$$

which in turn implies the ordering of the expectations $EX^{\mathscr{O}_m}_{m:m:n} \le EX^{\mathscr{R}}_{m:m:n} \le EX^{\mathscr{O}_1}_{m:m:n}$. This yields the following theorem due to Burkschat [228].

Theorem 26.2.2. Let $\psi(\mathscr{R}) = EX^{\mathscr{R}}_{m:m:n}$ denote the expected experimental time. Then, the unique optimal censoring plan is given by $\mathscr{O}_m$, i.e., right censoring yields the shortest expected experimental time. Moreover, $\psi(\mathscr{O}_m)$ is decreasing in n and increasing in m.

The result tells us that if we wish to finish the experiment in shortest (expected) time, then we have to conduct the experiment according to right censoring which is quite intuitive. Alternatively, we may consider the probability that the experiment will be finished with high probability before a given threshold T. In this case, we are naturally interested in identifying the censoring plan minimizing the probability $P(X^{\mathscr{R}}_{m:m:n} > T)$. As can be easily seen, everything to solve this problem is already done. The solution is directly connected to the stochastic ordering given in (26.5), and we get the following result. It is worth mentioning that T may also be a random variable independent of $X^{\mathscr{R}}_{m:m:n}$.

Theorem 26.2.3. Let $\psi(\mathscr{R}) = P(X^{\mathscr{R}}_{m:m:n} > T)$ with a given threshold T. Then, the optimal censoring plan is given by $\mathscr{O}_m$, i.e., right censoring yields the smallest exceedance probability. Moreover, $\psi(\mathscr{O}_m)$ is decreasing in n and increasing in m.

Burkschat [228] also considered the variance of the experimental time $\psi(\mathscr{R}) = \mathrm{Var}\, X^{\mathscr{R}}_{m:m:n}$ as an optimality criterion. He established a result for DFR distributions which is based on the dispersive ordering of the experimental time. In particular, given that the baseline distribution is DFR, the following result can be established along the lines of Theorem 3.4 in Hu and Zhuang [457].

Lemma 26.2.4. Let F be a DFR distribution. Then, $\mathscr{S} \preccurlyeq \mathscr{R}$ implies

$$X^{\mathscr{R}}_{m:m:n} \leq_{\mathsf{disp}} X^{\mathscr{S}}_{m:m:n}. \tag{26.6}$$

Proof. Let $\mathscr{S} \preccurlyeq \mathscr{R}$. Then, we get from (26.3) the ordering $\boldsymbol{\gamma}(\mathscr{S}) \geq \boldsymbol{\gamma}(\mathscr{R})$ so that $\boldsymbol{\gamma}(\mathscr{S}) \preccurlyeq_{\mathsf{p}} \boldsymbol{\gamma}(\mathscr{R})$. From Theorem 3.2.11, we then conclude the dispersive ordering in (26.6) □

Since dispersive ordering implies ordering of variance (see Shaked and Shanthikumar [799]), we obtain the following result. For the uniqueness part, we refer to Burkschat [228].

Theorem 26.2.5. Let F be a DFR distribution and $\psi(\mathscr{R}) = \mathsf{Var}\, X^{\mathscr{R}}_{m:m:n}$. Then, the unique optimal censoring plan is given by $\mathscr{O}_m$, i.e., right censoring yields the smallest variance of the experimental time. Moreover, $\psi(\mathscr{O}_m)$ is decreasing in n and increasing in m.

As noted by Burkschat [228], the entropy of $X^{\mathscr{R}}_{m:m:n}$, i.e., $\mathscr{H}(X^{\mathscr{R}}_{m:m:n}) = -E \log f_{m:m:n}\big(X^{\mathscr{R}}_{m:m:n}\big)$, may also serve as a measure of uncertainty. As shown by Oja [695, p. 160] (see also Ebrahimi et al. [346]), entropies of dispersively ordered random variables are ordered, too. Therefore, Lemma 26.2.4 yields the corresponding result for $\mathscr{H}(X^{\mathscr{R}}_{m:m:n})$.

Finally, Burkschat [228] has addressed the total time on test statistic

$$T^{\mathscr{R}} = \sum_{j=1}^{m} (R_j + 1) X^{\mathscr{R}}_{j:m:n}$$

with regard to both minimal expectation and minimal variance. In Theorem 4.7 of Burkschat [225], he presented an optimality result for the expected time on test provided that the population distribution is either IFR or DFR. In order to establish this result, he used the alternative expression in terms of spacings

$$T^{\mathscr{R}} = \sum_{j=1}^{m} \gamma_j(\mathscr{R})(X^{\mathscr{R}}_{j:m:n} - X^{\mathscr{R}}_{j-1:m:n}),$$

where $X^{\mathscr{R}}_{0:m:n} = 0$. The proof is based on the multivariate stochastic order of normalized spacings (see Theorem 3.2.29) as given in Theorem 4.5 of Burkschat [228].

Theorem 26.2.6. Let $\psi(\mathscr{R}) = ET^{\mathscr{R}} = \sum_{j=1}^{m} (R_j + 1) EX^{\mathscr{R}}_{j:m:n}$ denote the expected total time on test. Then, the unique optimal censoring plan is given by

(i) $\mathscr{O}_m$, i.e., right censoring yields the shortest total time on test if F is a DFR-distribution function. Moreover, $\psi(\mathscr{O}_m)$ is decreasing in n and increasing in m;
(ii) $\mathscr{O}_1$, i.e., first-step censoring yields the shortest total time on test if F is an IFR-distribution function. Moreover, $\psi(\mathscr{O}_1)$ is increasing in both n and m.

It is worth mentioning that in the case of exponential distribution—which is both IFR and DFR—the expected total time on test does not depend on the censoring plan. Thus, any censoring scheme yields the same expected time.

In case of the variance of the total time on test statistic, i.e., $\psi(\mathscr{R}) = \mathsf{Var}(T^{\mathscr{R}})$, Burkschat [228] also obtained an optimality result provided that F is a DFR-distribution function and that S_F as given in (3.9) exhibits a convex derivative. In that case, right censoring $\mathscr{O}_m$ is also optimal. Moreover, he showed that $\psi(\mathscr{O}_m)$ is decreasing in n and increasing in m.

26.3 Precision of Estimates

Precision of estimates as an optimality criterion was suggested by Balakrishnan and Aggarwala [86]. Suppose we are interested in a quantity θ which could be estimated by an estimator $\widehat{\theta}$. Then, it is recommended to identify a censoring plan such that the mean squared error of $\widehat{\theta}$ is as small as possible. This yields the optimization problem

$$\psi(\mathscr{R}) = \mathsf{MSE}(\widehat{\theta}) \longrightarrow \min_{\mathscr{R}\in\mathscr{C}^m_{m,n}} .$$

In order to solve this problem at least numerically, the objective function must be calculable for any censoring scheme $\mathscr{R}$. A closed-form expression is, of course, preferable. We follow the approach of Balakrishnan and Aggarwala [86] and present results based on best linear unbiased estimates in location–scale families of distribution. Results for location families or scale families are just mentioned. According to (11.1), we assume a location–scale family

$$\mathscr{F}_{ls} = \left\{F\left(\frac{\cdot - \mu}{\vartheta}\right)\middle|\mu \in \mathbb{R}, \vartheta > 0\right\},$$

where F is a given continuous distribution with mean zero and variance one. Thus, the BLUEs $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$ are used in the estimation process. The optimization problem is based on the variance–covariance matrix of these estimators, i.e., on the matrix

$$\vartheta^2\Sigma_*(\mathscr{R}) = \frac{\vartheta^2}{\Delta}\begin{pmatrix}\mathbf{b}'\Sigma^{-1}\mathbf{b} & -\mathbf{b}'\Sigma^{-1}\mathbb{1}\\ -\mathbf{b}'\Sigma^{-1}\mathbb{1} & \mathbb{1}'\Sigma^{-1}\mathbb{1}\end{pmatrix},$$

where $\Delta = \mathbb{1}'\Sigma^{-1}\mathbb{1}\mathbf{b}'\Sigma^{-1}\mathbf{b} - (\mathbb{1}'\Sigma^{-1}\mathbf{b})^2$, $\mathbf{b} = E\mathbf{Y}^{\mathscr{R}}$, $\Sigma = \mathsf{Cov}(\mathbf{Y}^{\mathscr{R}})$, and $\mathbf{Y}^{\mathscr{R}}$ forms a progressively Type-II censored sample based on the standard member F (see Sect. 11.1.1). Notice that the variance–covariance matrix is proportional to $\Sigma_*(\mathscr{R})$ which does not depend on the unknown parameters. It should also be mentioned that the same comment applies to criteria considered subsequently. Therefore, we may take $\vartheta = 1$ without loss of generality. Therefore, the optimization problem can be defined via some function ϕ defined on the set of all admissible matrices $\mathcal{S}_{m,n} = \{\Sigma_*(\mathscr{R}) | \mathscr{R} \in \mathscr{C}_{m,n}^m\}$, i.e.,

$$\psi(\mathscr{R}) = \phi(\Sigma_*(\mathscr{R})) \longrightarrow \min_{\mathcal{S}_{m,n}}.$$

The solution of this problem depends heavily on the choice of ϕ and, thus, a general solution will not be possible in most cases. However, if ϕ exhibits some monotonicity properties w.r.t. the matrices contained in $\mathcal{S}_{m,n}$, a general result will be possible. For nonnegative definite matrices, the so-called Löwner ordering(see Pukelsheim [733]) can be utilized to find such a general characterization.

Definition 26.3.1. Let $A, B \in \mathbb{R}^{n \times n}$ be nonnegative definite matrices. Then, A is called smaller than B in the Löwner ordering $(A \preceq_{\mathsf{L}} B)$ if $B - A$ is a nonnegative definite matrix.

Therefore, the problem can be solved in the family of objective functions ϕ which exhibits some monotonicity properties with respect to Löwner ordering provided that extremal elements exist in the set $\mathcal{S}_{m,n}$. Such a result can be obtained in the case of generalized Pareto distributions (see Sects. 11.2.2 and 26.3.2).

We now introduce some particular functions ϕ discussed in this section:

(i) A wide class of criteria can be defined by considering the objective function

$$\phi_a(\Sigma_*(\mathscr{R})) = a'\Sigma_*(\mathscr{R})a$$

for some given $a \in \mathbb{R}^2 \setminus \{0\}$. Notice that this approach means to minimize the variance of $a'\binom{\widehat{\mu}_{\mathsf{LU}}}{\widehat{\vartheta}_{\mathsf{LU}}}$ which is an unbiased estimator of $a'\binom{\mu}{\vartheta}$. Choosing $a' = (1, 0)$ or $a' = (0, 1)$ corresponds to minimizing the variance of $\widehat{\mu}_{\mathsf{LU}}$ and $\widehat{\vartheta}_{\mathsf{LU}}$, respectively.

An important example is given by the estimation of quantiles in the location–scale setting. Suppose we are interested in estimating the pth quantile. Then, the BLUE of ξ_p is given by

$$\widehat{\xi}_p = \widehat{\mu}_{\mathsf{LU}} + F^{\leftarrow}(p)\widehat{\vartheta}_{\mathsf{LU}}.$$

Hence, $a' = (1, F^{\leftarrow}(p))$ has to be used in the minimization process.

(ii) Another optimality criteria is given by the trace of $\Sigma_*(\mathscr{R})$, i.e.,

$$\phi(\Sigma_*(\mathscr{R})) = \mathrm{tr}(\Sigma_*(\mathscr{R})).$$

Using the terminology from experimental design in regression models, such an optimal censoring plan is referred to as A-optimal.

(iii) In a similar fashion, we may consider the determinant of $\Sigma_*(\mathscr{R})$, i.e.,

$$\phi(\Sigma_*(\mathscr{R})) = \det(\Sigma_*(\mathscr{R})).$$

Such an optimal censoring plan is referred to as D-optimal.

These criteria have a monotonicity property with regard to the Löwner ordering. Namely, if a censoring scheme $\mathscr{R}^*$ exists such that $\Sigma_*(\mathscr{R}^*) \preceq_{\mathsf{L}} \Sigma_*(\mathscr{R})$ for all $\mathscr{R} \in \mathscr{C}^m_{m,n}$, then $\mathscr{R}^*$ is optimal for any of these objective functions. Further criteria will be defined later. Notice that the results will be numerical in nature except for generalized Pareto-distributions.

26.3.1 Exponential Distribution

For the exponential distribution, the matrix $\Sigma_*(\mathscr{R})$ is given by (see (11.7))

$$\Sigma_*(\mathscr{R}) = \frac{1}{n^2(m-1)} \begin{pmatrix} m & -n \\ -n & n^2 \end{pmatrix}.$$

Obviously, this matrix does not depend on the censoring scheme employed. Therefore, any censoring plan yields the same value of the objective function and, thus, is optimal in the above sense. The same comment applies to the variances of the BLUEs in the location and scale cases, respectively.

26.3.2 Generalized Pareto Distributions

Burkschat et al. [235, 237] have addressed a location–scale family of generalized Pareto distributions which includes exponential, Pareto, and reflected power distributions (see Definition A.1.11). Since the exponential distribution has just been discussed, it is excluded from the following presentation of the material. Hence, we have to deal with the location–scale family

$$\mathscr{F}_q = \left\{ F\left(\frac{\cdot - \mu}{\vartheta}\right) \,\middle|\, \overline{F}(t) = (1-\operatorname{sgn}(q)t)^{1/q}, t \geq 0, 1 > \operatorname{sgn}(q)t, \mu \in \mathbb{R}, \vartheta > 0 \right\},$$

where $\operatorname{sgn}(q)$ denotes the sign of $q \neq 0$. In this case, the matrix of interest is given by (see Theorem 11.2.4)

$$\Sigma_*(\mathscr{R}) = \frac{q^2}{\Psi(n+2q)n-(n+2q)^2} \begin{pmatrix} \Psi & -\operatorname{sgn}(q)(\Psi + \frac{n+2q}{q}) \\ -\operatorname{sgn}(q)(\Psi + \frac{n+2q}{q}) & \Psi + \frac{(n+2q)^2}{q^2} \end{pmatrix},$$

where $\Psi = \sum_{j=1}^{m} \prod_{k=1}^{j} \left(1 + \frac{2q}{\gamma_k}\right)$. It is obvious that $\Sigma_*(\mathscr{R})$ depends on $\mathscr{R}$ only via the quantity Ψ so that we can write $\Sigma_*(\mathscr{R}) = S(\Psi)$. Now, Burkschat et al. [237] obtained the following ordering result which relates the Löwner ordering on the set $\mathcal{S}_{m,n}$ to the ordering of Ψ.

Lemma 26.3.2. Let $\Psi_1, \Psi_2 > \frac{n+2q}{n}$ be real numbers. Then,

$$\Psi_1 \geq \Psi_2 \iff S(\Psi_1) \preceq_{\mathsf{L}} S(\Psi_2).$$

Notice that $\boldsymbol{\gamma}(\mathscr{R}) \geq \boldsymbol{\gamma}(\mathscr{S})$ implies that

(i) $\Psi(\mathscr{R}) \leq \Psi(\mathscr{S})$, if $q > 0$, and
(ii) $\Psi(\mathscr{R}) \geq \Psi(\mathscr{S})$, if $q < 0$.

In particular, we conclude that $\mathscr{R} \preccurlyeq \mathscr{S}$ yields $\Psi(\mathscr{R}) \leq \Psi(\mathscr{S})$ for $q > 0$ and $\Psi(\mathscr{R}) \geq \Psi(\mathscr{S})$ for $q < 0$, respectively. Combining these findings, we get $\mathscr{R} \preccurlyeq \mathscr{S}$ which implies $\Sigma_*(\mathscr{R}) \preceq_{\mathsf{L}} \Sigma_*(\mathscr{S})$, $q > 0$, or $\Sigma_*(\mathscr{S}) \preceq_{\mathsf{L}} \Sigma_*(\mathscr{R})$, $q < 0$. Recalling (26.2), this yields the following result due to Burkschat et al. [237].

Theorem 26.3.3. For any objective function ϕ increasing w.r.t. $\preceq_{\mathsf{L}}$ on $\mathcal{S}_{m,n}$, the following assertions hold:

(i) If $q > 0$, then $\mathscr{O}_1$ is ϕ-optimal. This case includes the uniform distribution for $q = 1$;
(ii) If $q < 0$, i.e., for Pareto distributions, then $\mathscr{O}_m$ is ϕ-optimal provided that $n - m + 1 + 2q > 0$.

Moreover, for $q > 0$, the worst scheme is given by $\mathscr{O}_1$, whereas $\mathscr{O}_m$ is worst for $q < 0$.

Notice that the additional condition for Pareto distributions ensures the existence of involved moments. In particular, it guarantees the existence of the BLUEs for any censoring scheme $\mathscr{R} \in \mathscr{C}_{m,n}^{m}$.

Burkschat et al. [237] applied the so-called φ_p-criteria or matrix means used in experimental design in regression models to the optimization problem (cf. Pukelsheim [733, Chap. 6]). Let $T\Lambda T'$ be the spectral decomposition of $\Sigma_*(\mathscr{R})$ and let $\Sigma_*(\mathscr{R})^p = T\Lambda^p T'$ with $\Lambda^p = \mathrm{diag}(\lambda_{\max}^p, \lambda_{\min}^p)$. $\lambda_{\max}, \lambda_{\min} > 0$ denote the maximum and minimum eigenvalues of $\Sigma_*(\mathscr{R})$, respectively. Then, φ_p is defined as

$$\varphi_p(\Sigma_*(\mathscr{R})) = \left(\frac{1}{2}\,\mathrm{tr}(\Sigma_*(\mathscr{R})^p)\right)^{1/p}, \quad p \neq 0.$$

Taking the limits $p \to 0$ and $p \to \pm\infty$, we get the following supplementary criteria:

$$\varphi_0(\Sigma_*(\mathscr{R})) = \lim_{p\to 0} \varphi_p(\Sigma_*(\mathscr{R})) = \sqrt{\det \Sigma_*(\mathscr{R})},$$

$$\varphi_{\infty}(\Sigma_*(\mathscr{R})) = \lim_{p\to\infty} \varphi_p(\Sigma_*(\mathscr{R})) = \lambda_{\max}(\Sigma_*(\mathscr{R})),$$

$$\varphi_{-\infty}(\Sigma_*(\mathscr{R})) = \lim_{p\to-\infty} \varphi_p(\Sigma_*(\mathscr{R})) = \lambda_{\min}(\Sigma_*(\mathscr{R})).$$

Remark 26.3.4. Using the spectral decomposition, it is easy to see that φ_p is a function of the eigenvalues only. Thus, we write $\varphi_p(\lambda_{\max}, \lambda_{\min})$ instead and get the identity

$$\varphi_p(\lambda_{\max}, \lambda_{\min}) = \left(\frac{\lambda_{\max}^p + \lambda_{\min}^p}{2}\right)^{1/p}, \quad p \neq 0.$$

For $p = 0$, we get

$$\varphi_0(\lambda_{\max}, \lambda_{\min}) = \sqrt{\lambda_{\max}\lambda_{\min}}.$$

Obviously, the proposed criteria can be interpreted as means of the eigenvalues. For $p = 1$, we get the arithmetic mean which corresponds to the trace of $\Sigma_*(\mathscr{R})$. Hence, optimization w.r.t. φ_1 yields A-optimal censoring schemes. For $p = 0$, we have the geometric mean. This objective function leads to the same optimal censoring plans as the determinant of $\Sigma_*(\mathscr{R})$, and we thus arrive at D-optimal designs. Finally, $p = -1$ corresponds to the harmonic mean of the eigenvalues.

Now, taking the Löwner ordering established in Theorem 26.3.3 into account and using the fact that the φ_p-criteria are isotonic w.r.t. the Löwner ordering (see Pukelsheim [733, Sect. 5.5]), we arrive at the following result obtained by Burkschat [228]. It should be noted that the proof in the original paper proceeds by showing the monotonicity directly without using monotonicity properties of φ_p on $\mathcal{S}_{m,n}$ (see also Burkschat et al. [235]).

Theorem 26.3.5. Let $p \in [-\infty, \infty]$. Then,

(i) If $q > 0$, then $\mathscr{O}_1$ is φ_p-optimal. This case includes the uniform distribution for $q = 1$;
(ii) If $q < 0$, i.e., for Pareto distributions, then $\mathscr{O}_m$ is φ_p-optimal provided that $n - m + 1 + 2q > 0$.

Example 26.3.6. Burkschat et al. [237] have considered the censoring plans given in Table 26.1. They have computed the trace and determinant of the variance–covariance matrix for uniform, reflected power ($q = 2$), and Pareto distributions ($q = -\frac{1}{3}$). The plots given in Figs. 26.1–26.3 illustrate the characteristic behavior of the repeated measurements for increasing sample size n.

Figures 26.1–26.3 indicate that the choice of the employed censoring plan may highly influence the value of the objective function. The gain in efficiency of the optimal scheme w.r.t. other schemes may be significant. Suppose a reflected

Symbol	Censoring scheme
△	$\mathscr{O}_{10}$
○	$(\frac{n-10}{2}, 0^{*8}, \frac{n-10}{2})$
□	$(\frac{n-10}{10}^{*10})$
◇	$(0^{*4}, \frac{n-10}{2}, \frac{n-10}{2}, 0^{*4})$
▽	$\mathscr{O}_1$

Table 26.1 Symbols and censoring schemes used in the plots presented in Figs. 26.1–26.3 with $m = 10, n \geq 10$

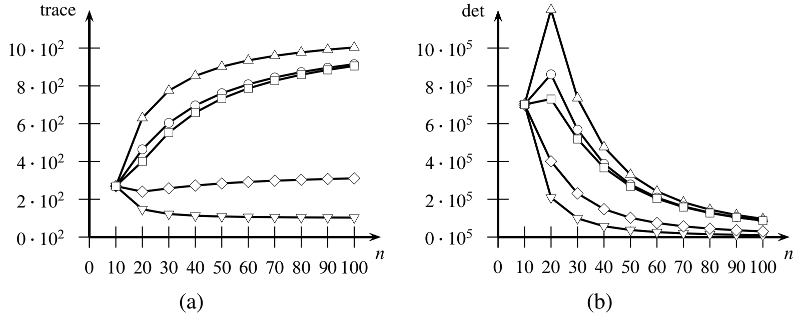

Fig. 26.1 Trace and determinant for censoring schemes given in Table 26.1 for uniform distribution ($q = 1$) . (**a**) Trace of censoring schemes. (**b**) Determinant of censoring schemes

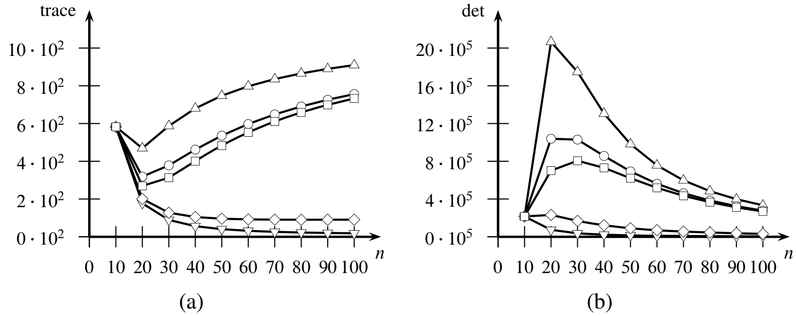

Fig. 26.2 Trace and determinant for censoring schemes given in Table 26.1 for reflected power distribution. ($q = 2$) (**a**) Trace of censoring schemes. (**b**) Determinant of censoring schemes

power distribution, i.e., $q > 0$, is given. Figure 26.2 shows that trace and determinant of the optimal censoring scheme $\mathscr{O}_1$ are decreasing in n. However, this is not necessarily true for other censoring plans employed. It is even worse as shown in Fig. 26.2a: The trace may increase if an unfavorable censoring plan is chosen. Thus, increasing the sample size n may not result in a better precision.

For Pareto distributions, the plots are decreasing in n for any considered censoring scheme. Thus, a higher sample size n improves the precision of the estimators. However, considering the trace as an optimality criterion, choosing

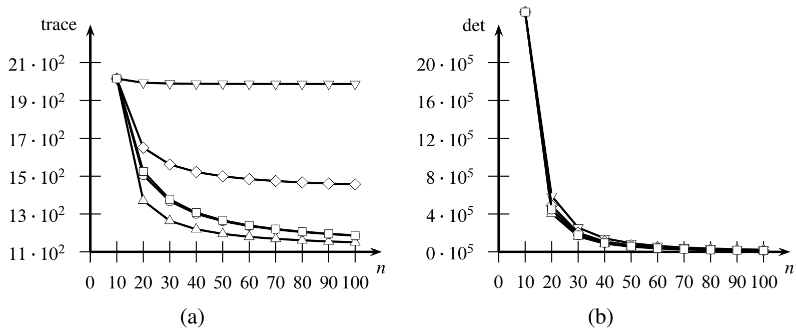

Fig. 26.3 Trace and determinant for censoring schemes given in Table 26.1 for Pareto distribution. ($q = -\frac{1}{3}$) (**a**) Trace of censoring schemes. (**b**) Determinant of censoring schemes

the optimal plan $\mathcal{O}_m$ yields a gain in precision of more than 70 % in comparison to the worst censoring scheme $\mathcal{O}_1$. Further plots with $q \in \{\frac{1}{3}, 4\}$ are presented in Burkschat et al. [237].

Remark 26.3.7. Results similar to those presented above have been established by Burkschat [227] using BLEEs of the unknown parameters. In Theorem 2.1, he proved a result similar to Lemma 26.3.2 for the mean squared error matrix. With this result, the same machinery as above can be applied. Finally, Theorems 26.3.3 and 26.3.5 hold for BLEEs, too. Additionally, uniqueness of the optimal censoring schemes is proved. Moreover, a multi-sample situation is discussed.

Remark 26.3.8. In the one-parameter settings, the same results can be established depending on the sign of the parameter q (see Burkschat et al. [235]).

26.3.3 Extreme Value Distribution

Optimal censoring schemes for the extreme value distribution (Type I) are reprinted in Table 26.2. Results for the extreme value distribution (Type II) are presented in Table 26.3. Similar tables are given in Balakrishnan and Aggarwala [86, pp. 200–210].

26.3.4 Further Distributions

Ng et al. [689] discussed A- and D-optimal censoring schemes for the Weibull distribution. Computational results on optimal censoring plans are provided. Tables for normal and log-normal distributions are presented in Balakrishnan and Aggarwala [86].

n / m	2	3	4	5	6	Criterion
15	13,0	12,0,0	11,0,0,0	0,10,0,0,0	0,9,0,0,0,0	D (1st)
	0.208880	0.070079	0.037306	0.023838	0.016252	
	12,1	11,1,0	10,1,0,0	10,0,0,0,0	1,8,0,0,0,0	D (2nd)
	0.270716	0.076208	0.038309	0.023859	0.016300	
	0.061836	0.006128	0.001003	0.000021	0.000048	Δ
	13,0	12,0,0	11,0,0,0	10,0,0,0,0	9,0,0,0,0,0	A (1st)
	1.658801	0.683595	0.443825	0.338027	0.277645	
	12,1	11,1,0	10,1,0,0	9,1,0,0,0	8,1,0,0,0,0	A (2nd)
	2.234228	0.751833	0.458666	0.341474	0.277809	
	0.575427	0.068237	0.014841	0.003447	0.000163	Δ
20	18,0	17,0,0	16,0,0,0	0,15,0,0,0	0,14,0,0,0,0	D (1st)
	0.188199	0.062780	0.033494	0.021174	0.014388	
	17,1	16,1,0	15,1,0,0	1,14,0,0,0	1,13,0,0,0,0	D (2nd)
	0.242727	0.068257	0.034400	0.021208	0.014426	
	0.054528	0.005477	0.000906	0.000034	0.000037	Δ
	18,0	17,0,0	16,0,0,0	15,0,0,0,0	14,0,0,0,0,0	A (1st)
	1.718838	0.688633	0.440165	0.331886	0.270708	
	17,1	16,1,0	15,1,0,0	14,1,0,0,0	13,1,0,0,0,0	A (2nd)
	2.309064	0.760167	0.456499	0.336229	0.271469	
	0.590226	0.071534	0.016334	0.004343	0.000761	Δ
25	23,0	22,0,0	21,0,0,0	0,20,0,0,0	0,19,0,0,0,0	D (1st)
	0.173793	0.057758	0.030885	0.019333	0.013115	
	22,1	21,1,0	20,1,0,0	1,19,0,0,0	1,18,0,0,0,0	D (2nd)
	0.223363	0.062796	0.031732	0.019364	0.013145	
	0.049570	0.005038	0.000846	0.000030	0.000029	Δ
	23,0	22,0,0	21,0,0,0	20,0,0,0,0	19,0,0,0,0,0	A (1st)
	1.760666	0.692381	0.438124	0.328188	0.266467	
	22,1	21,1,0	20,1,0,0	19,1,0,0,0	18,1,0,0,0,0	A (2nd)
	2.358473	0.765834	0.455397	0.333120	0.267633	
	0.597807	0.073454	0.017273	0.004932	0.001166	Δ
30	28,0	27,0,0	26,0,0,0	0,25,0,0,0	0,24,0,0,0,0	D (1st)
	0.162972	0.054020	0.028950	0.017963	0.012175	
	27,1	26,1,0	25,1,0,0	1,24,0,0,0	1,23,0,0,0,0	D (2nd)
	0.208880	0.058735	0.029754	0.017989	0.012199	
	0.045908	0.004715	0.000804	0.000027	0.000024	Δ
	28,0	27,0,0	26,0,0,0	25,0,0,0,0	24,0,0,0,0,0	A (1st)
	1.791522	0.695133	0.436760	0.325642	0.263531	
	27,1	26,1,0	25,1,0,0	24,1,0,0,0	23,1,0,0,0,0	A (2nd)
	2.393159	0.769775	0.454670	0.330992	0.264993	
	0.601637	0.074641	0.017911	0.005350	0.001462	Δ

Table 26.2 A- and D-optimal censoring plans for selected values of m and n for extreme value distribution (Type I). The values in small font denote the values of the criteria for the particular censoring scheme. The row Δ contains the difference between the best and second best censoring schemes

n / m	2	3	4	5	6	Criterion
15	11,2	9,0,3	6,0,0,5	0,5,0,0,5	0,3,0,0,0,6	D (1st)
	0.054830	0.025563	0.016220	0.011619	0.008976	
	12,1	8,0,4	5,1,0,5	1,4,0,0,5	0,2,1,0,0,6	D (2nd)
	0.055175	0.025601	0.016228	0.011630	0.008980	
	0.000345	0.000037	0.000007	0.000011	0.000004	Δ
	12,1	9,0,3	7,0,0,4	5,0,0,0,5	3,0,0,0,0,6	A (1st)
	1.005610	0.511761	0.353150	0.275009	0.228228	
	11,2	10,0,2	8,0,0,3	6,0,0,0,4	3,0,0,0,1,5	A (2nd)
	1.008464	0.512453	0.353706	0.275378	0.228463	
	0.002854	0.000692	0.000555	0.000369	0.000234	Δ
20	16,2	13,0,4	9,2,0,5	0,10,0,0,5	2,5,0,0,0,7	D (1st)
	0.043318	0.020091	0.012710	0.009044	0.006974	
	15,3	14,0,3	10,0,0,6	0,9,0,0,6	0,2,0,5,0,7	D (2nd)
	0.043747	0.020141	0.012710	0.009047	0.007012	
	0.000429	0.000050	0.000000	0.000003	0.000038	Δ
	16,2	14,0,3	12,0,0,4	10,0,0,0,5	2,5,0,0,0,7	A (1st)
	1.012956	0.506423	0.345463	0.266433	0.223349	
	17,1	15,0,2	11,0,0,5	9,0,0,0,6	5,2,0,1,4,2	A (2nd)
	1.015150	0.508953	0.346811	0.266986	0.223717	
	0.002194	0.002530	0.001348	0.000553	0.000368	Δ

Table 26.3 A- and D-optimal censoring plans for selected values of m and n for extreme value distribution (Type II). The values in small font denote the values of the criteria for the particular censoring scheme. The row Δ contains the difference between the best and second best censoring schemes

26.4 Maximum Fisher Information

Fisher information as a criterion for optimal progressive censoring has been introduced in Ng et al. [689] to identify optimal censoring schemes for a two-parameter Weibull distribution (see Sect. 26.4.2). They handled the problem computationally by using the missing information principle (see Sect. 9.1.2). Balakrishnan et al. [140] picked up the idea of optimal Fisher information by exploiting the hazard rate representation of Fisher information in progressively Type-II censored samples. This approach leads to some explicit results and also provides a computationally more efficient approach to address the problem. First, we present results for the single parameter case.

26.4.1 Single Parameter Case

Let $\mathcal{F} = \{F_\theta \mid \theta \in \Theta \subseteq \mathbb{R}\}$. To begin, it is worth mentioning that if $\mathcal{F}$ forms an exponential family as in (2.12), i.e., $F_\theta(x) = 1 - \mathrm{e}^{-\eta(\theta)d(x)}$, then, according to (9.11), the Fisher information about θ is given by

$$\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \frac{m}{\eta^2(\theta)}[\eta'(\theta)]^2$$

and is therefore independent of the censoring plan. Therefore, all censoring schemes yield the same information about the parameter and, thus, are optimal w.r.t. Fisher information. Examples of such distributions are, for instance, exponential distribution (scale parameter), extreme value distribution (location parameter), Weibull distribution (scale parameter), and Pareto distribution (shape parameter).

Using representation (9.1) obtained by Zheng and Park [943] and the results for Fisher information in a location or scale family, Balakrishnan et al. [140] showed that the optimal censoring plan can be calculated w.r.t. the standard member of the population distribution. In particular, they found the following result. The proof proceeds by using the partial order $\preccurlyeq$ on $\mathscr{C}_{m,n}^m$ introduced in Definition 7.1.9 and some stochastic ordering result obtained by Burkschat [228].

Theorem 26.4.1. Let w_θ be defined by

$$x \mapsto \left\{\frac{\partial}{\partial\theta}\log\lambda_\theta(x)\right\}^2. \tag{26.7}$$

Then,

(i) Right censoring $\mathscr{O}_m = (0^{*m-1}, n-m)$ has maximum Fisher information if w_θ is decreasing in θ;
(ii) First-step censoring $\mathscr{O}_1 = (n-m, 0^{*m-1})$ has maximum Fisher information if w_θ is increasing in θ.

Applying this result, Balakrishnan et al. [140] established that w_θ is decreasing for location families with log-concave hazard rate function. This applies to location families based on either a normal or a logistic distribution. Cramer and Ensenbach [293] showed that the same argument applies to a location family of Gumbel distributions.

For scale families, an analogous criterion can be formulated by considering the function w defined as

$$w(t) = t\lambda'(t)/\lambda(t), \quad t > 0, \tag{26.8}$$

where λ denotes the hazard rate of the standard member in $\mathcal{F}$. In particular, a decreasing w yields right censoring to be optimal. An example for a distribution satisfying this condition is given by a log-concave hazard rate function λ introduced by Block and Joe [207] (see also Pellerey et al. [715]). It is defined by

$$\lambda(t) = \frac{t}{1+t} + \log(1+t), \quad t \geq 0.$$

Notice that the corresponding cumulative distribution function F is given by $F(t) = 1-(1+t)^t$, $t \geq 0$, showing that $\mathcal{F}$ does not form an exponential family.

For an increasing function w given in (26.8), first-step censoring is optimal. A sufficient but not necessary condition for w to be increasing is log-convexity of the hazard rate λ. Examples for such distributions are given by a scale family of truncated extreme value distributions with cumulative distribution function F given by $F(t) = 1-\exp\{1-e^t\}$, $t \geq 0$, and a family defined by the cumulative distribution function F with $F(t) = 1-\exp\{-t^3/6-t^2/2-t\}$, $t \geq 0$.

Although many distributions satisfy the above criteria, some do not. For instance, scale families of extreme value and normal distributions (or shape families of Weibull distributions) do not satisfy the aforementioned conditions. In these cases, almost only computational results are available.

Balakrishnan et al. [140] found the following expression for the Fisher information in a scale parameter of an extreme value distribution given by $F_\theta(x) = 1-\exp\{-e^{\theta x}\}$, $\theta > 0$. They used the expression

$$\theta^2 \mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta) = \frac{m\pi^2}{6} + \sum_{i=1}^{m}(\gamma_i - \gamma_{i+1})c_{i-1}\sum_{j=1}^{i}\frac{a_{j,i}}{\gamma_j^2}(\gamma - 1 + \log\gamma_j)^2$$

to compute the Fisher information for any censoring scheme $\mathscr{R}$. Alternatively, (9.12) may be used. The computations suggest that one-step plans are optimal. Moreover, as indicated in Table 26.4, optimality moves quite regularly from first-step censoring $\mathscr{O}_1$ to right censoring $\mathscr{O}_m$. A similar behavior is observed for a scale family of normal and logistic distributions, respectively (see Tables 26.5 and 26.6).

These observations lead to the so-called one-step conjecture meaning that in many settings one-step censoring schemes are optimal plans. From a heuristic point of view, this seems to be quite reasonable because the convex hull of $\mathscr{C}_{m,n}^m$ forms a simplex and, thus, is a convex set. Moreover, its vertices are admissible and given by the one-step plans (see Sect. 1.1.1 and Fig. 1.3). If the objective function has some nice properties, e.g., convexity, the maximum is attained at the border of the convex hull so that the one-step plans are reasonable candidates for maximum Fisher information. As far as we know, there is only a proof for the extreme value distribution with $m = 2$ (see Balakrishnan et al. [140, Theorem 4.3]).

Optimal One-Step Plans

Due to the relevance of one-step plans in the above optimization problem, Balakrishnan et al. [140] established simple formulas to determine the Fisher information in a progressively Type-II censored sample for such censoring schemes. Following Park [704], the Fisher information in the complete sample $X_1, \ldots, X_n$ can be split into the part contained in the first k order statistics $X_{1:n}, \ldots, X_{k:n}$ and a remaining part as

	n		Fisher information (in units of $1/\theta^2$)			
			OSP		RC	
m	$n_1, \dots, n_2$	OSP	n_1	n_2	n_1	n_2
2	$3, \dots, 10$	$\mathscr{O}_1$	3.8	6.6	3.1	6.5
2	$11, \dots, 1000$	$\mathscr{O}_2$	7.1	74.4	7.1	74.4
3	$4, \dots, 10$	$\mathscr{O}_1$	5.7	8.1	4.5	7.0
3	$11, \dots, 19$	$\mathscr{O}_2$	8.5	12.3	7.6	12.2
3	$20, \dots, 300$	$\mathscr{O}_3$	12.7	63.2	12.7	63.2
4	$5, \dots, 10$	$\mathscr{O}_1$	7.6	9.5	6.0	7.3
4	$11, \dots, 19$	$\mathscr{O}_2$	9.9	13.7	7.9	12.9
4	$20, \dots, 28$	$\mathscr{O}_3$	14.1	18.0	13.5	17.9
4	$29, \dots, 100$	$\mathscr{O}_4$	18.5	43.1	18.5	43.1
5	$6, \dots, 10$	$\mathscr{O}_1$	9.5	11.0	7.5	7.7
5	$11, \dots, 19$	$\mathscr{O}_2$	11.4	15.1	8.2	13.2
5	$20, \dots, 28$	$\mathscr{O}_3$	15.5	19.4	13.8	18.7
5	$29, \dots, 37$	$\mathscr{O}_4$	19.8	23.7	19.3	23.7
5	$38, \dots, 50$	$\mathscr{O}_5$	24.2	29.9	24.2	29.9
6	$7, \dots, 10$	$\mathscr{O}_1$	11.3	12.4	9.1	8.5
6	$11, \dots, 19$	$\mathscr{O}_2$	12.8	16.5	8.8	13.4
6	$20, \dots, 28$	$\mathscr{O}_3$	16.9	20.7	14.0	19.2
6	$29, \dots, 35$	$\mathscr{O}_4$	21.2	24.1	19.8	23.4

Table 26.4 Optimal (max. Fisher information) progressive censoring plans for certain (small) m and n with respect to the scale parameter of an extreme value distribution (shape parameter of a Weibull distribution). The Fisher informations are given for the optimal one-step plans (OSP) and right censored samples (RC) (for both the minimum value n_1 and the maximum value n_2 of n in each row). The values are taken from Balakrishnan et al. [140]

$$\begin{aligned} n\mathscr{I}(X_1;\theta) &= \mathscr{I}(X_1,\dots,X_n;\theta) \\ &= \mathscr{I}(X_{1:n},\dots,X_{k:n};\theta) + \mathscr{I}(X_{k+1:n},\dots,X_{n:n}|X_{k:n};\theta). \end{aligned}$$

For a progressively censored sample with censoring scheme $\mathscr{O}_k$, this yields the identity

$$\mathscr{I}(\mathbf{X}^{\mathscr{O}_k};\theta) = \frac{1}{n-k}\left\{n(m-k)\mathscr{I}(X_1;\theta) + (n-m)\mathscr{I}(X_{1:n},\dots,X_{k:n};\theta)\right\}. \tag{26.9}$$

Notice that $\mathscr{I}(\mathbf{X}^{\mathscr{O}_k};\theta)$ can be written as a convex combination of the Fisher information in the complete sample $n\mathscr{I}(X_1;\theta)$ and the Fisher information in the first k order statistics. Hence, we get the bounds

$$\mathscr{I}(X_{1:n},\dots,X_{k:n};\theta) \le \mathscr{I}(\mathbf{X}^{\mathscr{O}_k};\theta) \le n\mathscr{I}(X_1;\theta).$$

	n		Fisher information (in units of $1/\theta^2$)			
			OSP		RC	
m	$n_1,\ldots,n_2$	OSP	n_1	n_2	n_1	n_2
2	$3,\ldots,11$	$\mathscr{O}_1$	4.4	9.4	4.0	9.3
2	$12,\ldots,200$	$\mathscr{O}_2$	10.1	77.3	10.1	77.3
3	$4,\ldots,11$	$\mathscr{O}_1$	6.5	10.8	5.8	9.8
3	$12,\ldots,22$	$\mathscr{O}_2$	11.5	18.5	10.6	18.5
3	$23,\ldots,200$	$\mathscr{O}_3$	19.2	95.3	19.2	95.3
4	$5,\ldots,11$	$\mathscr{O}_1$	8.5	12.2	7.6	10.3
4	$12,\ldots,22$	$\mathscr{O}_2$	12.9	19.9	11.0	19.0
4	$23,\ldots,34$	$\mathscr{O}_3$	20.6	28.5	19.8	28.4
4	$35,\ldots,100$	$\mathscr{O}_4$	29.2	68.5	29.2	68.5
5	$6,\ldots,11$	$\mathscr{O}_1$	10.6	13.6	9.5	11.0
5	$12,\ldots,22$	$\mathscr{O}_2$	14.3	21.2	11.6	19.3
5	$23,\ldots,34$	$\mathscr{O}_3$	21.9	29.8	20.1	29.1
5	$35,\ldots,46$	$\mathscr{O}_4$	30.5	38.4	29.9	38.4
5	$47,\ldots,50$	$\mathscr{O}_5$	39.2	41.4	39.2	41.4
6	$7,\ldots,11$	$\mathscr{O}_1$	12.6	15.0	11.5	12.0
6	$12,\ldots,22$	$\mathscr{O}_2$	15.7	22.6	12.5	19.5
6	$23,\ldots,34$	$\mathscr{O}_3$	23.2	31.1	20.3	29.4
6	$35,\ldots,40$	$\mathscr{O}_4$	31.8	35.5	30.3	34.4

Table 26.5 Optimal (max. Fisher information) progressive censoring plans for certain (small) m and n with respect to the scale parameter of a normal distribution. The Fisher informations are given for the optimal one-step plans (OSP) and right censored samples (RC) (for both the minimum value n_1 and the maximum value n_2 of n in each row). The values are taken from Balakrishnan et al. [140]

Finally, we get from (26.9)

$$\frac{m}{n-m}\mathscr{I}(X_1;\theta)\left(\frac{\mathscr{I}(\mathbf{X}^{\mathscr{O}_k};\theta)}{m\mathscr{I}(X_1;\theta)}-1\right)=\frac{\mathscr{I}(X_{1:n},\ldots,X_{k:n};\theta)-k\mathscr{I}(X_1;\theta)}{n-k}. \tag{26.10}$$

The right-hand side of (26.10) depends on $\mathscr{O}_k$ only via the number $k\in\{1,\ldots,m\}$. Thus, the Fisher information $\mathscr{I}(\mathbf{X}^{\mathscr{O}_k};\theta)$ is maximized for that k that maximizes the ratio on the right. Moreover, except for the condition $1\le k\le m$, the right-hand side is independent of m. Therefore, given $m_0\in\mathbb{N}$ and a number $k_{\text{opt}}(m_0)<m_0$ such that k_{opt} is optimal for that particular m_0, then k_{opt} is optimal for any $m\ge m_0$. In order to find this k, we have to compute the values of

$$\begin{aligned}\varrho\left(\frac{k}{n},n;\theta\right)&=\frac{\mathscr{I}(X_{1:n},\ldots,X_{k:n};\theta)}{(n-k)\mathscr{I}(X_1;\theta)}-\frac{k}{n-k}\\&=\frac{n}{n-k}V_k(\theta)-\frac{k}{n-k},\quad 1\le k\le m,\end{aligned} \tag{26.11}$$

			Fisher information (in units of $1/\theta^2$)			
	n		OSP		RC	
m	$n_1, \ldots, n_2$	OSP	n_1	n_2	n_1	n_2
2	$3, \ldots, 10$	$\mathscr{O}_1$	3.1	5.8	3.0	5.8
2	$11, \ldots, 1000$	$\mathscr{O}_2$	6.2	74.2	6.2	74.2
3	$4, \ldots, 10$	$\mathscr{O}_1$	4.6	6.8	4.3	6.2
3	$11, \ldots, 20$	$\mathscr{O}_2$	7.3	11.3	6.7	11.3
3	$21, \ldots, 300$	$\mathscr{O}_3$	11.8	62.4	11.8	62.4
4	$5, \ldots, 10$	$\mathscr{O}_1$	6.1	7.9	5.7	6.8
4	$11, \ldots, 20$	$\mathscr{O}_2$	8.4	12.3	7.2	11.4
4	$21, \ldots, 30$	$\mathscr{O}_3$	12.8	16.9	12.3	16.9
4	$31, \ldots, 100$	$\mathscr{O}_4$	17.4	41.3	17.4	41.3
5	$6, \ldots, 10$	$\mathscr{O}_1$	7.5	9.0	7.2	7.6
5	$11, \ldots, 20$	$\mathscr{O}_2$	9.4	13.3	7.8	12.0
5	$21, \ldots, 30$	$\mathscr{O}_3$	13.8	17.9	12.6	17.4
5	$31, \ldots, 40$	$\mathscr{O}_4$	18.4	22.5	17.9	22.5
5	$41, \ldots, 50$	$\mathscr{O}_5$	23.0	27.2	23.0	27.2
6	$7, \ldots, 10$	$\mathscr{O}_1$	8.9	10.0	8.6	8.6
6	$11, \ldots, 20$	$\mathscr{O}_2$	10.5	14.3	8.8	12.3
6	$21, \ldots, 30$	$\mathscr{O}_3$	14.8	18.9	12.8	17.7
6	$31, \ldots, 35$	$\mathscr{O}_4$	19.4	21.3	18.3	20.4

Table 26.6 Optimal (max. Fisher information) progressive censoring plans for certain (small) m and n with respect to the scale parameter of a logistic distribution. The Fisher informations are given for the optimal one-step plans (OSP) and right censored samples (RC) (for both the minimum value n_1 and the maximum value n_2 of n in each row). The values are taken from Balakrishnan et al. [140]

where $V_k(\theta) = \frac{\mathscr{I}(X_{1:n},\ldots,X_{k:n};\theta)}{n\mathscr{I}(X_1;\theta)}$. From expression (26.11), we deduce that the knowledge of $V_k(\theta)$ is sufficient to decide which censoring scheme is optimal. Moreover, we need to compute only the values of $V_k(\theta)$ until $V_k(\theta)$ decreases. Expression (26.11) shows that the objective function $\varrho\left(\frac{k}{n}, n; \theta\right)$ is connected to a relative version of the Fisher information $V_k(\theta)$. Based on (26.11) and following Park [704], Balakrishnan et al. [140] proposed information plots as well as a plot of the critical function. For illustration, we present these plots for the scale family of an extreme value distribution in Figs. 26.4 and 26.5. Notice that the quantities are independent of the scale parameter by Theorem 9.1.3. Similar plots for scale families of normal and logistic distributions can be found in Balakrishnan et al. [140].

Remark 26.4.2. The determination of optimal one-step censoring plans has been discussed in terms of the missing information principle in Park and Ng [708].

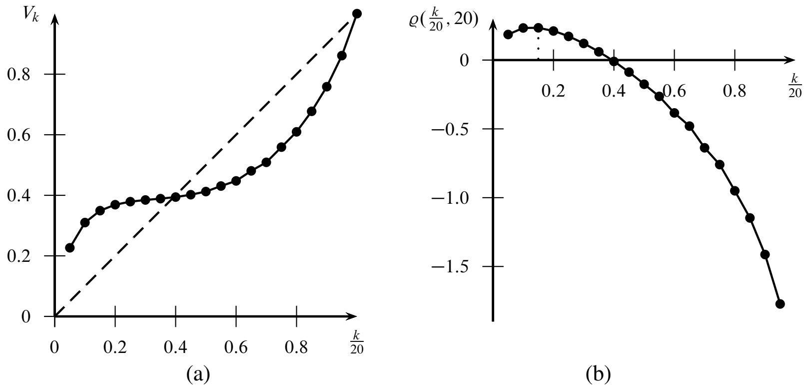

Fig. 26.4 Information plot (**a**) and plot of objective function (**b**) for $n = 20$. The optimal one-step censoring scheme is $\mathcal{O}_3$ with maximum 0.23445

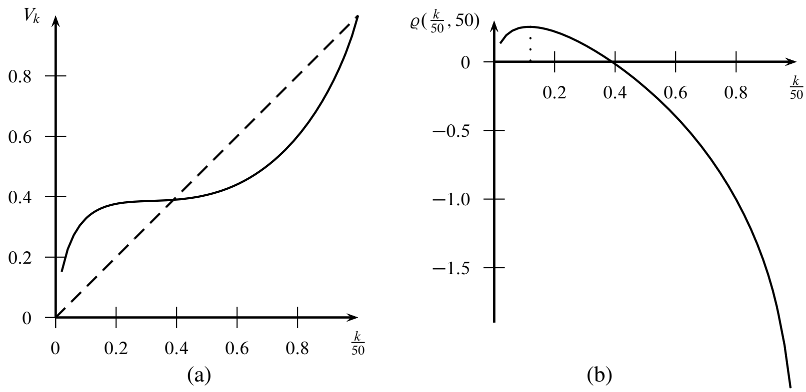

Fig. 26.5 Information plot (**a**) and plot of objective function (**b**) for $n = 50$. The optimal one-step censoring scheme is $\mathcal{O}_6$ with maximum 0.25496

State of the Art

In Table 26.7, we survey the recent results on maximum Fisher information plans for some selected lifetime distributions. It should be noted that the results on OSPs are conjectures. The results on optimal right censoring and invariance of the Fisher information are proven.

Further, it needs to be mentioned that in the framework of maximum Fisher information, other censoring scheme than those given in Table 26.7 may be optimal. As an example, we present results for the location parameter of a Laplace distribution with $m = 10$ and various values of n in Table 26.8. Notice that the Fisher information in the sample is increasing until $n = 22$ and then decreasing. A conjecture based on these values is presented in Table 26.9.

Distribution	Parameter	Optimal
Exponential: $F(x;\theta) = 1 - \mathrm{e}^{-\theta x}$, $\theta > 0$	Scale (θ)	ALL
Extreme value (Type I):	Location (μ)	ALL
$F(x;\mu,\theta) = 1 - \exp(-e^{\theta(x-\mu)})$, $\theta > 0$	Scale (θ)	OSP
Gumbel:	Location (μ)	RC
$F(x;\mu,\theta) = \exp(-e^{-\theta(x-\mu)})$, $\theta > 0$	Scale (θ)	OSP
Normal/Log-normal	Location (μ)	RC
	Scale (σ)	OSP
Weibull:	Scale (θ)	ALL
$F(x;\theta,\beta) = 1 - \exp(-(\theta x)^{\beta})$, $x > 0$, $\theta, \beta > 0$	Shape (β)	OSP
Pareto: $F(x;\alpha,\beta) = 1 - (\alpha/x)^{\beta}$, $x > \alpha$, $\alpha, \beta > 0$	Shape (β)	ALL
Logistic:	Location (μ)	RC
$F(x;\mu,\theta) = 1 - \frac{\exp(-\theta(x-\mu))}{1+\exp(-\theta(x-\mu))}$, $\theta > 0$	Scale (θ)	OSP

Table 26.7 Maximum Fisher information plans for selected distributions. The last column gives the optimal censoring plan w.r.t. the Fisher information criterion (RC $\hat{=}$ right censoring; OSP $\hat{=}$ one-step progressive censoring; ALL $\hat{=}$ Fisher information invariant)

n	Best scheme	Fisher information	Best OSP	Fisher information
11	$(0^{*9}, 1)$	10.998	$(0^{*9}, 1)$	10.998
12	$(0^{*9}, 2)$	11.988	$(0^{*9}, 2)$	11.988
13	$(0^{*9}, 3)$	12.956	$(0^{*9}, 3)$	12.956
14	$(0^{*9}, 4)$	13.878	$(0^{*9}, 4)$	13.878
15	$(0^{*9}, 5)$	14.729	$(0^{*9}, 5)$	14.729
16	$(0^{*9}, 6)$	15.481	$(0^{*9}, 6)$	15.481
17	$(0^{*9}, 7)$	16.113	$(0^{*9}, 7)$	16.113
18	$(0^{*9}, 8)$	16.611	$(0^{*9}, 8)$	16.611
19	$(0^{*9}, 9)$	16.970	$(0^{*9}, 9)$	16.970
20	$(0^{*9}, 10)$	17.196	$(0^{*9}, 10)$	17.196
21	$(0^{*9}, 11)$	17.300	$(0^{*9}, 11)$	17.300
22	$(0^{*9}12)$	17.300	$(0^{*9}, 12)$	17.300
23	$(1, 0^{*8}, 12)$	17.266	$(0^{*9}, 13)$	17.215
29	$(7, 0^{*8}, 12)$	17.105	$(0^{*9}, 19)$	15.885
30	$(9, 0^{*8}, 11)$	17.086	$(0^{*9}, 20)$	15.637
$1,000$	$(979, 0^{*8}, 11)$	16.496	$(0^{*9}, 990)$	10.111
$2,000$	$(1979, 0^{*8}, 11)$	16.486	$(0, 0^{*9}, 1990)$	10.055
$3,000$	$(2979, 0^{*8}, 11)$	16.483	$(0^{*9}0, 2990)$	10.037
$4,000$	$(3979, 0^{*8}, 11)$	16.481	$(0^{*9}, 3990)$	10.028

Table 26.8 Maximum Fisher information plans for Laplace distributions (location) including the value of the Fisher information

m	n	Optimal scheme
10	11–22	$(0^{*9}, n-10)$
	23–29	$(n-22, 0^{*8}, 12)$
	≥ 30	$(n-21, 0^{*8}, 11)$

Table 26.9 Conjecture on maximum Fisher information plans for Laplace distributions (location)

26.4.2 Two-Parameter Case

Optimal Fisher information based censoring plans in a two-parameter setting have been first addressed by Ng et al. [689] who considered Weibull and extreme value distributions. As optimality criteria, they used minimum trace/determinant of the variance–covariance matrix and maximum trace of the Fisher information matrix. Using the terminology as in experimental design of regression models (see, e.g., Pukelsheim [733]), A- and D-optimality of censoring schemes w.r.t. the Fisher information matrix was introduced in Dahmen et al. [320]. The problem has also been addressed by Abo-Eleneen [6] who provided some computational results for small sizes of m and n.

Definition 26.4.3. Let $\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta)$ denote the Fisher information matrix about θ in a progressively Type-II censored sample $\mathbf{X}^{\mathscr{R}}$.

A censoring plan $\mathscr{R} \in \mathscr{C}^n_{m,n}$ is said to be

(i) *A-optimal (w.r.t. Fisher information matrix)* if

$$\operatorname{tr}\Big(\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta)\Big) = \max_{\mathscr{S}\in\mathscr{C}^m_{m,n}} \operatorname{tr}\Big(\mathscr{I}(\mathbf{X}^{\mathscr{S}};\theta)\Big);$$

(ii) *D-optimal (w.r.t. Fisher information matrix)* if

$$\det\Big(\mathscr{I}(\mathbf{X}^{\mathscr{R}};\theta)\Big) = \max_{\mathscr{S}\in\mathscr{C}^m_{m,n}} \det\Big(\mathscr{I}(\mathbf{X}^{\mathscr{S}};\theta)\Big).$$

As in the one-parameter setting, Dahmen et al. [320] established a condition for A-optimality in terms of the hazard rate.

Theorem 26.4.4. Let $\theta = (\theta_1, \theta_2)$ and $\lambda_\theta = f_\theta/(1-F_\theta)$ be the hazard rate of F_θ. If w_θ, defined by

$$w_\theta : x \longmapsto \left\{\frac{\partial}{\partial\theta_1}\log\lambda_\theta(x)\right\}^2 + \left\{\frac{\partial}{\partial\theta_2}\log\lambda_\theta(x)\right\}^2,$$

is decreasing for all $\theta \in \Theta$, then right censoring is A-optimal. If w_θ is increasing for all $\theta \in \Theta$, then first-step censoring is A-optimal.

An example satisfying the above condition is provided by a two-parameter Lomax distribution with cumulative distribution function in (9.14). From the representation $\lambda_{q,\vartheta}(x) = q\vartheta \cdot (1 + \vartheta x)^{-1}$, $x > 0$, it can be seen that $w_{q,\vartheta}$ is decreasing in x for any $q, \vartheta > 0$. Hence, the trace of the Fisher information is maximized by right censoring and $\mathscr{O}_m$ is A-optimal. This is not true for D-optimality. In this case, the expression for the determinant of $\mathscr{I}\left(\mathbf{X}^{\mathscr{R}}; q, \vartheta\right)$ is (see (9.15))

$$\det(\mathscr{I}\left(\mathbf{X}^{\mathscr{R}}; q, \vartheta\right)) = \mathscr{I}_{11} \cdot \mathscr{I}_{22} - (\mathscr{I}_{12})^2 = \frac{1}{q^2\vartheta^2}\left(m \cdot \varphi_2 - \varphi_1^2\right), \qquad (26.12)$$

where

$$\varphi_1 = \sum_{s=1}^{m} \prod_{j=1}^{s} \left(1 - \frac{1}{q\gamma_j + 1}\right) \quad \text{and} \quad \varphi_2 = \sum_{s=1}^{m} \prod_{j=1}^{s} \left(1 - \frac{2}{q\gamma_j + 2}\right).$$

Notice that φ_1 as well as φ_2 are maximal for right censoring. To be more precise, φ_2 is maximal for right censoring, whereas $-\varphi_1$ is maximized for first-step censoring. Thus, both terms in (26.12) are competing. Moreover, both expressions depend on the parameter q so that the optimal scheme will generally depend on q. For details, we refer to Dahmen et al. [320].

Dahmen et al. [320] presented computational results for optimal censoring schemes for two-parameter Lomax, extreme value, Weibull, and normal distributions. As an example, we provide some tables for the extreme value distribution and Weibull distribution in Tables 26.10 and 26.11. Notice that the optimal plans given in Table 26.10 coincide with those presented in Table 26.4 for a scale family of extreme value distributions. This is not surprising because the part of the trace caused by the location parameter, i.e., $\mathscr{I}_{11}$, is constant w.r.t. the censoring scheme. Finally, it has to be mentioned that the two-parameter Lomax distribution provides an example for the situation when optimal censoring plans do not move from first-step censoring to right censoring (see Figs. 26.6 and 26.7). Details can be found in Dahmen et al. [320].

For Lomax distributions, Dahmen et al. [320] also computed the asymptotic Fisher information matrix for one-step plans $\mathscr{O}_j$, $1 \le j \le m$. Notice that the Fisher information in a one-step censoring scheme is given by

$$\mathscr{I}(\mathbf{X}^{\mathscr{O}_j}; q, \vartheta) = \begin{bmatrix} \frac{m}{q^2} & \frac{1}{q\vartheta}\varphi_1(j) \\ \frac{1}{q\vartheta}\varphi_1(j) & \frac{1}{\vartheta^2}\varphi_2(j) \end{bmatrix}.$$

m	n $n_1, \ldots, n_2$	OSP	Fisher information (in units of $1/\theta^2$) OSP n_1	OSP n_2	RC n_1	RC n_2
2	$3, \ldots, 10$	$\mathscr{O}_1$	5.79	8.63	5.05	8.51
	$11, \ldots, 1000$	$\mathscr{O}_2$	9.05	76.42	9.05	76.42
3	$4, \ldots, 10$	$\mathscr{O}_1$	8.72	11.08	7.47	9.98
	$11, \ldots, 19$	$\mathscr{O}_2$	11.50	15.28	10.59	15.21
	$20, \ldots, 300$	$\mathscr{O}_3$	15.74	66.15	15.74	66.15
3	$5, \ldots, 10$	$\mathscr{O}_1$	11.60	13.53	9.98	11.28
	$11, \ldots, 19$	$\mathscr{O}_2$	13.94	17.68	11.87	16.86
	$20, \ldots, 28$	$\mathscr{O}_3$	18.14	21.99	17.46	21.93
	$29, \ldots, 200$	$\mathscr{O}_4$	22.45	66.31	22.45	66.31
5	$6, \ldots, 10$	$\mathscr{O}_1$	14.46	15.98	12.54	12.72
	$11, \ldots, 19$	$\mathscr{O}_2$	16.39	20.08	13.21	18.18
	$20, \ldots, 28$	$\mathscr{O}_3$	20.53	24.37	18.82	23.72
	$29, \ldots, 37$	$\mathscr{O}_4$	24.83	28.71	24.29	28.66
	$38, \ldots, 200$	$\mathscr{O}_5$	29.18	75.27	29.18	75.27
⋮		⋮				
6	$7, \ldots, 10$	$\mathscr{O}_1$	17.30	18.43	15.14	14.46
	$11, \ldots, 19$	$\mathscr{O}_2$	18.83	22.47	14.76	19.37
	$20, \ldots, 28$	$\mathscr{O}_3$	22.93	26.74	20.03	25.18
	$29, \ldots, 37$	$\mathscr{O}_4$	27.21	31.08	25.79	30.53
	$38, \ldots, 46$	$\mathscr{O}_5$	31.54	35.43	31.10	35.39
	$47, \ldots, 200$	$\mathscr{O}_6$	35.90	83.11	35.90	83.11
10	$11, \ldots, 19$	$\mathscr{O}_2$	28.61	32.07	25.80	24.59
	$20, \ldots, 28$	$\mathscr{O}_3$	32.51	36.26	25.06	29.85
	$29, \ldots, 37$	$\mathscr{O}_4$	36.72	40.55	30.51	35.88
	$38, \ldots, 46$	$\mathscr{O}_5$	41.01	44.88	36.55	41.79
	$47, \ldots, 56$	$\mathscr{O}_6$	45.34	49.70	42.43	48.00
	$57, \ldots, 65$	$\mathscr{O}_7$	50.20	54.07	48.60	53.24
	$66, \ldots, 74$	$\mathscr{O}_8$	54.56	58.45	53.80	58.17
	$75, \ldots, 83$	$\mathscr{O}_9$	58.94	62.83	58.70	62.81
	$84, \ldots, 200$	$\mathscr{O}_{10}$	63.31	107.22	63.31	107.22

Table 26.10 A-Optimal (max. Fisher information) progressive censoring plans for certain (small) m and n and an extreme value distribution (Weibull distribution). The traces of the Fisher information matrices are given for the optimal one-step plans (OSP) and right censored samples (RC) (for both the minimum value n_1 and the maximum value n_2 of n in each row). Values in sans-serif font are computed by a genetic algorithm due to Vuong and Cramer [881]. The values are taken from Dahmen et al. [320]

			Fisher information (in units of $1/\theta^2$)			
	n		OSP		RC	
m	$n_1, \ldots, n_2$	OSP	n_1	n_2	n_1	n_2
2	$3, \ldots, 5000$	$\mathscr{O}_1$	7.48	79.12	6.09	5.58
3	$4, \ldots, 5000$	$\mathscr{O}_1$	16.51	159.86	13.41	11.56
4	$5, \ldots, 41$	$\mathscr{O}_1$	28.75	65.54	23.73	19.85
	$42, \ldots, 5000$	$\mathscr{O}_2$	66.13	281.42	19.84	19.54
5	6, 7	$\mathscr{O}_1$	44.21	47.01	37.11	35.26
	$8, \ldots, 5000$	$\mathscr{O}_2$	49.61	424.20	34.16	29.54
6	$7, \ldots, 884$	$\mathscr{O}_2$	63.15	371.44	53.59	41.58
	$885, \ldots, 5000$	$\mathscr{O}_3$	371.55	584.08	41.58	41.54
⋮		⋮				
10	11	$\mathscr{O}_2$	171.62		151.08	
	$12, \ldots, 116$	$\mathscr{O}_3$	178.46	464.55	143.71	111.54
	$117, \ldots, 5000$	$\mathscr{O}_4$	466.03	1473.28	111.54	109.56

Table 26.11 D-Optimal (max. Fisher information) progressive censoring plans for certain (small) m and n and an extreme value distribution (Weibull distribution). The determinants of the Fisher information matrices are given for the optimal one-step plans (OSP) and right censored samples (RC) (for both the minimum value n_1 and the maximum value n_2 of n in each row). Values in sans-serif font are computed by a genetic algorithm due to Vuong and Cramer [881]. The values are taken from Dahmen et al. [320]

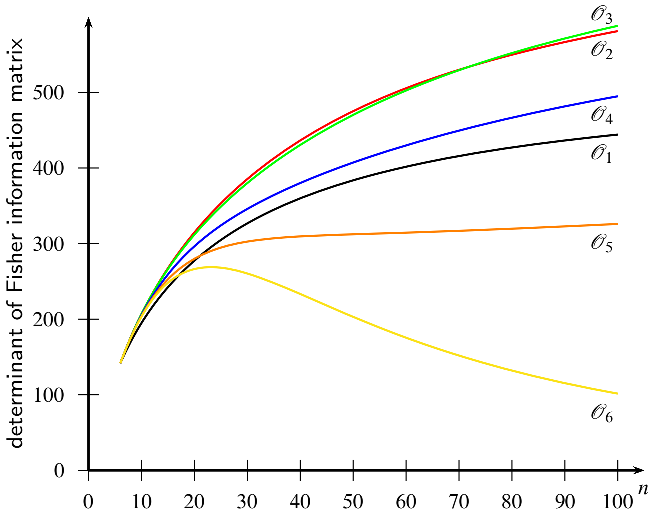

Fig. 26.6 Comparison of one-step censoring plans $\mathscr{O}_1, \ldots, \mathscr{O}_6$ for $m = 6$, $\vartheta = 1$, $q = 0.1$, and sample sizes $n = 6, \ldots, 100$ w.r.t. the determinant of the Fisher information matrix

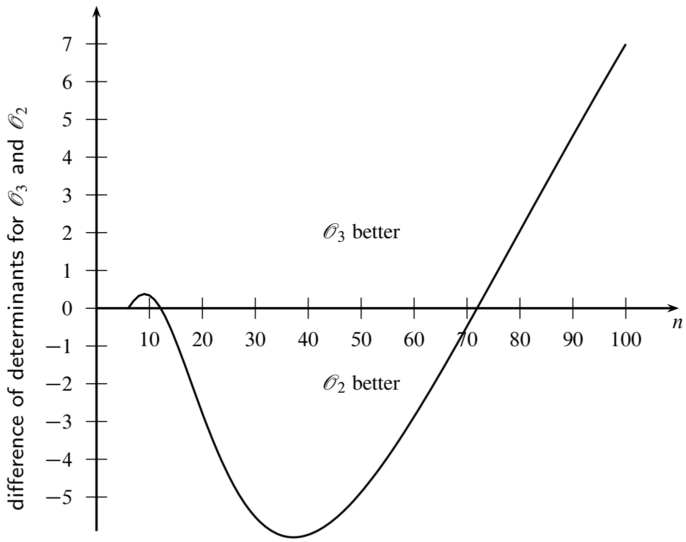

Fig. 26.7 Comparison of one-step censoring plans $\mathscr{O}_2$ and $\mathscr{O}_3$ for $m = 6$, $\vartheta = 1$, $q = 0.1$, and sample sizes $n = 6, \ldots, 100$ w.r.t. the differences of determinant of the Fisher information matrices

Then, $\gamma_i(\mathscr{O}_j) = n - i + 1$, $1 \leq i \leq j$, and $\gamma_i(\mathscr{O}_j) = m - j + 1$, $j + 1 \leq i \leq m$. Hence, for $n \to \infty$, the asymptotic Fisher information reads

$$\mathscr{I}_\infty(j; q, \vartheta) = \begin{bmatrix} \frac{m}{q^2} & \frac{1}{q\vartheta}\tau_1(j) \\ \frac{1}{q\vartheta}\tau_1(j) & \frac{1}{\vartheta^2}\tau_2(j) \end{bmatrix},$$

where

$$\tau_k(j) = \lim_{n\to\infty} \varphi_k(j) = j + \sum_{s=j+1}^{m} \prod_{\ell=j+1}^{s} \left(1 - \frac{k}{q(m-\ell+1)+k}\right), \quad k = 1, 2.$$

In order to find the asymptotically D-optimal plan, the optimal index j has been identified as illustrated by the following example taken from Dahmen et al. [320].

Example 26.4.5. For $q = 1$, the above expression simplifies to

$$\mathscr{I}_\infty(j; 1, \vartheta) = \begin{bmatrix} m & \frac{m+j}{2\vartheta} \\ \frac{m+j}{2\vartheta} & \frac{m+2j}{3\vartheta^2} \end{bmatrix},$$

which shows $12\vartheta^2 \det(\mathscr{I}_\infty(j; 1, \vartheta)) = 4m(m+2j) - 3(m+j)^2 = m^2 + 2mj - 3j^2$. Hence, the optimal j^* is included in the set $\{\lfloor m/3 \rfloor, \lfloor m/3 \rfloor + 1\}$ so that the optimal one-step censoring plan is $\mathscr{O}_{\lfloor m/3 \rfloor}$ or $\mathscr{O}_{\lfloor m/3 \rfloor+1}$. This means that optimal progressive censoring is carried out after having observed $1/3$ of the data intended to be observed. For $m = 6$, this yields $j^* = 2$, so that the D-optimal one-step censoring plan is given by $\mathscr{O}_2$. The corresponding determinant of the asymptotic Fisher information matrix is given by $\det(\mathscr{I}_\infty(2; 1, 1)) = 4$. For comparison, $\det(\mathscr{I}_\infty(3; 1, 1)) = \det(\mathscr{I}_\infty(1; 1, 1)) = 3.75$. This shows that $\mathscr{O}_1$ and $\mathscr{O}_3$ are asymptotically equivalent.

26.4.3 Asymptotically Optimal Censoring Schemes

From the preceding considerations, it is clear that the computation of optimal censoring plans is in most cases a very hard problem particularly for large m and n. On the other hand, it turns out that one-step plans are optimal in many cases as suggested by the small-sample results for Fisher information. In particular, the optimal censoring schemes seem to follow quite a regular pattern: For fixed m and increasing n, the optimal censoring plan moves from first-step censoring to second-step censoring, then to third-step censoring, etc., until it reaches finally the right censoring. This leads us to the conjecture that, for sufficiently large n, right censoring remains optimal. In order to study this behavior, Cramer and Ensenbach [293] introduced the idea of asymptotically optimal censoring plans (see Definition 26.4.6). This concept can be adapted to any optimality criteria but is illustrated here for maximum Fisher information.

Definition 26.4.6. For $n \in \mathbb{N}$, denote by $\mathscr{I}^*(n)$ the maximum and by $\mathscr{I}_*(n)$ the minimum of $\mathscr{I}(\mathbf{X}^{\mathscr{R}})$ for $\mathscr{R} \in \mathscr{C}^m_{n,m}$.

A sequence $(\mathscr{R}_n)_{n\in\mathbb{N}}$ of censoring schemes $\mathscr{R}_n \in \mathscr{C}^m_{n,m}$ is said to be *asymptotically optimal* (w. r. t. Fisher information) if $\mathscr{I}(\mathbf{X}^{\mathscr{R}_n}) \sim \mathscr{I}^*(n)$, i. e., if $\lim_{n\to\infty} \mathscr{I}(\mathbf{X}^{\mathscr{R}_n})/\mathscr{I}^*(n) = 1$.

Cramer and Ensenbach [293] presented a simple criterion in terms of the function (see (26.7))

$$w \equiv w_\theta : \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{\partial}{\partial \eta} \log \lambda_\eta(x)\bigg|_{\eta=\theta}.$$

First, the following result holds showing that right censoring is asymptotically optimal provided that the minimal Fisher information $\mathscr{I}_*(n)$ tends to ∞, $n \to \infty$.

Theorem 26.4.7. Let w^2 be decomposable into countably many monotone pieces. Then, there exists a constant $C \in \mathbb{R}$ independent of n such that

$$\mathscr{I}^*(n) \le \mathscr{I}(\mathbf{X}^{\mathscr{O}_{m,n}}) + C \quad \text{for every } n \ge m,$$

where $\mathscr{O}_{m,n} = (0^{*m-1}, n-m) \in \mathscr{C}^m_{m,n}$ denotes the one-step plan corresponding to right censoring and original sample size n.

Furthermore, for $\lim_{n\to\infty} \mathscr{I}_*(n) = \infty$, right censoring is asymptotically optimal, i.e., $\lim_{n\to\infty} \mathscr{I}(\mathbf{X}^{\mathscr{O}_{m,n}})/\mathscr{I}^*(n) = 1$.

Theorem 26.4.7 shows that, for certain distributions, the values of the Fisher information for right censoring and for the optimal censoring scheme differ at most by an additive constant independent of n. A sufficient condition ensuring $\lim_{n\to\infty} \mathscr{I}_*(n) = \infty$ is given by $\liminf_{x\to\infty} w(x)^2/|\log F(x)| > 0$ (see Cramer and Ensenbach [293, Theorem 2.7]). As pointed out by these authors, a simple criterion for asymptotic optimality of right censoring is given by the unboundedness of the sequence $(\mathscr{I}(X_{1:n}))_{n\in\mathbb{N}}$. Notice that $(\mathscr{I}(X_{1:n}))_{n\in\mathbb{N}}$ determines the Fisher information in the Type-II censored sample $(X_{1:n}, \ldots, X_{m:n})$, $1 \leq m \leq n$, completely (see Park [704], Park and Zheng [709], and Zheng et al. [945]). A similar result holds for progressively Type-II censored order statistics (see (9.16) and Abo-Eleneen [7, Theorem 4.1]).

Applying the above criteria to particular distributions, Cramer and Ensenbach [293] established that right censoring is asymptotically optimal for Type-I and Type-II extreme value distributions (scale), Weibull distributions (shape), logistic distributions (scale), normal distributions (scale), and Laplace distributions (location).

For a scale Laplace distribution, the situation is different. In this case, Cramer and Ensenbach [293] showed that all one-step censoring plans have the same asymptotic Fisher information and, thus, are asymptotically equivalent. In particular, $\mathscr{I}(\mathbf{X}^{\mathscr{O}_{k,n}}) \xrightarrow{n\to\infty} m$ holds for any $k \in \{1, \ldots, m\}$.

Finally, Cramer and Ensenbach [293] established the following result providing a characterization of asymptotic optimality of one-step censoring.

Theorem 26.4.8. Suppose there exists a function $g : [1, \infty) \to (0, \infty)$ with $g(n) = \mathscr{I}(X_{1:n})$ for all $n \in \mathbb{N}$ which satisfies some regularity conditions. Then,

$$\lim_{n\to\infty} \frac{\mathscr{I}(\mathbf{X}^{\mathscr{O}_{k,n}})}{g(n)} = k$$

holds

(i) for $k = 1$ and $k = m$ provided that $ng(n) \xrightarrow{n\to\infty} \infty$, and
(ii) for all $k \in \{1, \ldots, m\}$ provided that $g(n) \xrightarrow{n\to\infty} \infty$, respectively.

As a consequence of this result, the following corollaries hold.

Corollary 26.4.9. (i) Let g be as in Theorem 26.4.8 (i) and $m > 1$. Then, $(\mathscr{O}_{1,n})_{n\in\mathbb{N}}$ is not asymptotically optimal;
(ii) If additionally $g(n) \xrightarrow{n\to\infty} \infty$, then $(\mathscr{O}_{k,n})_{n\in\mathbb{N}}$ is not asymptotically optimal for any $1 \leq k < m$.

Corollary 26.4.10. For a scale-parameter extreme value distribution (Type I), right censoring is the only asymptotically optimal one-step censoring plan. Moreover,

$$\lim_{n\to\infty} \frac{\mathscr{I}(\mathbf{X}^{\mathscr{O}_{k,n}})}{(\log n)^2} = k, \quad 1 \le k \le m.$$

26.4.4 Maximum Fisher Information Plans in Progressive Hybrid Censoring

For Type-I and Type-II progressive hybrid censoring, Park et al. [711] have utilized the expressions in (9.18) and (9.19) to determine maximum Fisher information censoring plans. Using the expression in (9.18) and the stochastic ordering $\mathbf{X}^{\mathscr{O}_m} \le_{\text{st}} \mathbf{X}^{\mathscr{R}}$ for any censoring scheme $\mathscr{R} \in \mathscr{C}^m_{m,n}$ (see (26.5)), it follows readily that right censoring maximizes Fisher information in Type-II progressive hybrid censoring, too. However, the result can also be established for other distributions provided that the function w_θ given in (26.7), i.e., $x \mapsto w_\theta(x) = \left\{\frac{\partial}{\partial\theta}\log\lambda_\theta(x)\right\}^2$, is decreasing or increasing. Obviously, we can formulate the same results as given in Theorem 26.4.1. Notice that, now, the monotonicity of w_θ must only hold on the interval $[0, T]$. Hence, for location normal and logistic distributions, right censoring is optimal. Further examples can be taken from the discussion following Theorem 26.4.1.

For Type-II progressive hybrid censoring, (9.19) shows that right censoring maximizes Fisher information, too. Noticing that $R_m \le n - m$, a simple calculation yields the upper bound

$$\begin{aligned}\mathscr{I}_{m\vee T:m:n}(\vartheta) &= \mathscr{I}_{m\wedge T:m:n}(\vartheta) + \frac{1}{\vartheta^2}\sum_{i=n-R_m+1}^{n} F_{i:n}\Big(\frac{T}{\vartheta}\Big)\\ &\le \mathscr{I}_{m\wedge T:n}(\vartheta) + \frac{1}{\vartheta^2}\sum_{i=m+1}^{n} F_{i:n}\Big(\frac{T}{\vartheta}\Big),\end{aligned}$$

which is attained for right censoring.

26.5 Other Optimality Criteria and Approaches

26.5.1 Maximum Entropy Plans

Using the entropy expressions given in Sect. 9.4, Cramer and Bagh [291] established maximum/minimum entropy plans for the sample $\mathbf{X}^{\mathscr{R}}$. In particular, it is shown that the entropy $\mathscr{H}^{\mathscr{R}}_{1,\ldots,m:m:n}$ given in (9.20) satisfies the conditions of Theorem 26.1.2 provided F is a DFR distribution.

Theorem 26.5.1. Let F be a DFR-cumulative distribution function. Then, $\mathscr{O}_1$ is a maximum entropy plan and $\mathscr{O}_m$ is a minimum entropy plan.

Important examples for distributions included in Theorem 26.5.1 are exponential, mixtures of exponential, and Pareto distributions. Moreover, gamma and Weibull distributions with shape parameter less than one have the DFR-property. For IFR-distributions, the situation is more involved. It is shown by Cramer and Bagh [291] that for RPower(ν)-distributions, the situation is reversed: $\mathscr{O}_m$ is the unique maximum entropy plan and $\mathscr{O}_1$ is the unique minimum entropy plan provided $\nu \geq 1$. Notice that $\nu = 1$ corresponds to the uniform distribution. Moreover, it is shown that for $\nu \leq \frac{1}{n-1}$, the result is as in Theorem 26.5.1. In the remaining cases, inner censoring plans may be optimal. For details, we refer to Cramer and Bagh [291]. Similar results can be obtained for IFR-Weibull distributions.

Remark 26.5.2. (i) The same problem has also been addressed in Abo-Eleneen [9] who presented some numerical results for normal and logistic distributions;
(ii) Further optimality results in the same directions can be obtained using Kullback–Leibler divergence and $\mathscr{I}_\alpha$-divergence as an information criteria. For details, see Cramer and Bagh [291].

26.5.2 Optimal Estimation of Quantiles

For the extreme value distribution, Ng et al. [689] considered optimal estimation of quantiles w.r.t. the underlying censoring scheme. The maximum likelihood estimator of the pth quantile is given by

$$\widehat{\xi}_p = \widehat{\mu} + \widehat{\vartheta} F^{\leftarrow}(p),$$

where $F^{\leftarrow}(p) = \log(-\log(1-p))$, $p \in (0,1)$, and $\widehat{\mu}$ and $\widehat{\vartheta}$ are the maximum likelihood estimates of the location and scale parameters. Minimizing the variance of the quantile yields the optimization problem

$$\mathsf{Var}(\hat{\xi}_p) = \mathsf{Var}(\widehat{\mu}) + \big(F^{\leftarrow}(p)\big)^2 \mathsf{Var}(\widehat{\vartheta}) + 2F^{\leftarrow}(p)\,\mathsf{Cov}(\widehat{\mu}, \widehat{\vartheta}) \longrightarrow \min_{\mathscr{R} \in \mathscr{C}^m_{m,n}} . \quad (26.13)$$

They illustrated their approach by the scenario used to generate data 1.1.5. For $\xi_{0.95}$ and $\xi_{0.05}$ and fixed $n = 19$, $m = 8$, they found the optimal plans $\mathscr{O}_1$ and $\mathscr{O}_5$, respectively. They also provided the gain of efficiency using the optimal censoring plan in favor of the effectively applied censoring scheme $\mathscr{R} = (0^{*2}, 3, 0, 3, 0^{*2}, 5)$.

Since the optimal plan obtained from (26.13) will depend in most cases on the percentile p, several authors have proposed an integrated version of the criterion (see, e.g., Kundu [557], Pradhan and Kundu [727], and Pradhan and Kundu [728]). Inspired by Gupta and Kundu [423], they considered the MLE $\widehat{\xi}_p$ of the quantile ξ_p. Given a censoring scheme $\mathscr{R}$ and a probability measure W on the unit interval [0, 1], an information measure is defined via

$$\mathscr{I}_W(\mathscr{R}) = \int_{[0,1]} V_p(\mathscr{R}) dW(p),$$

where $V_p(\mathscr{R})$ denotes the asymptotic variance of $\widehat{\xi}_p$. Notice that a one-point measure W_{p_0} in p_0 yields the criterion proposed by Ng et al. [689]. A similar idea is present in Zhang and Meeker [940].

For generalized exponential distributions with cumulative distribution function as in (12.41), Pradhan and Kundu [727] obtained the asymptotic variance $\xi_p^2 \mathbf{v}'\mathbf{V}\mathbf{v}$ of the MLE for the quantile $\xi_p = -\frac{1}{\lambda}\log(1 - p^{1/\alpha})$, where

$$\mathbf{v}' = \Big(\frac{p^{1/\alpha}(-\log p)}{\alpha^2[-\log(1 - p^{1/\alpha})](1 - p^{1/\alpha})}, -\frac{1}{\lambda}\Big)$$

and $\mathbf{V}$ is the asymptotic variance–covariance matrix of the MLEs of (α, λ). They provided some computational results regarding optimal censoring plans. The same approach has been used by Pradhan and Kundu [728] for Birnbaum–Saunders distribution. Sultan et al. [827] applied the approach to the inverse Weibull or Fréchet distributions.

Remark 26.5.3. Bayesian versions of this criterion have been utilized in Kundu [557] (Weibull distribution) and Kundu and Pradhan [562] (generalized exponential distribution). For optimal censoring plans in the framework of competing risks from Weibull distributions, we refer to Pareek et al. [703] and Kundu and Pradhan [563]. Linear hazard rate distributions are discussed in Sen et al. [792].

Optimal estimation of quantiles in a nonparametric setting has been discussed in Balakrishnan and Han [98]. Given a progressively Type-II censored sample $Y_{1:m:n}^{\mathscr{R}}, \ldots, Y_{m:m:n}^{\mathscr{R}}$ and $\alpha, p \in (0, 1)$, denote $P(Y_{r:m:n}^{\mathscr{R}} \geq \xi_p)$ by $\zeta_r = \zeta_r(p)$. Then, Balakrishnan and Han [98] proposed to minimize the quantity

$$M_\alpha(\mathscr{R}) = \min_{(r,s)\in\mathcal{S}_\alpha} \{e_s - e_r\},$$

where $e_j = EF(Y_{j:m:n}^{\mathscr{R}}) = EU_{j:m:n}^{\mathscr{R}} = 1 - \prod_{i=1}^{j} \frac{\gamma_i(\mathscr{R})}{\gamma_i(\mathscr{R})+1}$ (see Theorem 7.2.3), $1 \leq j \leq m$, under the constraints

$$\mathcal{S}_\alpha = \{(k, \ell) \mid 1 \leq k < \ell \leq m, \zeta_\ell - \zeta_k \geq 1 - \alpha\}.$$

The results were illustrated by a numerical study.

26.5.3 Optimization Based on Pitman Closeness

In case of the exponential distribution, as seen above, most of the criteria discussed so far are invariant w.r.t. the censoring plans. Volterman et al. [877] proposed Pitman closeness (see Sect. 9.6) as a criterion to choose a preferable censoring scheme by comparing the BLUEs based on two censoring schemes $\mathscr{R}$ and $\mathscr{S}$. Given a censoring scheme $\mathscr{R}$, denote the BLUE for the scale parameter ϑ of the population distribution $\mathsf{Exp}(\vartheta)$ by $\widehat{\vartheta}_{\mathscr{R}}$. Then, the Pitman closeness of the BLUEs $\widehat{\vartheta}_{\mathscr{R}}$ and $\widehat{\vartheta}_{\mathscr{S}}$ is given by

$$\pi(\mathscr{R},\mathscr{S}) = P(|\widehat{\vartheta}_{\mathscr{R}}-\vartheta| < |\widehat{\vartheta}_{\mathscr{S}}-\vartheta|) = P(|\widehat{\vartheta}_{\mathscr{R}}/\vartheta-1| < |\widehat{\vartheta}_{\mathscr{S}}/\vartheta-1|).$$

Notice that this probability is independent of the parameter ϑ because the distributions of $\widehat{\vartheta}_{\mathscr{R}}/\vartheta$ and $\widehat{\vartheta}_{\mathscr{S}}/\vartheta$ are parameter-free. The calculation of $\pi(\mathscr{R},\mathscr{S})$ needs the joint distribution of the BLUEs $\widehat{\vartheta}_{\mathscr{R}} = \frac{1}{m}\sum_{j=1}^{m}(R_j+1)Z_{j:m:n}^{\mathscr{R}}$ and $\widehat{\vartheta}_{\mathscr{S}} = \frac{1}{m}\sum_{j=1}^{m}(S_j+1)Z_{j:m:n}^{\mathscr{S}}$ which are dependent in the present situation. This causes some difficulties in deriving an expression for $\pi(\mathscr{R},\mathscr{S})$. In order to find such an expression, Volterman et al. [877] considered the region where $\widehat{\vartheta}_{\mathscr{R}}$ is Pitman closer to ϑ than $\widehat{\vartheta}_{\mathscr{S}}$, i.e., where (w.l.o.g. let $\vartheta = 1$)

$$|\widehat{\vartheta}_{\mathscr{R}}-1| < |\widehat{\vartheta}_{\mathscr{S}}-1|.$$

Some simple calculations show that this inequality is satisfied iff

$$(\widehat{\vartheta}_{\mathscr{R}}-\widehat{\vartheta}_{\mathscr{S}})(\widehat{\vartheta}_{\mathscr{R}}+\widehat{\vartheta}_{\mathscr{S}}-2) < 0.$$

Therefore, the BLUE $\widehat{\vartheta}_{\mathscr{R}}$ will be Pitman closer than the BLUE $\widehat{\vartheta}_{\mathscr{S}}$ if either

$$\begin{aligned} &\widehat{\vartheta}_{\mathscr{R}} > \widehat{\vartheta}_{\mathscr{S}} \text{ and } \widehat{\vartheta}_{\mathscr{R}}+\widehat{\vartheta}_{\mathscr{S}} < 2 \quad \text{or} \\ &\widehat{\vartheta}_{\mathscr{R}} < \widehat{\vartheta}_{\mathscr{S}} \text{ and } \widehat{\vartheta}_{\mathscr{R}}+\widehat{\vartheta}_{\mathscr{S}} > 2. \end{aligned} \tag{26.14}$$

Assuming that the original sample is given by n IID random variables $Z_1,\ldots,Z_n$, it follows that the BLUEs are linear functions of the order statistics $Z_{1:n},\ldots,Z_{n:n}$. As pointed out by Volterman et al. [877], the BLUEs can be written in terms of these order statistics with a certain probability. To illustrate this approach, the following example is taken from Volterman et al. [877].

Example 26.5.4. Let $n = 4$, $m = 3$, and consider the one-step plans $\mathscr{O}_j$, $j = 1,2,3$. Then,

$$\widehat{\vartheta}_{\mathscr{O}_1} = \begin{cases} \frac{2}{3}Z_{1:4} + \frac{1}{3}Z_{2:4} + \frac{1}{3}Z_{3:4} & \text{with probability } \frac{1}{3} \\ \frac{2}{3}Z_{1:4} + \frac{1}{3}Z_{2:4} + \frac{1}{3}Z_{4:4} & \text{with probability } \frac{1}{3} \\ \frac{2}{3}Z_{1:4} + \frac{1}{3}Z_{3:4} + \frac{1}{3}Z_{4:4} & \text{with probability } \frac{1}{3} \end{cases},$$

$$\widehat{\vartheta}_{\mathscr{O}_2} = \begin{cases} \frac{1}{3}Z_{1:4} + \frac{2}{3}Z_{2:4} + \frac{1}{3}Z_{3:4} & \text{with probability } \frac{1}{2} \\ \frac{1}{3}Z_{1:4} + \frac{2}{3}Z_{2:4} + \frac{1}{3}Z_{4:4} & \text{with probability } \frac{1}{2} \end{cases},$$

$$\widehat{\vartheta}_{\mathscr{O}_3} = \frac{1}{3}Z_{1:4} + \frac{1}{3}Z_{2:4} + \frac{2}{3}Z_{3:4}.$$

These expressions can now be rewritten in terms of the normalized spacings $S_{1,4}, \ldots, S_{4,4}$ of the order statistics $Z_{1:4}, \ldots, Z_{4:4}$. This leads to the linear representations

$$\widehat{\vartheta}_{\mathscr{O}_1} = \begin{cases} \frac{1}{3}S_{1,4} + \frac{2}{9}S_{2,4} + \frac{1}{6}S_{3,4} & \text{with probability } \frac{1}{3} \\ \frac{1}{3}S_{1,4} + \frac{2}{9}S_{2,4} + \frac{1}{6}S_{3,4} + \frac{1}{3}S_{4,4} & \text{with probability } \frac{1}{3} \\ \frac{1}{3}S_{1,4} + \frac{2}{9}S_{2,4} + \frac{1}{3}S_{3,4} + \frac{1}{3}S_{4,4} & \text{with probability } \frac{1}{3} \end{cases},$$

$$\widehat{\vartheta}_{\mathscr{O}_2} = \begin{cases} \frac{1}{3}S_{1,4} + \frac{2}{3}S_{2,4} + \frac{1}{6}S_{3,4} & \text{with probability } \frac{1}{2} \\ \frac{1}{3}S_{1,4} + \frac{1}{3}S_{2,4} + \frac{1}{6}S_{3,4} + \frac{1}{3}S_{4,4} & \text{with probability } \frac{1}{2} \end{cases},$$

$$\widehat{\vartheta}_{\mathscr{O}_3} = \frac{1}{3}S_{1,4} + \frac{1}{3}S_{2,4} + \frac{1}{3}S_{3,4}.$$

Volterman et al. [877] obtained $\pi(\mathscr{O}_1, \mathscr{O}_2) = 0.5363$ and $\pi(\mathscr{O}_2, \mathscr{O}_3) = 0.5526$ showing that right censoring yields the Pitman closest BLUE.

In general, it is clear that both BLUEs are a random linear combination of the order statistics $Z_{1:n}, \ldots, Z_{n:n}$. Using the generation process and denoting the random indices by $K_1, \ldots, K_m$ and $L_1, \ldots, L_m$, respectively, we get

$$\widehat{\vartheta}_{\mathscr{R}} = \frac{1}{m}\sum_{j=1}^{m}(R_j + 1)Z_{K_j:n} \quad \text{and} \quad \widehat{\vartheta}_{\mathscr{S}} = \frac{1}{m}\sum_{j=1}^{m}(S_j + 1)Z_{L_j:n}. \tag{26.15}$$

Clearly, from (2.10), the estimators can always be expressed in terms of the normalized spacings $S_{1,n}, \ldots, S_{n,n}$ which are IID standard exponential random variables (see Theorem 2.3.2). Therefore, this leads to the expressions

$$\widehat{\vartheta}_{\mathscr{R}} = \sum_{j=1}^{n} w_j(\mathbf{K}_m)S_{j,n} \quad \text{and} \quad \widehat{\vartheta}_{\mathscr{S}} = \sum_{j=1}^{n} \widetilde{w}_j(\mathbf{L}_m)S_{j,n}.$$

By construction, $\mathbf{K}_m = (K_j)_j$ and $\mathbf{L}_m = (L_j)_j$ are independent random variables with probability mass function given in (10.9). In order to derive the Pitman closeness probability $\pi(\mathscr{R}, \mathscr{S})$, Volterman et al. [877] conditioned on $\mathbf{K}_m$ and $\mathbf{L}_m$ so that

$$\pi(\mathscr{R}, \mathscr{S}) = \sum_{\mathbf{k}_m, \mathbf{l}_m} P(\mathbf{K}_m = \mathbf{k}_m) P(\mathbf{L}_m = \mathbf{l}_m) \\ P\Big(\Big|\sum_{j=1}^{n} w_j(\mathbf{k}_m) S_{j,n} - 1\Big| < \Big|\sum_{j=1}^{n} \widetilde{w}_j(\mathbf{l}_m) S_{j,n} - 1\Big|\Big). \quad (26.16)$$

Hence, the desired Pitman closeness probability is a weighted sum of Pitman closeness probabilities for two L-statistics.

For any given censoring schemes $\mathscr{R}$ and $\mathscr{S}$, Volterman et al. [877] proposed a general algorithm based on the mixture representation of progressively Type-II censored order statistics (see Guilbaud [419]) to calculate the exact Pitman closeness probability.

Algorithm 26.5.5. Let $\mathscr{R}, \mathscr{S} \in \mathscr{C}_{m,n}^m$ be censoring schemes. Then, the following steps lead to the Pitman closeness probability $\pi(\mathscr{R}, \mathscr{S})$:

① Express the density function of the progressively Type-II censored order statistics as a mixture of density function of order statistics

$$f^{\mathbf{X}^{\mathscr{R}}}(\mathbf{t}_m) = \sum_{1 \le k_1 < \cdots < k_m \le n} w_{\mathbf{k}_m} f^{X_{k_1:n}, \ldots, X_{k_m:n}}(\mathbf{t}_m), \quad \mathbf{t}_m \in \mathbb{R}^m;$$

② For each $(X_{k_1:n}, \ldots, X_{k_m:n})$, let conditionally on $\mathbf{X}^{\mathscr{R}} = (X_{k_1:n}, \ldots, X_{k_m:n})$ (see (26.15))

$$\widehat{\vartheta}_{\mathscr{R}}(\mathbf{k}_m) = \frac{1}{m} \sum_{j=1}^{m} (R_j + 1) Z_{k_j:n}. \quad (26.17)$$

Then,

$$\widehat{\vartheta}_{\mathscr{R}} = \sum_{1 \le k_1 < \cdots < k_m \le n} w_{\mathbf{k}_m} \widehat{\vartheta}_{\mathscr{R}}(\mathbf{k}_m); \quad (26.18)$$

③ Rewrite expression (26.17) in terms of normalized spacings as

$$X_{i:n} = \sum_{\ell=1}^{i} \frac{1}{n - \ell + 1} S_{\ell,n},$$

where $S_{1,n}, \ldots, S_{n,n}$ are IID Exp(ϑ)-distributed random variables;

④ Define the constraints given in (26.14) in terms of $\widehat{\vartheta}_{\mathscr{R}}(\mathbf{k}_m)$ and $\widehat{\vartheta}_{\mathscr{S}}(\mathbf{l}_m)$;
⑤ Integrate the joint density function of $S_{1,n}, \ldots, S_{n,n}$ over $[0,\infty)^n$ subject to the constraints (26.14);
⑥ Evaluate the weighted sum given in (26.16).

Clearly, the computational effort grows fast while m (and n) increases. For this reason, Volterman et al. [877] proposed Monte Carlo simulation of $\pi(\mathscr{R},\mathscr{S})$ for large m.

Volterman et al. [877] provided many tables to investigate the Pitman closeness of the BLUEs. It turned out that right censoring seems to generate the Pitman closest BLUE among all BLUEs in the progressive censoring setting. Moreover, they observed that the neighbor $\mathscr{R}^* = (0^{*m-2}, 1, n-m-1)$ has the highest Pitman closeness probability among all comparisons studied in the simulation.

26.5.4 Optimal Block Censoring

Hofmann et al. [444] determined optimal censoring plans in the asymptotic block censoring model introduced in Sect. 3.4.6. As a measure of optimality, they chose minimal determinant of the variance–covariance matrix of the ABLUEs given in (11.19). They showed that it is equivalent to maximizing the expression $\mathcal{V} = K_1K_2 - K_3^2$. From (11.18), they obtained the following expression:

$$\mathcal{V} = \sum_{1\le j<k\le m} v_j^{-1}v_k^{-1}\Big((\Delta_j^{-1} - \Delta_{j-1}^{-1})(u_k\Delta_k^{-1} - u_{k-1}\Delta_{k-1}^{-1}) - (\Delta_k^{-1} - \Delta_{k-1}^{-1})(u_j\Delta_j^{-1} - u_{j-1}\Delta_{j-1}^{-1})\Big)^2,$$

where $\Delta_0^{-1} = 0$. Rewriting $\Delta_i^{-1} = f(u_i)/\prod_{j=1}^i p_j$ and v_i^{-1} as

$$v_i^{-1} = \frac{p_i t_0}{1-p_i}\prod_{j=1}^{i-1}(p_j t_j), \quad 1 \le i \le m,$$

this leads to the expression

$$\mathcal{V} = t_0^2 \sum_{1\le j<k\le m} \frac{1}{(1-p_j)p_j(1-p_k)p_k}\prod_{i=1}^{j-1}(p_i^{-1}t_i)\prod_{i=1}^{k-1}(p_i^{-1}t_i)(U_{jk} - U_{kj})^2, \tag{26.19}$$

where $U_{jk} = [f(u_j) - p_j f(u_{j-1})][u_k f(u_k) - p_k u_{k-1} f(u_{k-1})]$, $1 \le j < k \le m$, and $f(u_0) = 0$.

Since there is a one-to-one correspondence between $\lambda_0, \ldots, \lambda_m$, $\overline{\lambda}_1, \ldots, \overline{\lambda}_m$ $(\lambda_0 + \sum_{i=1}^m (\overline{\lambda}_i + \lambda_i) = 1)$ and $p_1, \ldots, p_m, t_0, \ldots, t_{m-1}$ (see (3.19)), i.e.,

$$\overline{\lambda}_i = t_0(1 - p_i) \prod_{j=1}^{i-1} (p_j t_j), \quad \lambda_i = t_0(1 - t_i) p_i \prod_{j=1}^{i-1} (p_j t_j), \quad 1 \le i \le m, \tag{26.20}$$

the optimization can be carried out w.r.t. $p_1, \ldots, p_m \in (0, 1)$ and $t_0, \ldots, t_{m-1} \in (0, 1]$. Hofmann et al. [444] pointed out that $\mathcal{V}$ is positive on the considered region of the parameters (see (26.19) and notice that $\mathcal{V}$ is the determinant of a positive definite matrix). However, from Example 26.5.7, it follows that $\mathcal{V}$ may be unbounded which also has been reported for the setting of order statistics (cf. Sarhan et al. [777], Saleh and Ali [767], and Ali and Umbach [33]). Theorem 26.5.6 provides conditions on the population cumulative distribution function F for $\mathcal{V}$ to be bounded.

Theorem 26.5.6 (Hofmann et al. [444]). Let $p_1, \ldots, p_m \in (0, 1)$ and F be an absolutely continuous cumulative distribution function with support $(\alpha(F), \omega(F))$, $-\infty \le \alpha(F) < \omega(F) \le \infty$, and continuous density function f on this support. For $t \in (0, 1]$, let

$$g_t(x) = f(F^{\leftarrow}(1-tx)), \quad h_t(x) = F^{\leftarrow}(1-tx) f(F^{\leftarrow}(1-tx)), \qquad x \in (0, 1).$$

Suppose g_1 and h_1 are differentiable on $(0, 1)$. Further, consider the following conditions:

$$\omega(F) < \infty \text{ and } \lim_{x \to 0} \frac{g_1(x)}{\sqrt{x}} \in \mathbb{R}, \tag{D1}$$

$$\omega(F) = \infty \text{ and } \lim_{x \to 0} \frac{h_1(x)}{\sqrt{x}} \in \mathbb{R}, \tag{D2}$$

$$\alpha(F) > -\infty \text{ and } \lim_{x \to 1} \frac{g_1(x)}{\sqrt{1-x}} \in \mathbb{R}, \tag{D3}$$

$$\alpha(F) = -\infty \text{ and } \lim_{x \to 1} \frac{h_1(x)}{\sqrt{1-x}} \in \mathbb{R}, \tag{D4}$$

$$\omega(F) < \infty \text{ and } \limsup_{x \to 0} |\sqrt{x} g_1'(x)| < \infty, \tag{D5}$$

$$\omega(F) = \infty \text{ and } \limsup_{x \to 0} |\sqrt{x} h_1'(x)| < \infty. \tag{D6}$$

If F satisfies either (D1) or (D2), and either (D3) or (D4), and either (D5) or (D6), then $\mathcal{V}$ is bounded. If neither (D3) nor (D4) hold, then $\mathcal{V}$ is unbounded.

Example 26.5.7. The following results are established by Hofmann et al. [444]:

(i) For exponential, uniform, and Pareto(α)-distributions, condition (D3) is violated so that $\mathcal{V}$ is unbounded;
(ii) For Weibull$(1, \beta)$-distributions, $\mathcal{V}$ is bounded iff $\beta \geq 2$;
(iii) $\mathcal{V}$ is bounded for the extreme value distribution with density function $f(x) = \exp(x - \exp(x))$, $x \in \mathbb{R}$, and for the normal distribution.

If $\mathcal{V}$ is unbounded, the optimal censoring plan must include the minimum (or maximum) of the sample. But, since the minimum and maximum do not have an asymptotic normal distribution, these cases cannot be handled in the present setup and have to be excluded.

Hofmann et al. [444] pointed out that $\mathcal{V}$ given in (26.19) is an increasing function of $t_0, \ldots, t_{m-1}$, so that, for any $p_1, \ldots, p_m$, the optimal solution is given by $t_i = 1$, $0 \leq i \leq m-1$. This implies $\lambda_i = 0$, $0 \leq i \leq m-1$. Thus, only λ_m is positive which corresponds to the usual Type-II right censoring. This means that the optimal solution of the asymptotic progressive block Type-II censoring scheme is equivalent to the corresponding asymptotic order statistics problem as discussed in Chan and Mead [241] (extreme value), Chan and Chan [240] (normal), and Hassanein [436] (Weibull). Hassanein [435] considered the trace of the covariance matrix (the total variance) for an underlying extreme value distribution. Here, we have to determine the quantiles $0 < \overline{\lambda}_1 < \overline{\lambda}_1 + \overline{\lambda}_2 < \cdots < \sum_{i=1}^m \overline{\lambda}_i < 1$ such that the BLUE based on the order statistics $X_{\overline{R}_1:n}, X_{\overline{R}_1+\overline{R}_2:n}, \ldots, X_{\sum_{i=1}^m \overline{R}_i:n}$ has minimum variance (here, minimum generalized variance). Further examples are provided by Ali and Umbach [33].

Since cost savings are often presented as a motivation for progressive Type-II censoring, Hofmann et al. [444] introduced a restriction to the progressive block Type-II censoring scheme by imposing an upper bound on the total proportion of observed blocks of failures:

$$\sum_{i=1}^m \overline{\lambda}_i \leq \tau \tag{26.21}$$

for some prefixed $\tau \in (0, 1]$. Notice that $\tau = 1$ corresponds to the unrestricted case. Then, rewriting (26.21) in terms of p_i and t_j, it follows from (26.20) that (26.21) is equivalent to the constraint

$$t_0 \sum_{i=1}^m (1 - p_i) \prod_{j=1}^{i-1} p_j t_j \leq \tau. \tag{26.22}$$

Now, Hofmann et al. [444] established the following result concerning the restricted optimization problem. It tells us that for an optimal solution, either all t_i's equal 1 such that right censoring is optimal or the constraint in (26.22) is sharp (i.e., equality holds).

Theorem 26.5.8. Let $(p_1, \ldots, p_m, t_0, \ldots, t_{m-1}) \in (0,1)^m \times (0,1]^m$ be an optimal solution to the problem of maximizing $\mathcal{V}$ in (26.19) under condition (26.22). Then, either $t_0 = \cdots = t_{m-1} = 1$ or equality holds in (26.22).

Computational Results

Hofmann et al. [444] provided some computational results for the maximization of $\mathcal{V}$ under restriction (26.22) for the extreme value distribution, the Weibull$(1,3)$-distribution, and the normal distribution.

First, let $\tau_c = \sum_{i=1}^m \overline{\lambda}_i = 1 - \prod_{i=1}^m p_i$, where the $\overline{\lambda}_i$'s (p_i's) correspond to the optimal solution in the unrestricted case, i.e., for $\tau = 1$. Obviously, the solution will not change as long as $\tau \in [\tau_c, 1]$. For $\tau < \tau_c$, Hofmann et al. [444] reported the following behavior. If τ falls slightly below τ_c, the optimal solution for $p_1, \ldots, p_m$ changes. However, the solution for $t_0, \ldots, t_{m-1}$ remains unchanged so that $\lambda_1 = \cdots = \lambda_{m-1} = 0$, i.e., progressive censoring is not applied. Moreover, it turns out that a critical value $\tau_t < \tau_c$ exists, such that the optimal solution is given by $\lambda_1 = \cdots = \lambda_{m-1} = 0$ for $\tau \in (\tau_t, 1]$, but not for $\tau < \tau_t$. Some (numerically calculated) critical τ-values are presented in Tables 26.12–26.14 for different values of m, which are taken from Hofmann et al. [444].

m	2	3	4	5	10	15	20
τ_c	0.9262	0.9671	0.9789	0.9878	0.9978	0.9993	0.9997
τ_t	0.8867	0.9054	0.9595	0.9741	0.9960	0.9988	0.9995

Table 26.12 Critical τ for the extreme value distribution (see Hofmann et al. [444])

m	2	3	4	5	10
τ_c	0.8602	0.8745	0.9482	0.9565	0.9903
τ_t	0.7202	0.7163	0.7212	0.8357	0.9485

Table 26.13 Critical τ for the Weibull$(1,3)$-distribution (see Hofmann et al. [444])

m	2	3	4	5	10	15	20
τ_c	0.8666	0.9167	0.9551	0.9701	0.9941	0.9979	0.9990
τ_t	0.7370	0.7414	0.8055	0.8669	0.9658	0.9873	0.9940

Table 26.14 Critical τ for the normal distribution (see Hofmann et al. [444])

For $\tau < \tau_t$, the optimal solutions suggested that one-step censoring plans are optimal. For illustration, we present a table for the extreme value distribution in Table 26.15 from Hofmann et al. [444]. There, results for the Weibull$(1,3)$ and the normal distribution can also be found.

m	τ	0.01	0.05	0.1	0.15	0.2	0.5	1
2	$\overline{\lambda}_1$	0.3998E−2	0.1831E−1	0.3462E−1	0.4988E−1	0.6435E−1	0.1398E+0	0.2390E+0
	$\overline{\lambda}_2$	0.6002E−2	0.3169E−1	0.6538E−1	0.1001E+0	0.1356E+0	0.3602E+0	0.6872E+0
	λ_1	0.9881E+0	0.9422E+0	0.8855E+0	0.8291E+0	0.7728E+0	0.4360E+0	
	λ_2	0.1869E−2	0.7838E−2	0.1454E−1	0.2094E−1	0.2718E−1	0.6404E−1	0.7378E−1
	Δ Var	90.8 %	82.9 %	76.4 %	70.8 %	65.6 %	35.7 %	
3	$\overline{\lambda}_1$	0.3960E−2	0.1789E−1	0.3339E−1	0.4751E−1	0.6050E−1	0.1164E+0	0.1902E+0
	$\overline{\lambda}_2$	0.4455E−2	0.2411E−1	0.5009E−1	0.7679E−1	0.1038E+0	0.2547E+0	0.5627E+0
	$\overline{\lambda}_3$	0.1586E−2	0.7994E−2	0.1651E−1	0.2570E−1	0.3566E−1	0.1289E+0	0.2142E+0
	λ_1	0.9893E+0	0.9472E+0	0.8947E+0	0.8421E+0	0.7895E+0	0.4683E+0	
	λ_3	0.6662E−3	0.2818E−2	0.5336E−2	0.7860E−2	0.1046E−1	0.3166E−1	0.3292E−1
	Δ Var	89.7 %	81.3 %	74.5%	68.8 %	63.6 %	34.4%	
4	$\overline{\lambda}_1$	0.3923E−2	0.1738E−1	0.3094E−1	0.4241E−1	0.5269E−1	0.9189E−1	0.9918E−1
	$\overline{\lambda}_2$	0.3668E−2	0.1863E−1	0.3012E−1	0.3959E−1	0.5205E−1	0.1383E+0	0.3017E+0
	$\overline{\lambda}_3$	0.1653E−2	0.9927E−2	0.2844E−1	0.5039E−1	0.7107E−1	0.2041E+0	0.4581E+0
	$\overline{\lambda}_4$	0.7547E−3	0.4069E−2	0.1051E−1	0.1761E−1	0.2419E−1	0.6570E−1	0.1199E+0
	λ_1	0.9897E+0	0.9485E+0	0.8964E+0	0.8442E+0	0.7923E+0	0.4811E+0	
	λ_4	0.3263E−3	0.1501E−2	0.3607E−2	0.5798E−2	0.7723E−2	0.1890E−1	0.2106E−1
	Δ Var	89.3 %	81.0 %	74.4 %	69.0 %	64.1 %	37.8 %	
5	$\overline{\lambda}_1$	0.3884E−2	0.3820E−2	0.7234E−2	0.1033E−1	0.1328E−1	0.2992E−1	0.7433E−1
	$\overline{\lambda}_2$	0.2947E−2	0.1461E−1	0.2721E−1	0.3831E−1	0.4867E−1	0.1033E+0	0.2314E+0
	$\overline{\lambda}_3$	0.1795E−2	0.1845E−1	0.3480E−1	0.4673E−1	0.5905E−1	0.1539E+0	0.4011E+0
	$\overline{\lambda}_4$	0.9299E−3	0.9268E−2	0.2217E−1	0.4000E−1	0.5838E−1	0.1601E+0	0.2169E+0
	$\overline{\lambda}_5$	0.4448E−3	0.3852E−2	0.8595E−2	0.1463E−1	0.2062E−1	0.5277E−1	0.6418E−1
	λ_1	0.9898E+0						
	λ_2		0.9486E+0	0.8970E+0	0.8451E+0	0.7933E+0	0.4845E+0	
	λ_5	0.1934E−3	0.1441E−2	0.2995E−2	0.4892E−2	0.6695E−2	0.1555E−1	0.1216E−1
	Δ Var	89.3 %	81.2 %	74.9 %	69.8 %	65.2 %	40.9 %	
10	λ_2	0.9900E+0						
	λ_3		0.9497E+0	0.8995E+0	.8493E+0	0.7991E+0	0.4980E+0	
	λ_{10}	0.4686E−4	0.2730E−3	0.5036E−3	0.7196E−3	0.9256E−3	0.2047E−2	0.2195E−2
	Δ Var	89.5 %	82.1%	76.5%	72.0%	67.9%	46.4%	
15	λ_4	0.9900E+0	0.9499E+0	0.8998E+0	0.8498E+0	0.7997E+0	0.4994E+0	
	λ_{15}	0.2075E−4	0.8390E−4	0.1513E−3	0.2132E−3	0.2717E−3	0.5872E−3	0.7139E−3
	Δ Var	89.7 %	82.4 %	77.0 %	72.6 %	68.7 %	47.8 %	
20	λ_5	0.9900E+0	0.9500E+0	0.8999E+0	0.8499E+0			
	λ_6					0.7999E+0	0.4997E+0	
	λ_{20}	0.8748E−5	0.3496E−4	0.6264E−4	0.8784E−4	0.1359E−3	0.2889E−3	0.3133E−3
	Δ Var	89.7 %	82.6 %	77.2 %	72.8 %	69.0 %	48.5 %	

Table 26.15 Optimal block sizes for the extreme value distribution and percentage reduction (Δ Var) of the generalized variance compared to right censoring ($\lambda_1 = \cdots = \lambda_{m-1} = 0$) (see Hofmann et al. [444]). The notation cE−k means $c \cdot 10^{-k}$

Table 26.15 shows the optimal censoring plans for selected values of $\tau \in \{.01, .05, .1, .15, .2, .5, 1\}$. For $m = 2, 3, 4, 5$, all sizes $\overline{\lambda}_i$ of the observed blocks and all nonzero censoring block sizes λ_i are given. For $m = 10, 15, 20$, only the nonzero λ_i's are provided. The $\Delta\,\mathsf{Var}$-rows give the percentage reduction in generalized variance of the optimal censoring scheme compared to Type-II right censoring ($\lambda_1 = \cdots = \lambda_{m-1} = 0$). Therefore,

$$\Delta\,\mathsf{Var} = 1 - \frac{\left|\mathsf{Var}\left(\widehat{\mu_{\text{opt}} \atop \vartheta_{\text{opt}}}\right)\right|}{\left|\mathsf{Var}\left(\widehat{\mu_{\text{right}} \atop \vartheta_{\text{right}}}\right)\right|},$$

where $\left(\widehat{\mu_{\text{opt}} \atop \vartheta_{\text{opt}}}\right)$, $\left(\widehat{\mu_{\text{right}} \atop \vartheta_{\text{right}}}\right)$ are the ABLUEs of the overall optimal censoring scheme and the optimal right censoring scheme, respectively. Notice that for fixed τ, the percentage reduction of the generalized variance by the use of optimal progressive censoring is quite stable over m. For small τ (heavy censoring), the reduction is as high as 75–90 %, 30–50 %, and 50–70 % for the extreme value distribution, the Weibull distribution, and the normal distribution, respectively. In the extreme value case, the reduction remains substantial even for τ as large as 0.5 (35–50 %), whereas for Weibull and normal, it becomes quite small as τ increases.

Remark 26.5.9. For $\tau = 1$, the problem reduces to the case of order statistics investigated in Ogawa [694] (see also Ali and Umbach [33]). Ogawa [694] discussed the extreme value distribution and calculated optimal solutions for small m ($m = 2, 3, 4$) which agree with our results. For the Weibull$(1, 3)$-distribution, Hofmann et al. [444] pointed out that the tables in Hassanein [436] seem to be in error. This may have been due to some computational difficulties because the formulas presented in Hassanein [436] yield the same results as obtained in Hofmann et al. [444].

26.5.5 Other Criteria for Optimal Censoring Plans

Optimal progressive censoring has also been discussed in special scenarios with progressively censored data. For instance, optimal Type-I interval censoring has been discussed already in Sect. 18.3. Under step-stress testing, optimal censoring results can be found in Sect. 23.1.1.

Appendix A
Distributions

A.1 Definitions of Distributions

Definition A.1.1 (Uniform distribution). The uniform distribution $\mathsf{U}(a,b)$, with parameters $a,b \in \mathbb{R}$, $a < b$, is defined by the density function

$$f(t) = \frac{1}{b-a}, \quad a < t < b.$$

For $a = 0$ and $b = 1$, it defines the standard uniform distribution $\mathsf{U}(0,1)$.

Definition A.1.2 (Beta distribution). The beta distribution $\mathsf{Beta}(\alpha,\beta)$, with parameters $\alpha,\beta > 0$, is defined by the density function

$$f(t) = \frac{1}{\mathsf{B}(\alpha,\beta)} t^{\alpha-1}(1-t)^{\beta-1}, \quad 0 < t < 1,$$

where $\mathsf{B}(\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}$ denotes the complete beta function and $\Gamma(\cdot)$ denotes the gamma function.

Definition A.1.3 (Power distribution). The power distribution $\mathsf{Power}(\alpha)$, with parameter $\alpha > 0$, is a particular beta distribution $\mathsf{Beta}(\alpha,1)$.

Definition A.1.4 (Reflected power distribution). The reflected power distribution $\mathsf{RPower}(\beta)$, with parameter $\beta > 0$, is a particular beta distribution $\mathsf{Beta}(1,\beta)$.

Definition A.1.5 (Exponential distribution). The two-parameter exponential distribution $\mathsf{Exp}(\mu,\vartheta)$, with parameters $\mu \in \mathbb{R}$ and $\vartheta > 0$, is defined by the density function

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7,

$$f(t) = \frac{1}{\vartheta} e^{-\frac{t-\mu}{\vartheta}}, \quad t > \mu.$$

For $\mu = 0$, it defines the exponential distribution $\mathsf{Exp}(\vartheta)$.

Definition A.1.6 (Weibull distribution). The three-parameter Weibull distribution Weibull(μ, ϑ, β), with parameters $\mu \in \mathbb{R}$ and $\vartheta, \beta > 0$, is defined by the density function

$$f(t) = \frac{\beta}{\vartheta}(t-\mu)^{\beta-1} e^{-\frac{(t-\mu)^\beta}{\vartheta}}, \quad t > \mu.$$

For $\mu = 0$, it defines the two-parameter Weibull distribution Weibull(ϑ, β).

Definition A.1.7 (Gamma distribution). The gamma distribution $\Gamma(\vartheta, \beta)$, with parameters $\vartheta, \beta > 0$, is defined by the density function

$$f(t) = \frac{1}{\Gamma(\beta)\vartheta^\beta} t^{\beta-1} e^{-\frac{t}{\vartheta}}, \quad t > 0,$$

where $\Gamma(\cdot)$ denotes the gamma function.

Definition A.1.8 (χ^2-distribution). The χ^2-distribution $\chi^2(n)$, with degrees of freedom $n \in \mathbb{N}$, is a particular gamma distribution $\Gamma(2, \frac{n}{2})$.

Definition A.1.9 (F-distribution). The F-distribution F(n, m), with parameters $n, m \in \mathbb{N}$, is defined by the density function

$$f(t) = \frac{\Gamma\left(\frac{n+m}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} \left(\frac{n}{m}\right)^{\frac{n}{2}} \frac{t^{\frac{n}{2}-1}}{(1+\frac{n}{m}t)^{\frac{n+m}{2}}}, \quad t > 0,$$

where $\Gamma(\cdot)$ denotes the gamma function.

Definition A.1.10 (Pareto distribution). The Pareto distribution Pareto(α), with parameter $\alpha > 0$, is defined by the density function

$$f(t) = \frac{\alpha}{t^{\alpha+1}}, \quad t \geq 1.$$

Definition A.1.11 (Generalized Pareto distributions). The family of generalized Pareto distribution (GP-distributions) is defined by

$$\mathcal{GPD} = \left\{ F_q \,\middle|\, F_q(t) = \begin{cases} 1-(1-\mathrm{sgn}(q)t)^{1/q}, & q \neq 0 \\ 1-e^{-t}, & q = 0 \end{cases}, t \geq 0, 1-\mathrm{sgn}(q)t > 0 \right\},$$

where $\mathrm{sgn}(q)$ denotes the sign of q, i.e.,

$$\mathrm{sgn}(q) = \begin{cases} 1, & q > 0 \\ 0, & q = 0 \\ -1, & q < 0 \end{cases}.$$

Remark A.1.12. The GP-family covers the

(i) Exponential distribution ($q = 0$),
(ii) Pareto distribution ($q < 0$),
(iii) Reflected power distribution ($q > 0$).

Definition A.1.13 (Lomax distribution). The Lomax distribution Lomax(α), with parameter $\alpha > 0$, is defined by the density function

$$f(t) = \frac{\alpha}{(t+1)^{\alpha+1}}, \quad t \geq 0.$$

Definition A.1.14 (Extreme value distribution (Type I)). The extreme value distribution (Type I) is defined by the density function

$$f(t) = \mathrm{e}^{t-\mathrm{e}^t}, \quad t \in \mathbb{R}.$$

Definition A.1.15 (Extreme value distribution (Type II)/Gumbel distribution). The extreme value distribution (Type II) or Gumbel distribution is defined by the density function

$$f(t) = \mathrm{e}^{-t-\mathrm{e}^{-t}}, \quad t \in \mathbb{R}.$$

Definition A.1.16 (Cauchy distribution). The Cauchy distribution Cauchy (μ, α), with parameters $\alpha > 0, \mu \in \mathbb{R}$, is defined by the density function

$$f(t) = \frac{\alpha}{\pi(1 + \alpha^2(t-\mu)^2)}, \quad t \in \mathbb{R}.$$

For $\mu = 0$ and $\alpha = 1$, it defines the standard Cauchy distribution Cauchy$(0, 1)$.

Definition A.1.17 (Laplace distribution). The Laplace distribution Laplace (μ, ϑ), with parameters $\vartheta > 0, \mu \in \mathbb{R}$, is defined by the density function

$$f(t) = \frac{1}{\vartheta}\mathrm{e}^{-\frac{|t-\mu|}{\vartheta}}, \quad t \in \mathbb{R}.$$

For $\mu = 0$ and $\vartheta = 1$, it defines the standard Laplace distribution Laplace$(0, 1)$.

Definition A.1.18 (Logistic distribution (Type I)). The logistic distribution is defined by the density function

$$f(t) = \frac{\mathrm{e}^t}{(1 + \mathrm{e}^t)^2}, \quad t \in \mathbb{R}.$$

Definition A.1.19 (Normal distribution). The normal distribution $\mathsf{N}(\mu, \sigma^2)$, with parameters $\sigma^2 > 0, \mu \in \mathbb{R}$, is defined by the density function

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}, \quad t \in \mathbb{R}.$$

For $\mu = 0$ and $\sigma^2 = 1$, it defines the standard normal distribution $\mathsf{N}(0, 1)$.

Definition A.1.20 (Log-normal distribution). The log-normal distribution log-$\mathsf{N}(\mu, \sigma^2)$, with parameters $\sigma^2 > 0, \mu \in \mathbb{R}$, is defined by the density function

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}\, t} \mathrm{e}^{-\frac{(\log t-\mu)^2}{2\sigma^2}}, \quad t > 0.$$

Definition A.1.21 (Binomial distribution). The binomial distribution $\mathrm{bin}(n, p)$, with parameters $n \in \mathbb{N}, p \in [0, 1]$, is defined by the probability mass function

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{0, \ldots, n\}.$$

A.2 Definitions and Preliminaries

A.2.1 Quantile Function

Definition A.2.1. Let F be a cumulative distribution function. Then, the quantile function $F^{\leftarrow} : [0, 1] \to \overline{\mathbb{R}}$ is defined by

$$F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad 0 < t < 1.$$

The values at zero and one are defined by the respective one-sided limits $F^{\leftarrow}(0) = F^{\leftarrow}(0+)$ and $F^{\leftarrow}(1) = F^{\leftarrow}(1-)$, respectively.

Lemma A.2.2. Let F be a cumulative distribution function and let $F^{\leftarrow}$ be the associated quantile function. Then, $F^{\leftarrow}$ is a left-continuous and nondecreasing function. Moreover, for all $y \in (0, 1)$ and $x \in \mathbb{R}$, the following assertions hold:

(i) $F(x) \geq y \iff x \geq F^{\leftarrow}(y)$,
(ii) $F(x-) \leq y \iff x \leq F^{\leftarrow}(y+)$,
(iii) $F(F^{\leftarrow}(y)-) \leq y \leq F(F^{\leftarrow}(y))$,
(iv) $F^{\leftarrow}(F(x)) \leq x \leq F^{\leftarrow}(F(x)+)$,
(v) If F is continuous, then $F(F^{\leftarrow}(y)) = y$.

A.2.2 Stochastic Orders

Univariate Stochastic Orders

The univariate stochastic orderings are defined for lifetime distributions with support included in the positive real line. However, they may also be defined on $\mathbb{R}$.

Definition A.2.3 (Stochastic order). Let $X \sim F$, $Y \sim G$ be random variables.

X is said to be stochastically smaller than Y, i.e., $X \leq_{\text{st}} Y$ or $F \leq_{\text{st}} G$, iff

$$\overline{F}(x) \leq \overline{G}(x) \qquad \text{for all } x \geq 0.$$

An important characterization of the stochastic order in terms of nondecreasing functions $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ is that

$$X \leq_{\text{st}} Y \iff E\phi(X) \leq E\phi(Y) \quad \text{for all nondecreasing functions } \phi : \mathbb{R} \longrightarrow \mathbb{R}$$

provided the expectations exist. This property is useful in the definition of the multivariate version of the stochastic order (see p. 577).

Definition A.2.4 (Failure rate/hazard rate order). Let $X \sim F$, $Y \sim G$ be random variables. Then, X is said to be smaller than Y in the hazard rate order, i.e., $X \leq_{\text{hr}} Y$ or $F \leq_{\text{hr}} G$, iff

$$\overline{F}(x)\overline{G}(y) \leq \overline{F}(y)\overline{G}(x) \qquad \text{for all } 0 \leq y \leq x.$$

The preceding condition means that the ratio $\frac{\overline{F}(x)}{\overline{G}(x)}$ is nonincreasing in $x \geq 0$, where $\frac{a}{0}$ is defined to be ∞. If F and G are absolutely continuous cumulative distribution functions with density functions f and g, respectively, then hazard rate ordering is equivalent to increasing hazard rates

$$\lambda_F(x) = \frac{f(x)}{1 - F(x)} \leq \frac{g(x)}{1 - G(x)} = \lambda_G(x) \qquad \text{for all } x \geq 0.$$

Definition A.2.5 (Reversed hazard rate order). Let $X \sim F$, $Y \sim G$ be random variables. Then, X is said to be smaller than Y in the reversed hazard rate order, i.e., $X \leq_{\text{rh}} Y$ or $F \leq_{\text{rh}} G$, iff

$$F(x)G(y) \leq F(y)G(x) \qquad \text{for all } 0 \leq y \leq x.$$

The preceding condition means that the ratio $\frac{F(x)}{G(x)}$ is nonincreasing in $x \geq 0$, where $\frac{a}{0}$ is defined to be ∞.

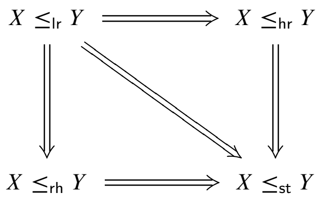

Fig. A.1 Relations between stochastic orders.

Definition A.2.6 (Likelihood ratio order). Let $X \sim F$, $Y \sim G$ be random variables with absolutely continuous cumulative distribution functions and density functions f and g, respectively. Then, X is said to be smaller than Y in the likelihood ratio order, i.e., $X \leq_{\mathsf{lr}} Y$ or $F \leq_{\mathsf{lr}} G$, iff

$$f(x)g(y) \leq f(y)g(x) \qquad \text{for all } 0 \leq y \leq x.$$

The preceding condition means that the ratio $\frac{f(x)}{g(x)}$ is nonincreasing in $x \geq 0$.

These orders are related as depicted in Fig. A.1 (cf. Shaked and Shanthikumar [799] and Müller and Stoyan [659]).

An intuitive approach to some partial orders is to compare quantile differences (cf. Lewis and Thompson [584], Shaked [797, 798]) as below.

Definition A.2.7 (Dispersive order). Let $X \sim F$, $Y \sim G$ be random variables. Then, X is said to be smaller than Y in the dispersive order, i.e., $X \leq_{\mathsf{disp}} Y$ or $F \leq_{\mathsf{disp}} G$, iff

$$F^{\leftarrow}(x) - F^{\leftarrow}(y) \leq G^{\leftarrow}(x) - G^{\leftarrow}(y) \qquad \text{for all } 0 < y < x < 1.$$

If G is continuous, then the preceding condition can be written as

$$F^{\leftarrow}(G(x)) - F^{\leftarrow}(G(y)) \leq x - y \qquad \text{for all } 0 \leq y < x. \tag{A.1}$$

This means that $F^{\leftarrow}(G(x)) - x$ is a nonincreasing function in $x \geq 0$. In fact, this property defines the so-called tail ordering introduced by Doksum [343]. It was shown by Deshpande and Kochar [336] that the tail order and dispersive order coincide under weak conditions (cf. Kamps [498, p. 181]).

Lorenz ordering of distributions is defined via the Lorenz curve $L_X : [0,1] \longrightarrow [0,1]$ of a random variable $X \sim F$ given by

$$L_X(u) = \frac{\int_0^u F^{\leftarrow}(t)\,dt}{\int_0^1 F^{\leftarrow}(t)\,dt}, \quad u \in [0,1].$$

Definition A.2.8 ((Increasing) convex order, Lorenz order). Let X and Y be nonnegative random variables with finite expectations and cumulative distribution functions F and G, respectively. Then,

(i) X is smaller than Y in convex order, i.e., $X \leq_{\mathsf{cx}} Y$ or $F \leq_{\mathsf{cx}} G$, if $Ef(X) \leq Ef(Y)$ for all convex functions such that the expectations exist;
(ii) X is smaller than Y in increasing convex order, i.e., $X \leq_{\mathsf{icx}} Y$ or $F \leq_{\mathsf{icx}} G$, if $Ef(X) \leq Ef(Y)$ for all increasing convex functions such that the expectations exist;
(iii) X and Y are ordered w.r.t. the Lorenz order, i.e., $X \leq_{\mathsf{L}} Y$ or $F \leq_{\mathsf{L}} G$, iff

$$L_X(u) \geq L_Y(u) \quad \text{for all } u \geq 0.$$

The Lorenz order $\leq_{\mathsf{L}}$ is related to the convex order $\leq_{\mathsf{cx}}$ of random variables as follows:

$$X \leq_{\mathsf{L}} Y \iff \frac{X}{EX} \leq_{\mathsf{cx}} \frac{Y}{EY}$$

indicating that convex ordering supposes equal expectations of the considered random variables (cf. Shaked and Shanthikumar [799, p. 116]).

A.2.2.1 Multivariate Stochastic Orderings

Definition A.2.9 (Multivariate stochastic order). Let $\mathbf{X} = (X_1, \ldots, X_n)'$, $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ be random vectors. Then, $\mathbf{X}$ is said to be stochastically smaller than $\mathbf{Y}$, i.e., $\mathbf{X} \leq_{\mathsf{st}} \mathbf{Y}$ or $F_{\mathbf{X}} \leq_{\mathsf{st}} F_{\mathbf{Y}}$, iff

$$E\phi(\mathbf{X}) \leq E\phi(\mathbf{Y}) \quad \text{for all nondecreasing functions } \phi : \mathbb{R}^n \longrightarrow \mathbb{R}$$

provided the expectations exist.

Definition A.2.10 (Multivariate likelihood ratio order). Let $\mathbf{X} = (X_1, \ldots, X_n)'$, $\mathbf{Y} = (Y_1, \ldots, Y_n)'$ be random vectors with density functions $f_{\mathbf{X}}$ and $f_{\mathbf{Y}}$. Then, $\mathbf{X}$ is said to be smaller than $\mathbf{Y}$ in the multivariate likelihood ratio order, i.e., $\mathbf{X} \leq_{\mathsf{lr}} \mathbf{Y}$ or $F_{\mathbf{X}} \leq_{\mathsf{lr}} F_{\mathbf{Y}}$, iff

$$f^{\mathbf{X}}(\mathbf{x}_n) f^{\mathbf{Y}}(\mathbf{y}_n) \leq f^{\mathbf{X}}(\mathbf{x}_n \wedge \mathbf{y}_n) f^{\mathbf{Y}}(\mathbf{x}_n \vee \mathbf{y}_n) \tag{A.2}$$

for all $\mathbf{x} = (x_1, \ldots, x_n)', \mathbf{y} = (y_1, \ldots, y_n)' \in \mathbb{R}^n$.

Orderings of Real Vectors

The following partial orders for real vectors in the n-dimensional Euclidean space $\mathbb{R}^n$ are used in the book.

Definition A.2.11 (Majorization). Let $\mathbf{x} = (x_1, \ldots, x_m)', \mathbf{y} = (y_1, \ldots, y_m)' \in \mathbb{R}^m$ be vectors. Then, $\mathbf{x}$ is said to majorize $\mathbf{y}$, i.e., $\mathbf{y} \preccurlyeq_{\mathrm{m}} \mathbf{x}$, iff

$$\sum_{i=1}^{j} x_{i:m} \leq \sum_{i=1}^{j} y_{i:m} \text{ for } j = 1, \ldots, m-1 \text{ and } \sum_{i=1}^{m} x_{i:m} = \sum_{i=1}^{m} y_{i:m};$$

$\mathbf{x}$ is said to weakly majorize $\mathbf{y}$, i.e. $\mathbf{y} \preccurlyeq_{\mathrm{w}} \mathbf{x}$, iff

$$\sum_{i=1}^{j} x_{i:m} \leq \sum_{i=1}^{j} y_{i:m} \text{ for } j = 1, \ldots, m.$$

A different ordering concept was introduced by Bon and Păltănea [213] (see also Khaledi and Kochar [520]).

Definition A.2.12 (p-larger order). Let $\mathbf{x} = (x_1, \ldots, x_m)', \mathbf{y} = (y_1, \ldots, y_m)' \in [0, \infty)^m$ be nonnegative vectors. Then, $\mathbf{x}$ is said to p-majorize $\mathbf{y}$, i.e., $\mathbf{y} \preccurlyeq_{\mathrm{p}} \mathbf{x}$, iff

$$\prod_{i=1}^{j} x_{i:m} \leq \prod_{i=1}^{j} y_{i:m} \text{ for } j = 1, \ldots, m.$$

As indicated by Bon and Păltănea [213], majorization implies p-majorization.

Appendix B
Additional Demonstrative Data Sets

Many data sets have been used throughout for illustrating various inferential procedures. In the following, we introduce some further data that are available in the literature.

B.1 Progressively Type-II Censored Data

Herd [440] presented the following progressively Type-II censored data set which has also been revisited by Sarhan and Greenberg [776, p. 355] (see also Cohen [273, Sect. 4.5]).

Data B.1.1 (Herd's gyroscope data). Eleven gyroscopes are put on a progressively censored life test with censoring scheme $\mathscr{R} = (3, 2^{*2}, 0)$. Therefore, four failures are observed and seven units are progressively censored. The data set is given by

i	1	2	3	4
$x_{i:4:11}$	34	133	169	237
R_i	3	2	2	0

The following data is taken from Cohen [270]. The original sample comprises 100 simulated data points from a three-parameter Weibull distribution Weibull$(100, 100, 2)$.

Data B.1.2 (Cohen's simulated Weibull data). The original sample comprises 100 simulated data points from a three-parameter Weibull distribution Weibull$(100, 100, 2)$. It includes 68 life-span observations, 32 values are progressively Type-II censored according to the censoring scheme $\mathscr{R} = (0^{*5}, 10, 0^{*33}, 15, 0^{*27}, 7)$. The experiment is terminated at the failure time $X_{68:68:100} = 249.35$ with seven survivors.

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7,

109.12	130.53	144.09	158.31	177.19	198.11	222.11
113.37	131.98	148.83	158.92	180.57	199.23	224.83
117.73	133.14	150.23	160.13	181.99	203.27	227.27
119.56	134.52	150.79	161.31	184.02	206.55	230.88
119.82	135.73	151.88	162.09	185.43	208.76	235.14
*124.63	136.71	153.07	165.45	187.21	210.69	237.43
125.21	137.88	154.18	166.62	189.77	213.32	246.08
126.93	138.63	154.97	168.23	191.63	215.08	*249.35
128.25	141.11	155.26	169.98	194.88	218.43	
129.41	142.33	156.82	*174.22	196.91	219.37	

* indicates failure times with progressively censored items

The following simulated progressive Type-II censored data sets are given in Balakrishnan and Aggarwala [86].

Data B.1.3 (Laplace data from Balakrishnan and Aggarwala [86], p. 33). The following progressively Type-II censored sample has been simulated from a Laplace$(25, 5)$-distribution with $n = 20$, $m = 10$, and censoring scheme $\mathscr{R} = (2, 0^{*2}, 2, 0^{*3}, 2, 0, 4)$:

19.2117	21.9736	23.4178	23.6625	23.8022
24.2302	25.6207	25.8699	26.4800	27.5534

Data B.1.4 (Exponential data from Balakrishnan and Aggarwala [86], p. 40). The following general progressively Type-II censored sample has been simulated from an exponential distribution Exp$(25, 10)$ with $n = 50$ and censoring scheme $\mathscr{R}_{\triangleleft}5 = (2, 1, 0, 2, 1, 3, 2, 3, 4, 3, 0, 2, 1, 2, 4)$ $(r = 5, m = 15)$:

25.99609	26.17323	26.55884	26.65558	27.32842
27.52826	28.58114	28.58350	28.68850	29.09515
29.17521	29.47387	29.61337	33.44267	35.74206

Data B.1.5 (Progressively Type-II censored data from Spinelli and Stephens [821]). The following progressively Type-II censored sample has been generated by Balakrishnan et al. [131] from data reported in Spinelli and Stephens [821] using the censoring scheme $\mathscr{R} = (0, 2, 0^{*2}, 2, 0^{*5}, 2^{*2}, 0^{*4}, 1^{*2}, 0, 2)$ $(n = 32, m = 20)$. The original data consist of 32 observations on measurements of modulus of repute of wood beams. The quantity of interest measures the breaking strength of lumber.

43.19	49.44	51.55	56.63	67.27
78.47	86.59	90.63	94.38	98.21
98.39	99.74	100.22	103.48	105.54
107.13	108.14	108.94	110.81	116.39

Data B.1.6 (Nelson's progressively Type-II censored data as given in Pakyari and Balakrishnan [701]). Pakyari and Balakrishnan [701] generated a second progressively Type-II censored sample from Data 1.1.2 with censoring scheme $\mathscr{R} = (0, 1, 0, 1, 0, 1, 0, 1, 0, 1)$. The values are given as log-times.

i	1	2	3	4	5
x_i	−1.3093	−0.9163	−0.3711	−0.2357	1.0116
R_i	0	1	0	1	0
i	6	7	8	9	10
x_i	1.3635	2.7682	3.3250	3.9748	5.3711
R_i	1	0	1	0	1

Data B.1.7 (King's progressively Type-II censored wire strength connection data). Pakyari and Balakrishnan [701] generated a progressively Type-II censored sample from wire strength connection data reported by King [533] (see also Nelson [676], Table 5.1, p. 111). They applied the censoring scheme $\mathscr{R} = (0, 2, 1, 0, 3, 0^{*2}, 2, 0, 2)$ to generate $m = 10$ observations from a total of $n = 20$ available data points. The original data with sample size $n = 23$ were adjusted by eliminating three observations due to validity suspicion (see Nelson [676]).

i	1	2	3	4	5
x_i	550	750	950	1150	1150
R_i	0	2	1	0	3
i	6	7	8	9	10
x_i	1150	1350	1450	1550	1850
R_i	0	0	2	0	2

B.2 Progressively Type-I Censored Data

Nelson [674] introduced the following progressively Type-I censored data (see also Nelson [676, p. 318]).

Data B.2.1 (Nelson's fan data). Seventy generator fans were placed on a life test. The failure and censoring times as well as the censoring numbers are given in Table B.1.

Failure time x_j	Censoring time T_j	R_j	Failure time x_j	Censoring time T_j	R_j
4,500			46,000		
	4,600	1		48,500	4
11,500				50,000	3
11,500			61,000		
	15,600	1		61,000	3
16,000				63,000	1
	16,600	1		64,500	2
	18,500	5		67,000	1
	20,300	3		74,500	1
20,700				78,000	2
20,700				81,000	2
20,800				82,000	1
	22,000	1		85,000	3
	30,000	4	87,500		
31,000				87,500	2
	32,000	1		94,000	1
34,500				99,000	1
	37,500	2		101,000	3
	41,500	4		115,000	1
	43,000	4			

Table B.1 Nelson's [674] fan data (times are measured in hours)

Data B.2.2 (Nelson's progressively Type-I censored insulating fluid data). The measurements result from $n = 19$ breakdown times (in minutes) of an insulating fluid at 36 kV. The failure and censoring times as well as the censoring numbers are given in Tables B.2 and B.3. In Table B.2, the effectively applied censoring plan is $\mathscr{R} = (3, 0, 3, 5)$ and the censoring times are $T_j = 2j$, $j = 1, \ldots, 5$. The sample size and the number of observations are given by $n = 19$ and $m = d_{\bullet 4} = 8$.

Failure time x_j	Censoring time T_j	ζ_j	Failure time x_j	Censoring time T_j	ζ_j
0.19				4	0
0.78			4.67		
0.96			4.85		
1.31				6	3
	2	3	7.35		
2.78				8	5

Table B.2 Nelson's [677] insulating fluid data progressively Type-I censored as in Balakrishnan et al. [150] (times are measured in minutes)

In Table B.3, the initially planned censoring plan is $\mathscr{R}^0 = (3, 4)$ to be carried out at censoring times $T_1 = 1$, $T_2 = 5$. The termination of the experiment is scheduled for $T_3 = 15$. The effectively applied censoring plan is $\mathscr{R} = (3, 4, 3)$. The sample size and the number of observations are given by $n = 19$ and $m = d_{\bullet 4} = 9$.

Failure time x_j	Censoring time T_j	ζ_j	Failure time x_j	Censoring time T_j	ζ_j
0.19			4.67		
0.78			4.85		
0.96				5	4
	1	3	6.50		
1.31			12.06		
4.15				15	3

Table B.3 Nelson's [677] insulating fluid data progressively Type-I censored at $T_1 = 1$, $T_2 = 5$, and $T_3 = 15$ (times are measured in minutes)

Data B.2.3 (Wingo's life test data). The data resulted from a life test of $n = 50$ specimens. $m = 33$ failure times were observed (see Table B.4), whereas 17 items were progressively censored during the experiment. The censoring times are given by $T_1 = 70$, $T_2 = 80$, $T_3 = 99$, $T_4 = 121$, and $T_5 = 150$. The effectively applied censoring plan was $\mathscr{R} = (4, 5, 4, 3, 1)$.

37	55	64	72	74	87	88	89	91	92	94
95	97	98	100	101	102	102	105	105	107	113
117	120	120	120	122	124	126	130	135	138	182

Table B.4 Wingo's [898] life-test data (times are measured in hours)

Data B.2.4 (Wingo's pain relief data). The data were taken from a clinical trial that was conducted to assess the effectiveness of an anesthetic antibiotic ointment in relieving pain caused by superficial skin wounds. $n = 30$ patients were included in the study where measurements of $m = 20$ patients are available and given in Table B.5. Some patients dropped out of the study at censoring times $T_1 = 0.25$, $T_2 = 0.50$, and $T_3 = 0.75$. A final termination was not fixed so that we assume T_4 to be very large ensuring that no observation is right censored. The effectively applied censoring plan was $\mathscr{R} = (5, 1, 4, 0)$.

0.828	0.881	1.138	0.879	0.554	0.653	0.698	0.566	0.665	0.917
0.529	0.786	1.110	0.866	1.037	0.788	1.050	0.899	0.683	0.829

Table B.5 Wingo's [901] pain relief data (times are measured in hours)

Data B.2.5 (Montanari and Cacciari's XLPE-isolated cable data). Montanari and Cacciari [655] considered testing of XLPE-insulated cable models which were subjected to electrical and combined thermal–electrical stresses. One of the aims of the study was the analysis of aging mechanisms. The data are given in Table B.6.

Failure times x_i	445	479	489	607	692	969				
Censoring times T_i	445	453	479	489	588	607	766	780	969	1,121
Censoring numbers R_i	1	1	1	2	1	1	2	1	1	1

Table B.6 Montanari and Cacciari's [655] XLPE-isolated cable data (times are measured in hours)

These data suggest that the censoring procedure might be a mixture of Type-I and Type-II progressive censoring since at the time of the first, second, third, and sixth failures, units have been removed from the life test. However, the data will be handled as Type-I censored data.

Notation

The following notation are used throughout the book. In selected cases, a page number is provided in order to refer to the definition.

Notation	Explanation	Page
Progressively censored order statistics		
$X_{i:m:n}^{\mathscr{R}}, X_{i:m:n}$	ith progressively Type-II censored order statistic based on censoring scheme $\mathscr{R}$	
$U_{i:m:n}^{\mathscr{R}}, U_{i:m:n}$	ith uniform progressively Type-II censored order statistic based on censoring scheme $\mathscr{R}$	
$Z_{i:m:n}^{\mathscr{R}}, Z_{i:m:n}$	ith exponential progressively Type-II censored order statistic based on censoring scheme $\mathscr{R}$	
$\mathbf{X}^{\mathscr{R}}, \mathbf{U}^{\mathscr{R}}, \mathbf{Z}^{\mathscr{R}}$	Vector of progressively Type-II censored order statistics	
$S_i^{\mathscr{R}}$	ith normalized spacing of progressively Type-II censored order statistic based on censoring scheme $\mathscr{R}$	26
$S_i^{\star\mathscr{R}}$	ith spacing of progressively Type-II censored order statistic based on censoring scheme $\mathscr{R}$	28
$S_{i,j}^{\star\mathscr{R}}$	Generalized spacing of progressively Type-II censored order statistic based on censoring scheme $\mathscr{R}$	46
$\mathscr{R}, \mathscr{S}$	Censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$, etc.	
$\mathscr{O}_k$	One-step censoring scheme alternatively represented as $(0^{*k-1}, n-m, 0^{*m-k})$	7
$\mathscr{R}_{\rhd r}$	Right truncated censoring scheme $(R_1, \ldots, R_{r-1})$	10
$\mathscr{R}_{\lhd r}$	Left truncated censoring scheme $(R_{r+1}, \ldots, R_m)$	10
$\mathscr{C}_{m,n}^m$	Set of admissible censoring schemes in progressive Type-II censoring	5
$\gamma_i, \gamma_i(\mathscr{R})$	$\gamma_i = \sum_{j=i}^m (R_j + 1)$, $1 \le i \le m$, for a censoring scheme $\mathscr{R} = (R_1, \ldots, R_m)$	
c_{r-1}	$\prod_{j=1}^r \gamma_j$	
$c(\mathscr{R})$	$\prod_{j=1}^m \gamma_j(\mathscr{R})$	

(continued)

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7,

(continued)

Notation	Explanation	Page
$a_{j,r}$	$\prod_{\substack{i=1\\i\neq j}}^{r} \frac{1}{\gamma_i-\gamma_j}$, $1 \le j \le r$	
$a_{j,k_2}^{(k_1)}$	$\prod_{\substack{\nu=k_1+1\\\nu\neq j}}^{k_2} \frac{1}{\gamma_\nu-\gamma_j}$, $k_1+1 \le j \le k_2$	
$X_{i:n}$	ith order statistic in a sample of size n	
$U_{i:n}$	ith uniform order statistic in a sample of size n	
$Z_{i:n}$	ith exponential order statistic in a sample of size n	
$S_{i:n}$	ith spacing in a sample of size n	
$X_{i:m:n}^{\mathscr{R},T}$	ith progressively Type-I censored order statistic with censoring scheme $\mathscr{R}$ and threshold T	
$X_i^{(\mathrm{I})}, Z_i^{(\mathrm{I})}$	ith (exponential) Type-I progressively hybrid censored order statistic	
$W_i^{(\mathrm{I})}$	ith normalized spacing of exponential Type-I progressively hybrid censored order statistic	
$X_i^{(\mathrm{II})}, Z_i^{(\mathrm{II})}$	ith (exponential) Type-II progressively hybrid censored order statistic	
$W_i^{(\mathrm{II})}$	ith normalized spacing of exponential Type-II progressively hybrid censored order statistic	
$\mathbf{X}, \mathbf{X}_n$	Vector of random variables $X_1, \dots, X_n$	
f^X, F^X	Density function/cumulative distribution function of a random variable X	
$F_{\exp}$	Cumulative distribution function of an exponential distribution	
$f_{\chi^2(d)}, F_{\chi^2(d)}$	Density function/cumulative distribution function of a χ^2-distribution with d degrees of freedom	
$f_{j:m:n}, f_{j:m:n}^{\mathscr{R}}$	Density function of $X_{j:m:n}^{\mathscr{R}}$	
$F_{j:m:n}, F_{j:m:n}^{\mathscr{R}}$	Cumulative distribution function of $X_{j:m:n}^{\mathscr{R}}$	
$f_{j:n}, F_{j:n}$	Density function/cumulative distribution function of $X_{j:n}$	
$f_{1,\dots,m:m:n}, F_{1,\dots,m:m:n}$	Joint density function/cumulative distribution function of $X_{1:m:n}^{\mathscr{R}}, \dots, X_{m:m:n}^{\mathscr{R}}$	
$\mathscr{I}(\mathbf{X}^{\mathscr{R}}; \theta), \mathscr{I}(\mathbf{X}^{\mathscr{R}}; \boldsymbol{\theta})$	Fisher information	201
$\mathscr{H}_{1,\dots,m:m:n}^{\mathscr{R}}$	Joint entropy of $X_{1:m:n}^{\mathscr{R}}, \dots, X_{m:m:n}^{\mathscr{R}}$	216
$\mathscr{I}_{\mathscr{R}}(f \,\Vert\, g)$	Kullback–Leibler information	221
Distributions		
$\mathsf{U}(a,b)$	Uniform distribution	571
$\mathsf{Beta}(\alpha,\beta)$	Beta distribution	571
$\mathsf{Power}(\alpha)$	Power distribution	571
$\mathsf{RPower}(\beta)$	Reflected power distribution	571
$\mathsf{Exp}(\mu,\vartheta)$	Exponential distribution	572
$\Gamma(\vartheta,\beta)$	Gamma distribution	572
$\mathsf{N}(\mu,\sigma^2)$	Normal distribution	574
$\chi^2(n)$	χ^2-distribution	572

(continued)

(continued)

Notation	Explanation	Page
$\mathsf{F}(n, m)$	F-distribution	572
$\mathsf{Pareto}(\alpha)$	Pareto distribution	572
$\mathsf{Cauchy}(\mu, \alpha)$	Cauchy distribution	573
$\mathsf{Laplace}(\mu, \vartheta)$	Laplace distribution	573
$\mathsf{bin}(n, p)$	Binomial distribution	574
Special functions		
$\exp(t)$, e^t	Exponential function	
$\log(t)$	Natural logarithm	
$\Gamma(\alpha)$	Gamma function	
$\mathsf{IG}(t;\alpha)$	Incomplete gamma function ratio defined as $\mathsf{IG}(t;\alpha) = \frac{1}{\Gamma(\alpha)} \int_0^t z^{\alpha-1}\mathrm{e}^{-z} dz$, $t \geq 0$	
$\mathsf{B}(\alpha, \beta)$	Beta function	
$\mathsf{B}_t(\alpha, \beta)$	Incomplete beta function defined as $\mathsf{B}_t(\alpha, \beta) = \int_0^t x^{\alpha-1}(1-x)^{\beta-1} dx$, $0 < t < 1$	
$n!$	n factorial defined by $\prod_{j=1}^n j$, $n \in \mathbb{N}_0$, where $\prod_{j=1}^0 j = 1$	
$\binom{n}{k}$	Binomial coefficient $\frac{n!}{k!(n-k)!}$, $k, n \in \mathbb{N}_0$, $k \leq n$	
$\binom{n}{k_1,\ldots,k_r}$	Multinomial coefficient $\frac{n!}{k_1!\cdots k_r!}$, $k_i, n \in \mathbb{N}_0$, $\sum_{i=1}^r k_i = n$	
$F(a, b; c; t)$	Hypergeometric function defined by $F(a, b; c; t) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{x^j}{j!}$, where Pochhammer's symbol $(\cdot)_j$ is defined by $(x)_j = \prod_{i=1}^j (x + i - 1)$, $j \in \mathbb{N}$, $x \in \mathbb{R}$; $(x)_0 = 1$	
$B_k(s \mid a_k, \ldots, a_1)$	Univariate B-Spline B_k of degree k with knots $a_k < \cdots < a_1$	131
$\mathbb{1}_A$	Indicator function on the set A	
$[x]_+ = \max(x, 0)$	Positive part of x	
$[x]_- = \min(x, 0)$	Negative part of x	
$\lfloor x \rfloor$	Denotes the largest integer k satisfying $k \leq x$	
$\mathrm{sgn}(x)$	Sign of x	
$h[x_j, \ldots, x_\nu]$	Divided differences of order $\nu - j$ at $x_1 > \cdots > x_m$ for function h	43
Sets		
$\mathbb{N}$	Integers $\{1, 2, 3, \ldots\}$	
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \ldots\}$	
$\mathbb{R}$	Real numbers	
$\mathbb{R}^n$	n-fold Cartesian product of $\mathbb{R}$	
$\mathbb{R}^n_{\leq}$	$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_1 \leq \cdots \leq x_n\}$	
$\mathfrak{S}_n$	Set of all permutations of $(1, \ldots, n)$	
$(\Omega, \mathfrak{A})$	Measurable space with σ-algebra $\mathfrak{A}$	
$(\Omega, \mathfrak{A}, P)$	Probability space with σ-algebra $\mathfrak{A}$ and probability measure P	

(continued)

(continued)

Notation	Explanation	Page
Symbols		
$X \sim F$	X is distributed according to a cumulative distribution function F	
$\overset{\text{iid}}{\sim}$	Independent and identically distributed	
$\overset{\text{d}}{\longrightarrow}$	Convergence in distribution	
$\overset{\text{P}}{\longrightarrow}$	Convergence in probability	
$\overset{\text{a.e.}}{\longrightarrow}$	Convergence almost everywhere	
$F^{\leftarrow}$	Quantile function of F	
$\overline{F}$	Survival/reliability function $\overline{F} = 1 - F$	
$\alpha(F)$	Left endpoint of the support of F	
$\omega(F)$	Right endpoint of the support of F	
$\xi_p = F^{\leftarrow}(p)$, $p \in (0, 1)$,	pth quantile of F	
λ^r	r-dimensional Lebesgue measure	
#	Counting measure	
ε_T	One-point distribution/Dirac measure in T	
$\text{tr}(A)$	Trace of a matrix A	
$\det(A)$	Determinant of a matrix A	
$\text{med}(F)$	Median of cumulative distribution function F	
$\mathbf{a}_k$	$(a_1, \ldots, a_k) \in \mathbb{R}^k$	
(a^{*k})	$(a, \ldots, a) \in \mathbb{R}^k$	
$\mathbb{1}$	$\mathbb{1} = (1, \ldots, 1)$	
$g(t-)$	$\lim_{x \nearrow t} g(x)$	
$g(t+)$	$\lim_{x \searrow t} g(x)$	
const	Represents all additive terms of a function which do not contain the variable of the function	
Orderings		
$\overset{\text{d}}{=}$	Identical distribution	
$\leq_{\text{st}}$	Stochastic order	
$\leq_{\text{hr}}$	Hazard rate order	
$\leq_{\text{rh}}$	Reversed hazard rate order	
$\leq_{\text{lr}}$	Likelihood ratio order	
$\leq_{\text{disp}}$	Dispersive order	
$\leq_{\text{cx}}$	Convex order	
$\leq_{\text{icx}}$	Increasing convex order	
$\leq_{\text{L}}$	Lorenz order	
$\preccurlyeq_{\text{m}}$	Majorization order	
$\preccurlyeq_{\text{w}}$	Weak majorization order	
$\preccurlyeq_{\text{p}}$	p-larger order	
Operations		
$x \wedge y$, $x, y \in \mathbb{R}$	$\min\{x, y\}$	
$x \vee y$, $x, y \in \mathbb{R}$	$\max\{x, y\}$	
$\mathbf{x} \wedge \mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$	$(x_1 \wedge y_1, \ldots, x_n \wedge y_n)'$	
$\mathbf{x} \vee \mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$	$(x_1 \vee y_1, \ldots, x_n \vee y_n)'$	
$a_{\bullet r}$	$\sum_{j=1}^{r} a_j$ for $\mathbf{a}_r \in \mathbb{R}^r$	
$A \otimes B$	Kronecker product of matrices A and B	

(continued)

(continued)

Notation	Explanation	Page
Abbreviations		
IID	Independent and identically distributed	
INID	Independent but not necessarily identically distributed	
a.s.	Almost surely	
a.e.	Almost everywhere	
e.g.	For example/exempli gratia	
i.e.	id est	
w.r.t.	With respect to	
IFR/DFR	Increasing/decreasing failure rate	
IFRA/DFRA	Increasing/decreasing failure rate on average	
NBU/NWU	New better/worse than used	
BLUE	Best linear unbiased estimator	
ABLUE	Asymptotic best linear unbiased estimator	
BLIE	Best linear invariant estimator	
BLEE	Best linear equivariant estimator	
MLE	Maximum likelihood estimator	
AMLE	Approximate maximum likelihood estimator	
MMLE	Modified maximum likelihood estimator	
UMVUE	Uniform minimum unbiased estimator	
MRE	Minimum risk equivariant estimator	
IE	Inverse estimator	
BUP	Best unbiased predictor	
MLP	Maximum likelihood predictor	
MMLP	Modified maximum likelihood predictor	
AMLP	Approximate maximum likelihood predictor	
MUP	Median unbiased predictor	
CMP	Conditional median predictor	
BLUP	Best linear unbiased predictor	
BLEP	Best linear equivariant predictor	
PLF	Predictive likelihood function	
HPD	Highest probability density	
MSE	Mean squared error	
MSPE	Mean squared predictive error	
OSP	One-step censoring plan	
FSP	First-step censoring plan	
AQL	Acceptance quality level	
RQL	Rejectable quality level	

References

1. Aarset MV (1987) How to identify a bathtub hazard rate. IEEE Trans Reliab R-36:106–108
2. Abdel-Aty Y, Franz J, Mahmoud MAW (2007) Bayesian prediction based on generalized order statistics using multiply type-II censoring. Statistics 41:495–504
3. Abdel-Hamid AH (2009) Constant-partially accelerated life tests for Burr type-XII distribution with progressive type-II censoring. Comput Stat Data Anal 53:2511–2523
4. Abdel-Hamid AH, AL-Hussaini EK (2011) Inference for a progressive stress model from Weibull distribution under progressive type-II censoring. J Comput Appl Math 235: 5259–5271
5. Abdel-Hamid AH, AL-Hussaini EK (2014) Bayesian prediction for Type-II progressive censored data from the Rayleigh distribution under progressive-stress model. J Stat Comput Simul 84:1297–1312
6. Abo-Eleneen ZA (2007) Fisher information and optimal schemes in progressive Type-II censored samples. Model Assist Stat Appl 2:153–163
7. Abo-Eleneen ZA (2008) Fisher information in type II progressive censored samples. Comm Stat Theory Meth 37:682–691
8. Abo-Eleneen ZA (2011) The entropy of progressively censored samples. Entropy 13:437–449
9. Abo-Eleneen ZA (2012) A novel approach for optimal schemes in progressive censoring plans. J Comm Comput 9:426–433
10. Abushal TA, Soliman AA (2013) Estimating the Pareto parameters under progressive censoring data for constant-partially accelerated life tests. J Stat Comput Simul (to appear)
11. Aggarwala R (2001) Progressive interval censoring: some mathematical results with applications to inference. Comm Stat Theory Meth 30:1921–1935
12. Aggarwala R, Balakrishnan N (1996) Recurrence relations for single and product moments of progressive type-II right censored order statistics from exponential and truncated exponential distributions. Ann Inst Stat Math 48:757–771
13. Aggarwala R, Balakrishnan N (1998) Some properties of progressive censored order statistics from arbitrary and uniform distributions with applications to inference and simulation. J Stat Plan Infer 70:35–49
14. Aggarwala R, Balakrishnan N (2002) Maximum likelihood estimation of the Laplace parameters based on progressive Type-II censored samples. In: Balakrishnan N (ed) Advances on methodological and applied aspects of probability and statistics. Taylor & Francis, London, pp 159–167
15. Aggarwala R, Childs A (2000) Conditional inference for the parameters of Pareto distributions when observed samples are progressively censored, Chap. 17. In: Balakrishnan N, Melas VB, Ermakov S (eds) Advances in stochastic simulation methods. Birkhäuser, Boston, pp 293–302
16. Ahmadi J (2007) Some results based on entropy properties of progressive Type-II censored data. J Stat Res Iran 4:191–202

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7,

17. Ahmadi J, MirMostafaee SM, Balakrishnan N (2012) Bayesian prediction of k-record values based on progressively censored data from exponential distribution. J Stat Comput Simul 82:51–62
18. Ahmadi MV, Doostparast M, Ahmadi J (2013) Estimating the lifetime performance index with Weibull distribution based on progressive first-failure censoring scheme. J Comput Appl Math 239:93–102
19. Ahmadi K, Rezaei M, Yousefzadeh F (2014) Estimation for the generalized half-normal distribution based on progressive type-II censoring. J Stat Comput Simul (to appear)
20. Ahmed EA (2014) Bayesian estimation based on progressive Type-II censoring from two-parameter bathtub-shaped lifetime model: an Markov chain Monte Carlo approach. J Appl Stat 41:752–758
21. Ahsanullah M (1981) On characterizations of the exponential distribution by spacings. Stat Papers 22:316–320
22. Ahsanullah M (1984) A characterization of the exponential distribution by higher order gap. Metrika 31:323–326
23. Ahsanullah M (1996) Linear prediction of generalized order statistics from two parameter exponential distributions. J Appl Stat Sci 3:1–9
24. Ahsanullah M, Hamedani G (2010) Exponential distribution: theory and methods. Nova Science Publisher, New York
25. Akritas MG, Basak I, Lee MH (1993) Robust estimation under Type II censoring. Aust N Z J Stat 35:303–318
26. Al-Aboud FM (2009) Bayesian estimations for the extreme value distribution using progressive censored data and asymmetric loss. Int Math Forum 4:1603–1622
27. AL-Hussaini EK (1999) Predicting observables from a general class of distributions. J Stat Plan Infer 79:79–91
28. AL-Hussaini EK, Ahmad AEBA (2003) On Bayesian predictive distributions of generalized order statistics. Metrika 57:165–176
29. AL-Hussaini EK, Jaheen ZF (1992) Bayesian estimation of the parameters, reliability and failure rate functions of the Burr type XII failure model. J Stat Comput Simul 41:31–40
30. AL-Hussaini EK, Jaheen ZF (1994) Approximate Bayes estimators applied to the Burr model. Comm Stat Theory Meth 23:99–121
31. Alam K (1972) Unimodality of the distribution of an order statistic. Ann Math Stat 43: 2041–2044
32. Albing M (2006) Process capability analysis with focus on indices for one-sided specification limits. Master's thesis, Luleå University of Technology, Luleå, Sweden
33. Ali MM, Umbach D (1998) Optimal linear inference using selected order statistics in location-scale models. In: Balakrishnan N, Rao CR (eds) Order statistics: applications. Handbook of statistics, vol 17. Elsevier, Amsterdam, pp 183–213
34. Ali Mousa MAM (2001) Inference and prediction for Pareto progressively censored data. J Stat Comput Simul 71:163–181
35. Ali Mousa MAM, Al-Sagheer S (2005) Bayesian prediction for progressively type-II censored data from the Rayleigh model. Comm Stat Theory Meth 34:2353–2361
36. Ali Mousa MAM, Al-Sagheer S (2006) Statistical inference for the Rayleigh model based on progressively Type-II censored data. Statistics 40:149–157
37. Ali Mousa MAM, Jaheen ZF (2002) Bayesian prediction for progressively censored data from the Burr model. Stat Papers 43:587–593
38. Ali Mousa MAM, Jaheen ZF (2002) Statistical inference for the Burr model based on progressively censored data. Comput Math Appl 43:1441–1449
39. Alimohammadi M, Alamatsaz MH (2011) Some new results on unimodality of generalized order statistics and their spacings. Stat Probab Lett 81:1677–1682
40. Alvarez-Andrade S, Bordes L (2004) Empirical quantile process under type-II progressive censoring. Stat Probab Lett 68:111–123

41. Alvarez-Andrade S, Bordes L (2008) Type-II progressive censoring and related processes. Rev Roumaine Math Pures Appl 63:267–276
42. Alvarez-Andrade S, Balakrishnan N, Bordes L (2007) Homogeneity tests based on several progressively Type-II censored samples. J Multivariate Anal 98:1195–1213
43. Alvarez-Andrade S, Balakrishnan N, Bordes L (2009) Proportional hazards regression under progressive Type-II censoring. Ann Inst Stat Math 61:887–903
44. Amin ZH (2008) A note on the parameter estimation for the lognormal distribution based on progressively Type-I interval censored samples. Model Assist Stat Appl 3:169–176
45. Amin ZH (2008) Bayesian inference for the Pareto lifetime model under progressive censoring with binomial removals. J Appl Stat 35:1203–1217
46. Andersen PK, Borgan Ø, Gill RD, Keiding N (1993) Statistical models based on counting processes. Springer, New York
47. Arias-Nicolás JP, Belzunce F, Núñez Barrera O, Suárez-Llorens A (2009) A multivariate IFR notion based on the multivariate dispersive ordering. Appl Stoch Models Bus Ind 25:339–358
48. Armero C, Bayarri MJ (1994) Prior assessments for prediction in queues. The Statistician 43:139–153
49. Arnold BC (1983) Pareto distributions. Statistical distributions in scientific work, vol 5. International Co-operative Publishing House, Burtonsville
50. Arnold BC (1985) p-Norm bounds on the expectation of the maximum of possibly dependent sample. J Multivariate Anal 17:316–332
51. Arnold BC, Balakrishnan N (1989) Relations, bounds and approximations for order statistics. Lecture notes in statistics, vol 53. Springer, New York
52. Arnold BC, Nagaraja HN (1991) Lorenz ordering of exponential order statistics. Stat Probab Lett 11:485–490
53. Arnold BC, Press SJ (1983) Bayesian inference for Pareto populations. J Econometrics 21:287–306
54. Arnold BC, Press SJ (1989) Bayesian estimation and prediction for Pareto data. J Am Stat Assoc 84:1079–1084
55. Arnold BC, Villaseñor JA (1991) Lorenz ordering of order statistics. Lecture notes–monograph series, vol 19. Institute of Mathematical Statistics, Hayward, pp 38–47
56. Arnold BC, Villaseñor JA (1998) Lorenz ordering of order statistics and record values. In: Balakrishnan N, Rao CR (eds) Order Statistics: theory and methods. Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 75–87
57. Arnold BC, Becker A, Gather U, Zahedi H (1984) On the Markov property of order statistics. J Stat Plan Infer 9:147–154
58. Arnold BC, Balakrishnan N, Nagaraja HN (1992) A first course in order statistics. Wiley, New York
59. Arnold BC, Balakrishnan N, Nagaraja HN (1998) Records. Wiley, New York
60. Arnold BC, Castillo E, Sarabia JM (2009) Multivariate order statistics via multivariate concomitants. J Multivar Anal 100:946–951
61. Asgharzadeh A (2006) Point and interval estimation for a generalized logistic distribution under progressive type II censoring. Comm Stat Theory Meth 35:1685–1702
62. Asgharzadeh A (2009) Approximate MLE for the scaled generalized exponential distribution under progressive type-II censoring. J Korean Stat Soc 38:223–229
63. Asgharzadeh A, Valiollahi R (2008) Estimation based on progressively censored data from the Burr model. Int Math Forum 3:2113–2121
64. Asgharzadeh A, Valiollahi R, Raqab MZ (2011) Stress-strength reliability of Weibull distribution based on progressively censored samples. Stat Oper Res Trans 35:103–124
65. Ashour SK, Hassen A, Youssef A (1991) Bayesian estimation of a linear failure rate. IAPQR Trans 16:9–16
66. Bagai I, Kochar SC (1986) On tail-ordering and comparison of failure rates. Comm Stat Theory Meth 15:1377–1388

67. Bagdonavičius V (1978) Testing the hypothesis of additive accumulation of damages. Probab Theory Appl 23:403–408
68. Bagdonavičius V, Nikulin M (2002) Accelerated life models: modeling and statistical analysis. Chapman & Hall/CRC Press, Boca Raton/Florida
69. Bagdonavičius V, Cheminade O, Nikulin M (2004) Statistical planning and inference in accelerated life testing using the CHSS model. J Stat Plan Infer 126:535–551
70. Bai DS, Hong YW (1992) Estimation of $Pr(X < Y)$ in the exponential case with common location parameter. Comm Stat Theory Meth 21:269–282
71. Bai DS, Kim MS, Lee SH (1989) Optimum simple step-stress accelerated life tests with censoring. IEEE Trans Reliab 38:528–532
72. Bain LJ, Antle CE (1970) Inferential procedures for the Weibull and generalized gamma distributions. ARL report no. 70-0266. Wright Patterson Air Force Base, Dayton
73. Bain LJ, Engelhardt M (1973) Interval estimation for the two-parameter double exponential distribution. Technometrics 15:875–887
74. Bain LJ, Engelhardt M (1991) Statistical analysis of reliability and life-testing models, 2nd edn. Marcel Dekker, New York
75. Bairamov I (1996) Some distribution free properties of statistics based on record values and characterizations of the distributions through a record. J Appl Stat Sci 5:17–25
76. Bairamov I (2006) Progressive type II censored order statistics for multivariate observations. J Multivar Anal 97:797–809
77. Bairamov I, Eryılmaz S (2000) Distributional properties of statistics based on minimal spacing and record exceedance statistics. J Stat Plan Infer 90:21–33
78. Bairamov I, Eryılmaz S (2006) Spacings, exceedances and concomitants in progressive Type II censoring scheme. J Stat Plan Infer 136:527–536
79. Bairamov I, Gebizlioglu OL (1998) On the ordering of random vectors in a norm sense. J Appl Stat Sci 6:77–86
80. Bairamov I, Parsi S (2011) On flexible progressive censoring. J Comput Appl Math 235:4537–4544
81. Balakrishnan N (ed) (1992) Handbook of the logistic distribution. Dekker, New York
82. Balakrishnan N (1993) A simple application of binomial-negative binomial relationship in the derivation of sharp bounds for moments of order statistics based on greatest convex minorants. Stat Probab Lett 18:301–305
83. Balakrishnan N (2007) Permanents, order statistics, outliers, and robustness. Rev Math Comput 20:7–107
84. Balakrishnan N (2007) Progressive censoring methodology: an appraisal (with discussions). TEST 16:211–296
85. Balakrishnan N (2009) A synthesis of exact inferential results for exponential step-stress models and associated optimal accelerated life-tests. Metrika 69:351–396
86. Balakrishnan N, Aggarwala R (2000) Progressive censoring: theory, methods, and applications. Birkhäuser, Boston
87. Balakrishnan N, Asgharzadeh A (2005) Inference for the scaled half-logistic distribution based on progressively type-II censored samples. Comm Stat Theory Meth 34:73–87
88. Balakrishnan N, Basu AP (eds) (1995) The exponential distribution: theory, methods, and applications. Taylor & Francis, Newark
89. Balakrishnan N, Bordes L (2004) Non-parametric hazard rate estimation under progressive Type-II censoring. In: Balakrishnan N, Rao CR (eds) Advances in survival analysis. Handbook of statistics, vol 23. Elsevier, Amsterdam, pp 227–249
90. Balakrishnan N, Chan PS (1995) Maximum likelihood estimation for the log-gamma distribution under Type-II censored samples and associated inference. In: Balakrishnan N (ed) Recent advances in life-testing and reliability: a volume in honor of Alonzo Clifford Cohen Jr. CRC Press, Boca Raton, pp 409–438
91. Balakrishnan N, Chan PS (1995) Maximum likelihood estimation for the three-parameter log-gamma distribution under Type-II censoring. In: Balakrishnan N (ed) Recent advances in

life-testing and reliability: a volume in honor of Alonzo Clifford Cohen Jr. CRC Press, Boca Raton, pp 439–454
92. Balakrishnan N, Cohen AC (1991) Order statistics and inference: estimation methods. Academic, Boston
93. Balakrishnan N, Cramer E (2008) Progressive censoring from heterogeneous distributions with applications to robustness. Ann Inst Stat Math 60:151–171
94. Balakrishnan N, Cutler CD (1996) Maximum likelihood estimation of Laplace parameters based on Type-II censored samples. In: Morrison DF, Nagaraja HN, Sen PK (eds) Statistical theory and applications. papers in honor of Herbert A. David. Springer, Berlin, pp 145–151
95. Balakrishnan N, Dembińska A (2007) Ordered random variables from discontinuous distributions. J Iran Stat Soc (JIRSS) 6:31–46
96. Balakrishnan N, Dembińska A (2008) Progressively Type-II right censored order statistics from discrete distributions. J Stat Plan Infer 138:845–856
97. Balakrishnan N, Dembińska A (2009) Erratum to 'progressively Type-II right censored order statistics from discrete distributions' [J Stat Plan Infer 138:845–856 (2008)]. J Stat Plan Infer 139:1572–1574
98. Balakrishnan N, Han D (2007) Optimal progressive Type-II censoring schemes for nonparametric confidence intervals of quantiles. Comm Stat Simul Comput 36:1247–1262
99. Balakrishnan N, Han D (2009) Optimal step-stress testing for progressively Type-I censored data from exponential distribution. J Stat Plan Infer 139:1782–1798
100. Balakrishnan N, Hossain A (2007) Inference for the Type-II generalized logistic distribution under progressive Type-II censoring. J Stat Comput Simul 77:1013–1031
101. Balakrishnan N, Iliopoulos G (2009) Stochastic monotonicity of the MLE of exponential mean under different censoring schemes. Ann Inst Stat Math 61:753–772
102. Balakrishnan N, Iliopoulos G (2010) Stochastic monotonicity of the MLEs of parameters in exponential simple step-stress models under Type-I and Type-II censoring. Metrika 72: 89–109
103. Balakrishnan N, Joshi PC (1984) Product moments of order statistics from the doubly truncated exponential distribution. Naval Res Logist Quart 31:27–31
104. Balakrishnan N, Kannan N (2001) Point and interval estimation for parameters of the logistic distribution based on progressively Type-II censored samples. In: Balakrishnan N, Rao CR (eds) Advances in reliability. Handbook of statistics, vol 20. Elsevier, Amsterdam, pp 431–456
105. Balakrishnan N, Kateri M (2008) On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. Stat Probab Lett 78:2971–2975
106. Balakrishnan N, Kim JA (2005) EM algorithm and optimal censoring schemes for progressively Type-II censored bivariate normal data. In: Balakrishnan N, Nagaraja HN, Kannan N (eds) Advances in ranking and selection, multiple comparisons, and reliability. Birkhäuser, Boston, pp 21–45
107. Balakrishnan N, Kim JA (2005) Point and interval estimation for bivariate normal distribution based on progressively Type-II censored data. Comm Stat Theory Meth 34:1297–1347
108. Balakrishnan N, Kundu D (2013) Hybrid censoring: models, inferential results and applications (with discussions). Comput Stat Data Anal 57:166–209
109. Balakrishnan N, Leung MY (1988) Order statistics from the type I generalized logistic distribution. Comm Stat Simul Comput 17:25–50
110. Balakrishnan N, Lin CT (2002) Exact linear inference and prediction for exponential distributions based on general progressively type-II censored samples. J Stat Comput Simul 72:677–686
111. Balakrishnan N, Lin CT (2003) On the distribution of a test for exponentiality based on progressively type-II right censored spacings. J Stat Comput Simul 73:277–283
112. Balakrishnan N, Malov SV (2005) Some characterizations of exponential distribution based on progressively censored order statistics. In: Balakrishnan N, Bairamov IG, Gebizlioglu OL (eds) Advances on models, characterizations and applications. Chapman & Hall/CRC, Boca Raton, pp 97–109

113. Balakrishnan N, Mi J (2003) Existence and uniqueness of the MLEs for normal distribution based on general progressively type-II censored samples. Stat Probab Lett 64:407–414
114. Balakrishnan N, Ng HKT (2006) Precedence-type tests and applications. Wiley, Hoboken
115. Balakrishnan N, Rao CR (1997) Large-sample approximations to the best linear unbiased estimation and best linear unbiased prediction based on progressively censored samples and some applications. In: Panchapakesan S, Balakrishnan N (eds) Advances in statistical decision theory and applications. Birkhäuser, Boston, pp 431–444
116. Balakrishnan N, Rao CR (eds) (1998) Order statistics: applications. Handbook of statistics, vol 17. Elsevier, Amsterdam
117. Balakrishnan N, Rao CR (eds) (1998) Order statistics: theory and methods. Handbook of statistics, vol 16. Elsevier, Amsterdam
118. Balakrishnan N, Rasouli A (2008) Exact likelihood inference for two exponential populations under joint Type-II censoring. Comput Stat Data Anal 52:2725–2738
119. Balakrishnan N, Saleh HM (2011) Relations for moments of progressively Type-II censored order statistics from half-logistic distribution with applications to inference. Comput Stat Data Anal 55:2775–2792
120. Balakrishnan N, Saleh HM (2012) Relations for moments of progressively Type-II censored order statistics from log-logistic distribution with applications to inference. Comm Stat Theory Meth 41:880–906
121. Balakrishnan N, Saleh HM (2013) Recurrence relations for single and product moments of progressively Type-II censored order statistics from a generalized half-logistic distribution with application to inference. J Stat Comput Simul 83:1704–1721
122. Balakrishnan N, Sandhu RA (1995) A simple simulational algorithm for generating progressive Type-II censored samples. Am Stat 49:229–230
123. Balakrishnan N, Sandhu RA (1996) Best linear unbiased and maximum likelihood estimation for exponential distributions under general progressive Type-II censored samples. Sankhyā Ser B 58:1–9
124. Balakrishnan N, Stepanov A (2005) A note on the number of observations near an order statistic. J Stat Plan Infer 134:1–14
125. Balakrishnan N, Stepanov A (2008) Asymptotic properties of numbers of near minimum observations under progressive Type-II censoring. J Stat Plan Infer 138:1010–1020
126. Balakrishnan N, Sultan KS (1998) Recurrence relations and identities for moments of order statistics. In: Balakrishnan N, Rao CR (eds) Order statistics: theory and Methods, Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 149–228
127. Balakrishnan N, Varadan J (1991) Approximate MLEs for the location and scale parameters of the extreme value distribution with censoring. IEEE Trans Reliab 40:146–151
128. Balakrishnan N, Chandramouleeswaran M, Ambagaspitiya R (1996) BLUEs of location and scale parameters of Laplace distribution based on type-II censored samples and associated inference. Microelectron Reliab 36:371–374
129. Balakrishnan N, Cramer E, Kamps U (2001) Bounds for means and variances of progressive type II censored order statistics. Stat Probab Lett 54:301–315
130. Balakrishnan N, Cramer E, Kamps U, Schenk N (2001) Progressive type II censored order statistics from exponential distributions. Statistics 35:537–556
131. Balakrishnan N, Ng HKT, Kannan N (2002) A test of exponentiality based on spacings for progressively type-II censored data. In: Huber-Carol C, Balakrishnan N, Nikulin M, Mesbah M (eds) Goodness-of-fit tests and model validity. Birkhäuser, Boston, pp 89–111
132. Balakrishnan N, Childs A, Chandrasekar B (2002) An efficient computational method for moments of order statistics under progressive censoring. Stat Probab Lett 60:359–365
133. Balakrishnan N, Brain C, Mi J (2002) Stochastic order and MLE of the mean of the exponential distribution. Methodol Comput Appl Probab 4:83–93
134. Balakrishnan N, Kannan N, Lin CT, Ng HKT (2003) Point and interval estimation for Gaussian distribution, based on progressively Type-II censored samples. IEEE Trans Reliab 52:90–95

135. Balakrishnan N, Ng HKT, Kannan N (2004) Goodness-of-fit tests based on spacings for progressively type-II censored data from a general location-scale distribution. IEEE Trans Reliab 53:349–356
136. Balakrishnan N, Kannan N, Lin CT, Wu SJS (2004) Inference for the extreme value distribution under progressive Type-II censoring. J Stat Comput Simul 74:25–45
137. Balakrishnan N, Cramer E, Kamps U (2005) Relation for joint densities of progressively censored order statistics. Statistics 39:529–536
138. Balakrishnan N, Habibi Rad A, Arghami NR (2007) Testing exponentiality based on Kullback-Leibler information with progressively Type-II censored data. IEEE Trans Reliab 56:301–307
139. Balakrishnan N, Burkschat M, Cramer E (2008) Best linear equivariant estimation and prediction in location-scale families. Sankhyā Ser B 70:229–247
140. Balakrishnan N, Burkschat M, Cramer E, Hofmann G (2008) Fisher information based progressive censoring plans. Comput Stat Data Anal 53:366–380
141. Balakrishnan N, Tripathi RC, Kannan N (2008) On the joint distribution of placement statistics under progressive censoring and applications to precedence test. J Stat Plan Infer 138:1314–1324
142. Balakrishnan N, Davies KF, Keating JP (2009) Pitman closeness of order statistics to population quantiles. Comm Stat Simul Comput 38:802–820
143. Balakrishnan N, Iliopoulos G, Keating JP, Mason RL (2009) Pitman closeness of sample median to population median. Stat Probab Lett 79:1759–1766
144. Balakrishnan N, Beutner E, Cramer E (2010) Exact two-sample non-parametric confidence, prediction, and tolerance intervals based on ordinary and progressively Type-II right censored data. TEST 19:68–91
145. Balakrishnan N, Bordes L, Zhao X (2010) Minimum-distance parametric estimation under progressive Type-I censoring. IEEE Trans Reliab 59:413–425
146. Balakrishnan N, Davies KF, Keating JP, Mason RL (2010) Simultaneous closeness among order statistics to population quantiles. J Stat Plan Infer 140:2408–2415
147. Balakrishnan N, Tripathi RC, Kannan N, Ng, HKT (2010) Some nonparametric precedence-type tests based on progressively censored samples and evaluation of power. J Stat Plan Infer 140:559–573
148. Balakrishnan N, Belzunce F, Hami N, Khaledi BE (2010) Univariate and multivariate likelihood ratio ordering of generalized order statistics and associated conditional variables. Probab Eng Inform Sci 24:441–455
149. Balakrishnan N, Cramer E, Dembińska A (2011) Characterizations of geometric distribution through progressively Type-II right censored order statistics. Statistics 59:559–573
150. Balakrishnan N, Han D, Iliopoulos G. (2011) Exact inference for progressively Type-I censored exponential failure data. Metrika 73:335–358
151. Balakrishnan N, AL-Hussaini EK, Saleh HM (2011) Recurrence relations for moments of progressively censored order statistics from logistic distribution with applications to inference. J Stat Plan Infer 141:17–30
152. Balakrishnan N, Kamps U, Kateri M (2012) A sequential order statistics approach to step-stress testing. Ann Inst Stat Math 64:303–318
153. Balakrishnan N, Belzunce F, Sordo MA, Suárez-Llorens A (2012) Increasing directionally convex orderings of random vectors having the same copula, and their use in comparing ordered data. J Multivariate Anal 105:45–54
154. Balakrishnan N, Volterman W, Zhang L (2013) A meta-analysis of multisample Type-II censored data with parametric and nonparametric results. IEEE Trans Reliab 62:2–12
155. Balakrishnan N, Beutner E, Cramer E (2013) Computational aspects of statistical intervals based on two Type-II censored samples. Comput Stat 28:893–917
156. Balakrishnan N, Cramer E, Iliopoulos G (2014) On the method of pivoting the CDF for exact confidence intervals with illustration for exponential mean under life-test with time constraints. Stat Probab Lett 89:124–130

157. Balasooriya U (1995) Failure-censored reliability sampling plans for the exponential distribution. J Stat Comput Simul 52:337–349
158. Balasooriya U, Balakrishnan N (2000) Reliability sampling plans for the lognormal distribution, based on progressively censored samples. IEEE Trans Reliab 49:199–203
159. Balasooriya U, Low CK (2004) Competing causes of failure and reliability tests for Weibull lifetimes under type I progressive censoring. IEEE Trans Reliab 53:29–36
160. Balasooriya U, Low CK (2007) Comments on: progressive censoring methodology: an appraisal. TEST 16:284–286
161. Balasooriya U, Saw SLC (1998) Reliability sampling plans for the two-parameter exponential distribution under progressive censoring. J Appl Stat 25:707–714
162. Balasooriya U, Saw SLC (1999) A note on approximate moments of progressively censored order statistics. Metron 57:117–130
163. Balasooriya U, Saw SLC, Gadag V (2000) Progressively censored reliability sampling plans for the Weibull distribution. Technometrics 42:160–167
164. Balasubramanian K, Beg MI (1990) On expectations of functions of order statistics. Sankhyā 52:103–114
165. Ball F (1985) Deterministic and stochastic epidemics with several kinds of susceptibles. Adv Appl Probab 17:1–22
166. Barbour A, Lindvall T, Rogers L (1991) Stochastic ordering of order statistics. J Appl Probab 28:278–286
167. Barlow RE, Proschan F (1966) Inequalities for linear combinations of order statistics from restricted families. Ann Math Stat 37:1574–1592
168. Barlow RE, Proschan F (1975) Statistical theory of reliability and life testing, probability models. Holt-Rinehart and Winston, New York
169. Barlow RE, Madansky A, Proschan F, Scheuer EM (1968) Statistical estimation procedures for the 'burn-in' process. Technometrics 10:51–62
170. Barlow RE, Bartholomew DJ, Bremner JM, Brunk HD (1972) Statistical inference under order restrictions. Wiley, New York
171. Barnett V, Lewis T (1994) Outliers in statistical data, 3rd edn. Wiley, Chichester
172. Bartoszewicz J (1977) Estimation of $P(Y < X)$ in the exponential case. Zastosowania Matematyki 16:1–8
173. Bartoszewicz J (1985) Moment inequalities for order statistics from ordered families of distributions. Metrika 32:383–389
174. Basak I, Balakrishnan N (2003) Robust estimation under progressive censoring. Comput Stat Data Anal 44:349–376
175. Basak I, Balakrishnan N (2009) Predictors of failure times of censored units in progressively censored samples from normal distribution. Sankhyā 71-B:222–247
176. Basak I, Balakrishnan N (2012) Estimation for the three-parameter gamma distribution based on progressively censored data. Stat Meth 9:305–319
177. Basak P, Balakrishnan N (2012) Estimation for the three-parameter inverse Gaussian distribution under progressive Type-II censoring. J Stat Comput Simul 82:1055–1072
178. Basak I, Basak P, Balakrishnan N (2006) On some predictors of times to failure of censored items in progressively censored samples. Comput Stat Data Anal 50:1313–1337
179. Basak P, Basak I, Balakrishnan N (2009) Estimation for the three-parameter lognormal distribution based on progressively censored data. Comput Stat Data Anal 53:3580–3592
180. Basirat M, Baratpour S, Ahmadi J (2013) Statistical inferences for stress-strength in the proportional hazard models based on progressive Type-II censored samples. J Stat Comput Simul (to appear)
181. Basu D (1955) On statistics independent of a complete sufficient statistic. Sankhyā 15: 377–380
182. Basu AP (1995) Accelerated life testing with applications. In: Balakrishnan N, Basu AP (eds) The exponential distribution: theory, methods and applications. Gordon and Breach, Newark, pp 377–383

183. Bebiano N (1982) On the evaluation of permanents. Pacific J Math 101:1–9
184. Beg MA (1980) Estimation of $Pr(Y < X)$ for exponential-family. IEEE Trans Reliab R 29:158–159
185. Beg MI, Ahsanullah M (2008) Concomitants of generalized order statistics from Farlie-Gumbel-Morgenstern distributions. Stat Meth 5:1–20
186. Belzunce F, Lillo RE, Ruiz JM, Shaked M (2001) Stochastic comparisons of nonhomogeneous processes. Probab Eng Inform Sci 15:199–224
187. Belzunce F, Mercader JA, Ruiz JM (2003) Multivariate aging properties of epoch times of nonhomogeneous processes. J Multivariate Anal 84:335–350
188. Belzunce F, Mercader JA, Ruiz JM (2005) Stochastic comparisons of generalized order statistics. Probab Eng Inform Sci 19:99–120
189. Belzunce F, Ruiz JM, Suárez-Llorens A (2008) On multivariate dispersion orderings based on the standard construction. Stat Probab Lett 78:271–281
190. Berger JO (1985) Statistical decision theory and Bayesian analysis, 2nd edn. Springer, New York
191. Berger JO, Sun D (1993) Bayesian analysis for the Poly-Weibull distribution. J Am Stat Assoc 88:1412–1418
192. Beutner E (2007) Progressive type-II censoring and transition kernels. Comm Dependability Qual Manag 10:25–32
193. Beutner E, Cramer E (2010) Nonparametric meta-analysis for minimal-repair systems. Aust N Z J Stat 52:383–401
194. Beutner E, Cramer E (2011) Confidence intervals for quantiles in a minimal repair set-up. Int J Appl Math Stat 24:86–97
195. Beutner E, Kamps U (2009) Order restricted statistical inference for scale parameters based on sequential order statistics. J Stat Plan Infer 139:2963–2969
196. Bhattacharya PK (1974) Convergence of sample paths of normalized sums of induced order statistics. Ann Stat 2:1034–1039
197. Bhattacharya PK (1984) Induced order statistics: theory and applications. In: Krishnaiah PR, Sen PK (eds) Nonparametric methods. Handbook of statistics, vol 4. Elsevier, Amsterdam, pp 383–403
198. Bhattacharya B (2007) Testing for ordered failure rates under general progressive censoring. J Stat Plan Infer 137:1775–1786
199. Bhattacharyya GK, Johnson RA (1974) Estimation of reliability in a multicomponent stress-strength model. J Am Stat Assoc 69:966–970
200. Bhattacharyya GK, Soejoeti Z (1989) A tampered failure rate model for step-stress accelerated life test. Comm Stat Theory Meth 18:1627–1643
201. Bickel PJ (1967) Some contributions to the theory of order statistics. In: Proceedings of 5th Berkeley symposium on mathematical statistics and probability (Berkeley, CA, 1965/1966), vol I: statistics. University of California Press, Berkeley, pp 575–591
202. Bieniek M (2007) On characterizations of distributions by regression of adjacent generalized order statistics. Metrika 66:233–242
203. Bieniek, M. (2007) Variation diminishing property of densities of uniform generalized order statistics. Metrika 65:297–309
204. Bieniek M, Szynal D (2003) Characterizations of distributions via linearity of regression of generalized order statistics. Metrika 58:259–271
205. Birnbaum ZW (1956) On a use of Mann-Whitney statistics. In: Proceedings of the 3rd Berkeley symposium on mathematical statistics and probability (Berkeley, CA, 1954/1955), vol I: contributions to the theory of statistics. University of California Press, Berkeley, pp 13–17
206. Block HW (1975) Continuous multivariate exponential extensions. In: Barlow RE, Fussel JB, Singpurwalla ND (eds) Reliability and fault tree analysis. SIAM, Philadelphia, pp 285–306
207. Block HW, Joe H (1997) Tail behaviour of the failure rate functions of mixtures. Lifetime Data Anal 3:269–288

208. Blom G (1958) Statistical estimates and transformed beta-variables. Almqvist & Wiksell, Stockholm
209. Blyth CR (1972) Some probability paradoxes in choice from among random alternatives: rejoinder. J Am Stat Assoc 67:379–381
210. Bobotas P, Kourouklis S (2011) Improved estimation of the scale parameter, the hazard rate parameter and the ratio of the scale parameters in exponential distributions: an integrated approach. J Stat Plan Infer 141:2399–2416
211. Boland PJ, Hollander M, Joag-Dev K, Kochar SC (1996) Bivariate dependence properties of order statistics. J Multivariate Anal 56:75–89
212. Boland PJ, Shaked M, Shanthikumar JG (1998) Stochastic ordering of order statistics. In: Balakrishnan N, Rao CR (eds) Order statistics: theory and methods. Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 89–103
213. Bon JL, Păltănea E (1999) Ordering properties of convolutions of exponential random variables. Lifetime Data Anal 5:185–192
214. Bondesson L (1979) On best equivariant and best unbiased equivariant estimators in the extended Gauss-Markov model. Scand J Stat 6:22–28
215. Boos DD, Hughes-Oliver JM (1998) Applications of Basu's theorem. Am Stat 52:218–221
216. Bordes L (2004) Non-parametric estimation under progressive censoring. J Stat Plan Infer 119:171–189
217. Botta RF, Harris CM, Marchal WG (1987) Characterizations of generalized hyperexponential distribution functions. Comm Stat Stoch Models 3:115–148
218. Bradley DM, Gupta RC (2002) On the distribution of the sum of n non-identically distributed uniform random variables. Ann Inst Stat Math 54:689–700
219. Brascamp HJ, Lieb EH (1975) Some inequalities for Gaussian measures and the long-range order of the one-dimensional plasma. In: Arthurs AM (ed) Functional integration and its applications, proceedings of the conference on functional integration, Cumberland Lodge, England. Clarendon Press, Oxford, pp 1–14
220. Brockwell PJ, Davis RA (2009) Time series: theory and methods, 2nd edn. Springer, New York
221. Brown L (1968) Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. Ann Math Stat 39:29–48
222. Brunk HD (1962) On the range of the difference between hypothetical distribution function and Pyke's modified empirical distribution function. Ann Math Stat 33:525–532
223. Bryson MC (1974) Heavy-tailed distributions: properties and tests. Technometrics 16:61–68
224. Buckle N, Kraft C, Eeden Cv (1969) An approximation to the Wilcoxon-Mann-Whitney distribution. J Am Stat Assoc 64:591–599
225. Burke M (2011) Nonparametric estimation of a survival function under progressive Type-I multistage censoring. J Stat Plan Infer 141:910–923
226. Burkschat M (2006) Estimation with generalized order statistics. Ph.D. thesis, RWTH Aachen University, Aachen, Germany
227. Burkschat M (2008) On optimal linear equivariant estimation under progressive censoring. Statistics 42:383–392
228. Burkschat M (2008) On optimality of extremal schemes in progressive Type-II censoring. J Stat Plan Infer 138:1647–1659
229. Burkschat M (2009) Multivariate dependence of spacings of generalized order statistics. J Multivariate Anal 100:1093–1106
230. Burkschat M (2009) Systems with failure-dependent lifetimes of components. J Appl Probab 46:1052–1072
231. Burkschat M (2010) Linear estimators and predictors based on generalized order statistics from generalized Pareto distributions. Comm Stat Theory Meth 39:311–326
232. Burkschat M, Cramer E (2012) Fisher information in generalized order statistics. Statistics 46:719–743
233. Burkschat M, Navarro J (2011) Aging properties of sequential order statistics. Probab Eng Inform Sci 25:449–467

234. Burkschat M, Cramer E, Kamps U (2003) Dual generalized order statistics. Metron 61: 13–26
235. Burkschat M, Cramer E, Kamps U (2006) On optimal schemes in progressive censoring. Stat Probab Lett 76:1032–1036
236. Burkschat M, Cramer E, Kamps U (2007) Linear estimation of location and scale parameters based on generalized order statistics from generalized Pareto distributions. In: Ahsanullah M, Raqab M (eds) Recent developments in ordered random variables. Nova Science Publisher, New York, pp 253–261
237. Burkschat M, Cramer E, Kamps U (2007) Optimality criteria and optimal schemes in progressive censoring. Comm Stat Theory Meth 36:1419–1431
238. Calabria R, Pulcini G (1996) Point estimation under asymmetric loss functions for left-truncated exponential samples. Comm Stat Theory Meth 25:585–600
239. Casella G, Berger RL (2002) Statistical inference, 2nd edn. Duxbury Press, Pacific Grove
240. Chan LK, Chan NN (1973) On the optimum best linear unbiased estimates of the parameters of the normal distribution based on selected order statistics. Scand Actuar J 1973:120–128
241. Chan LK, Mead E (1971) Linear estimation of the parameters of the extreme-value distribution based on suitably chosen order statistics. IEEE Trans Reliab R-20:74–83
242. Chan W, Proschan F, Sethuraman J (1991) Convex-ordering among functions, with applications to reliability and mathematical statistics. In: Block HW, Sampson AR, Savits TH (eds) Topics in statistical dependence. IMS lecture notes: monograph series. IMS, Hayward, pp 121–134
243. Chandrasekar B, Balakrishnan N (2002) On a multiparameter version of Tukey's linear sensitivity measure and its properties. Ann Inst Stat Math 54:796–805
244. Chandrasekar B, Leo Alexander T, Balakrishnan N (2002) Equivariant estimation for parameters of exponential distributions based on Type-II progressively censored samples. Comm Stat Theory Meth 31:1675–1686
245. Chandrasekar B, Childs A, Balakrishnan N (2004) Exact likelihood inference for the exponential distribution under generalized Type-I and Type-II hybrid censoring. Naval Res Logist 51:994–1004
246. Chatterjee SK, Sen PK (1973) Nonparametric testing under progressive censoring. Calcutta Stat Assoc Bull 22:13–50
247. Chen Z (1996) Joint confidence region for the parameters of Pareto distribution. Metrika 44:191–197
248. Chen G, Balakrishnan N (1995) A general purpose approximate goodness-of-fit test. J Qual Tech 27:154–161
249. Chen SM, Bhattacharyya GK (1988) Exact confidence bounds for an exponential parameter under hybrid censoring. Comm Stat Theory Meth 16:2429–2442
250. Chen J, Hu T (2007) Multivariate dispersive ordering of generalized order statistics. J Iran Stat Soc (JIRSS) 6:61–75
251. Chen DG, Lio YL (2009) A note on the maximum likelihood estimation for the generalized gamma distribution parameters under progressive Type-II censoring. Int J Intell Tech Appl Stat 2:57–64
252. Chen DG, Lio YL (2010) Parameter estimations for generalized exponential distribution under progressive Type-I interval censoring. Comput Stat Data Anal 54:1581–1591
253. Chen JJ, Novick MR (1984) Bayesian analysis for binomial models with generalized beta prior distributions. J Educ Stat 9:163–175
254. Chen H, Xie H, Hu T (2009) Log-concavity of generalized order statistics. Stat Probab Lett 79:396–399
255. Cheng C, Chen J, Li Z (2010) A new algorithm for maximum likelihood estimation with progressive Type-I interval censored data. Comm Stat Simul Comput 39:750–766
256. Childs A, Balakrishnan N (1996) Conditional inference procedures for the Laplace distribution based on type-II right censored samples. Stat Probab Lett 31:31–39

257. Childs A, Balakrishnan N (1997) Maximum likelihood estimation of Laplace parameters based on general Type-II censored examples. Stat Papers 38:343–349
258. Childs A, Balakrishnan N (2000) Conditional inference procedures for the Laplace distribution when the observed samples are progressively censored. Metrika 52:253–265
259. Childs A, Chandrasekar B, Balakrishnan N, Kundu D (2003) Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution. Ann Inst Stat Math 55:319–330
260. Childs A, Chandrasekar B, Balakrishnan N (2008) Exact likelihood inference for an exponential parameter under progressive hybrid censoring schemes. In: Vonta F, Nikulin M, Limnios N, Huber-Carol C (eds) Statistical models and methods for biomedical and technical systems. Birkhäuser, Boston, pp 323–334
261. Childs A, Balakrishnan N, Chandrasekar B (2012) Exact distribution of the MLEs of the parameters and of the quantiles of two-parameter exponential distribution under hybrid censoring. Statistics 46:441–458
262. Chiou WJ, Cohen A (1984) Estimating the common location parameter of exponential distributions with censored samples. Naval Res Logist 31:475–482
263. Cho Y, Cho E (2001) The volume of simplices clipped by a half space. Appl Math Lett 14:731–735
264. Christensen R (2011) Plane answers to complex questions: the theory of linear models, 4th edn. Springer, New York
265. Čibisov DM (1964) Limit distributions for order statistics. Theory Probab Appl 9:142–148
266. Cohen AC (1951) Estimating parameters of logarithmic-normal distributions by maximum likelihood. J Am Stat Assoc 46:206–212
267. Cohen AC (1963) Progressively censored samples in life testing. Technometrics 5:327–329
268. Cohen AC (1965) Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. Technometrics 7:579–588
269. Cohen AC (1966) Life testing and early failure. Technometrics 17:347–351
270. Cohen AC (1975) Multi-censored sampling in the three parameter Weibull distribution. Technometrics 17:347–351
271. Cohen AC (1976) Progressively censored sampling in the three parameter log-normal distribution. Technometrics 18:99–103
272. Cohen AC (1991) Truncated and censored samples. theory and applications. Marcel Dekker, New York
273. Cohen AC (1995) MLEs under censoring and truncation and inference. In: Balakrishnan N, Basu AP (eds) The exponential distribution: theory, methods, and applications. Gordon and Breach, Newark, pp 33–51
274. Cohen AC, Norgaard NJ (1977) Progressively censored sampling in the three-parameter gamma distribution. Technometrics 19:333–340
275. Cohen A, Sackrowitz HB (1992) Some remarks on a notion of positive dependence, association, and unbiased testing. In: Shaked M, Tong YL (eds) Stochastic inequalities: papers from the AMS-IMS-SIAM joint summer research conference held in Seattle, Washington, July 1991. IMS lecture notes—monograph series. Institute of Mathematical Statistics, Hayward, pp 33–37
276. Cohen AC, Whitten BJ (1982) Modified maximum likelihood and modified moment estimators for the three-parameter Weibull distribution. Comm Stat Theory Meth 11: 2631–2656
277. Cohen AC, Whitten BJ (1988) Parameter estimation in reliability and life span models. Marcel Dekker, New York
278. Cole RH (1951) Relations between moments of order statistics. Ann Math Stat 22:308–310
279. Constantine K, Tse SK, Karson M (1986) Estimation of $P(X < Y)$ in the gamma case. Comm Stat Simul Comput 15:365–388
280. Corana A, Marchesi M, Martini C, Ridella S (1987) Minimizing multimodal functions of continuous variables with the “Simulated Annealing” algorithm. ACM Trans Math Software 13:262–280. See also Corana et al. [281]

281. Corana A, Marchesi M, Martini C, Ridella S (1989) Corrigenda: "minimizing multimodal functions of continuous variables with the 'Simulated Annealing' algorithm". ACM Trans Math Software 15:287. See Corana et al. [280]
282. Cover TM, Thomas JA (2006) Elements of information theory, 2nd edn. Wiley, Hoboken
283. Cramer E (2001) Inference for stress-strength systems based on Weinman multivariate exponential samples. Comm Stat Theory Meth 30:331–346
284. Cramer E (2002) A note on moments of progressively Type II censored order statistics. Comm Stat Theory Meth 31:1301–1307
285. Cramer E (2003) Contributions to generalized order statistics. Habilitationsschrift. University of Oldenburg, Oldenburg
286. Cramer E (2004) Logconcavity and unimodality of progressively censored order statistics. Stat Probab Lett 68:83–90
287. Cramer E (2006) Dependence structure of generalized order statistics. Statistics 40: 409–413
288. Cramer E (2006) Sequential order statistics. In: Kotz S, Balakrishnan N, Read CB, Vidakovic B, Johnson NL (eds) Encyclopedia of statistical sciences, vol 12. Wiley, Hoboken, pp 7629–7634
289. Cramer E (2009) Hermite interpolation polynomials and distributions of ordered data. Stat Meth 6:337–343
290. Cramer E (2014) Extreme value analysis for progressively Type-II censored order statistics. Comm Stat Theory Meth 43:2135–2155
291. Cramer E, Bagh C (2011) Minimum and maximum information censoring plans in progressive censoring. Comm Stat Theory Meth 40:2511–2527
292. Cramer E, Balakrishnan N (2013) On some exact distributional results based on Type-I progressively hybrid censored data from exponential distributions. Stat Meth 10:128–150
293. Cramer E, Ensenbach M (2011) Asymptotically optimal progressive censoring plans based on Fisher information. J Stat Plan Infer 141:1968–1980
294. Cramer E, Iliopoulos G (2010) Adaptive progressive Type-II censoring. TEST 19:342–358
295. Cramer E, Kamps U (1996) Sequential order statistics and k-out-of-n systems with sequentially adjusted failure rates. Ann Inst Stat Math 48:535–549
296. Cramer E, Kamps U (1997) A note on the UMVUE of $Pr(X < Y)$ in the exponential case. Comm Stat Theory Meth 26:1051–1055
297. Cramer E, Kamps U (1997) The UMVUE of $P(X < Y)$ based on Type-II censored samples from Weinman multivariate exponential distributions. Metrika 46:93–121
298. Cramer E, Kamps U (1998) Sequential k-out-of-n systems with Weibull components. Econ Qual Control 13:227–239
299. Cramer E, Kamps U (2001) Estimation with sequential order statistics from exponential distributions. Ann Inst Stat Math 53:307–324
300. Cramer E, Kamps U (2001) Sequential k-out-of-n systems, Chap. 12. In: Balakrishnan N, Rao CR (eds) Handbook of statistics: advances in reliability, vol 20. Elsevier, Amsterdam, pp 301–372
301. Cramer E, Kamps U (2003) Marginal distributions of sequential and generalized order statistics. Metrika 58:293–310
302. Cramer E, Kamps U (2005) Characterization of the exponential distribution by conditional expectations of generalized spacings. In: Balakrishnan N, Bairamov IG, Gebizlioglu OL (eds) Advances on models, characterizations, and applications. Chapman & Hall/CRC, Boca Raton, pp 83–96
303. Cramer E, Lenz U (2010) Association of progressively Type-II censored order statistics. J Stat Plan Infer 140:576–583
304. Cramer E, Naehrig G (2012) Laplace record data. J Stat Plan Infer 142:2179–2189
305. Cramer E, Schmiedt AB (2011) Progressively Type-II censored competing risks data from Lomax distributions. Comput Stat Data Anal 55:1285–1303

306. Cramer E, Tamm M (2013) On a correction of the scale MLE for a two-parameter exponential distribution under progressive Type-I censoring. Comm Stat Theory Meth (to appear)
307. Cramer E, Tran TTH (2009) Generalized order statistics from arbitrary distributions and the Markov chain property. J Stat Plan Infer 139:4064–4071
308. Cramer E, Kamps U, Schenk N (2001) On the joint completeness of two independent statistics. Stat Decisions 20:269–277
309. Cramer E, Kamps U, Rychlik T (2002) Evaluations of expected generalized order statistics in various scale units. Appl Math 29:285–295
310. Cramer E, Kamps U, Rychlik T (2002) On the existence of moments of generalized order statistics. Stat Probab Lett 59:397–404
311. Cramer E, Kamps U, Raqab MZ (2003) Characterizations of exponential distributions by spacings of generalized order statistics. Appl Math 30:257–265
312. Cramer E, Kamps U, Keseling C (2004) Characterizations via linear regression of ordered random variables: a unifying approach. Comm Stat Theory Meth 33:2885–2911
313. Cramer E, Kamps U, Rychlik T (2004) Unimodality of uniform generalized order statistics, with applications to mean bounds. Ann Inst Stat Math 56:183–192
314. Cramer E, Herle K, Balakrishnan N (2009) Permanent expansions and distributions of order statistics in the INID case. Comm Stat Theory Meth 38:2078–2088
315. Cramer E, Burkschat M, Górny J (2014) On some exact distributional results based on Type-II progressively hybrid censored data from exponential distributions (submitted)
316. Crowder MJ (2001) Classical competing risks. Chapman & Hall/CRC, Boca Raton
317. Crowley J, Hu M (1977) Covariance analysis of heart transplant survival data. J Am Stat Assoc 72:27–36
318. D'Agostino RB, Stephens MA (eds) (1986) Goodness-of-fit techniques. Marcel Dekker, New York
319. Dahmen W, Micchelli CA (1986) Statistical encounters with B-splines. In: Marron JS (ed) Function estimates (Arcata, California, 1985). American Mathematical Society, Providence, pp 17–48
320. Dahmen K, Burkschat M, Cramer E (2012) A- and D-optimal progressive Type-II censoring designs based on Fisher information. J Stat Comput Simul 82:879–905
321. Dallas AC (1976) Characterizing the Pareto and power distributions. Ann Inst Stat Math 28:491–497
322. David HA (1973) Concomitants of order statistics. Bull Inter Stat Inst 45:295–300
323. David HA, Groeneveld RA (1982) Measures of local variation in a distribution: expected length of spacings and variances of order statistics. Biometrika 69:227–232
324. David FN, Johnson NL (1954) Statistical treatment of censored data: I. Fundamental formulae. Biometrika 41:228–240
325. David HA, Joshi PC (1968) Recurrence relations between moments of order statistics for exchangeable variates. Ann Math Stat 39:272–274
326. David HA, Nagaraja HN (1998) Concomitants of order statistics. In: Balakrishnan N, Rao CR (eds) Handbook of statistics: order statistics: theory and methods, vol 16. Elsevier, Amsterdam, pp 487–513
327. David HA, Nagaraja HN (2003) Order statistics, 3rd edn. Wiley, Hoboken
328. David HA, Shu VS (1978) Robustness of location estimators in the presence of an outlier. In: David HA (ed) Contributions to survey sampling and applied statistics. Academic, New York, pp 235–250
329. Davis DJ (1952) An analysis of some failure data. J Am Stat Assoc 47:113–150
330. de Boor C (2001) A practical guide to splines, revised edn. Springer, New York
331. De Haan L, Ferreira A (2006) Extreme value theory. an introduction. Springer, New York
332. DeGroot MH, Goel PK (1979) Bayesian estimation and optimal designs in partially accelerated life testing. Naval Res Logist Quart 26:223–235
333. Dembińska A, Wesołowski J (1998) Linearity of regression from non-adjacent order statistics. Metrika 90:215–222

334. Dembińska A, Wesołowski J (2000) Linearity of regression from non-adjacent record values. J Stat Plan Infer 90:195–205
335. Dembo A, Cover TM (1991) Information theoretic inequalities. IEEE Trans Inform Theory 37:1501–1518
336. Deshpande JV, Kochar SC (1983) Dispersive ordering is the same as tail ordering. Adv Appl Probab 15:686–687
337. Devroye L (1986) Non-uniform random variate generation. Springer, New York
338. Dey S, Dey T (2014) Statistical inference for the Rayleigh distribution under progressively Type-II censoring with binomial removal. Appl Math Model 38:974–982
339. Dharmadhikari S, Joag-dev K (1988) Unimodality, convexity, and applications. Academic, Boston
340. Ding C, Tse SK (2013) Design of accelerated life test plans under progressive Type II interval censoring with random removals. J Stat Comput Simul 83:1330–1343
341. Ding C, Yang C, Tse SK (2010) Accelerated life test sampling plans for the Weibull distribution under Type I progressive interval censoring with random removals. J Stat Comput Simul 80:903–914
342. Doganaksoy N, Balakrishnan N (1997) A useful property of best linear unbiased predictors with applications to life-testing. Am Stat 51:22–28
343. Doksum K (1969) Starshaped transformations and the power of rank tests. Ann Math Stat 40:1167–1176
344. Durbin J (1969) Tests for serial correlation in regression analysis based on the periodogram of least-squares residuals. Biometrika 56:1–15
345. Eaton ML (1982) A review of selected topics in multivariate probability inequalities. Ann Stat 10:11–43
346. Ebrahimi N, Maasoumi E, Soofi ES (1999) Ordering univariate distributions by entropy and variance. J Econometrics 90:317–336
347. Ebrahimi N, Soofi E, Zahedi H (2004) Information properties of order statistics and spacings. IEEE Trans Inform Theory 50:177–183
348. Efron B, Johnstone IM (1990) Fisher's information in terms of the hazard rate. Ann Stat 18:38–62
349. Elkahlout GR (2010) Statistical inference using general progressively censored data from the generalized exponential distribution with asymmetric loss. Far East J Theor Stat 30: 117–144
350. El-Neweihi E, Govindarajulu Z (1979) Characterizations of geometric distribution and discrete IFR (DFR) distributions using order statistics. J Stat Plan Infer 3:85–90
351. Engelhardt M, Bain LJ (1978) Tolerance limits and confidence limits on reliability for the two-parameter exponential distribution. Technometrics 20:37–39
352. Epstein B (1954) Truncated life tests in the exponential case. Ann Math Stat 25:555–564
353. Epstein B (1960) Estimation from life test data. Technometrics 2:447–454
354. Esary JD, Proschan F, Walkup DW (1967) Association of random variables, with applications. Ann Math Stat 38:1466–1474
355. Escobar LA, Meeker WQ (1986) Algorithm AS 218: elements of the Fisher information matrix for the smallest extreme value distribution and censored data. J Roy Stat Soc C (Appl Stat) 35:80–86
356. Escobar LA, Meeker WQ (2001) The asymptotic equivalence of the Fisher information matrices for type I and type II censored data from location-scale families. Comm Stat Theory Meth 30:2211–2225
357. Evans I, Nigm AHM (1980) Bayesian prediction for the left truncated exponential distribution. Technometrics 22:201–204
358. Falk M (1989) A note on uniform asymptotic normality of intermediate order statistics. Ann Inst Stat Math 41:19–29
359. Fan DY (1991) The distribution of the product of independent beta variables. Comm Stat Theory Meth 20:4043–4052

360. Fan TH, Wang WL, Balakrishnan N (2008) Exponential progressive step-stress life-testing with link function based on Box-Cox transformation. J Stat Plan Infer 138:2340–2354
361. Feller W (1971) An introduction to probability theory and its applications, vol II, 2nd edn. Wiley, New York
362. Ferguson TS (1967) On characterizing distributions by properties of order statistics. Sankhyā A 29:265–278
363. Fernández AJ (2000) Bayesian inference from type II doubly censored Rayleigh data. Stat Probab Lett 48:393–399
364. Fernández AJ (2004) On estimating exponential parameters with general type II progressive censoring. J Stat Plan Infer 121:135–147
365. Fernández AJ (2004) One- and two-sample prediction based on doubly censored exponential data and prior information. TEST 13:403–416
366. Fernández AJ (2005) Progressively censored variables sampling plans for two-parameter exponential distributions. J Appl Stat 32:823–829
367. Fernández AJ (2009) Bayesian estimation and prediction based on Rayleigh sample quantiles. Qual Quant 44:1–10
368. Fernández AJ (2014) Computing optimal confidence sets for Pareto models under progressive censoring. J Comput Appl Math 258:168–180
369. Fertig KW, Mann NR (1980) Life-test sampling plans for two-parameter Weibull populations. Technometrics 22:165–177
370. Fischer T, Kamps U (2011) On the existence of transformations preserving the structure of order statistics in lower dimensions. J Stat Plan Infer 141:536–548
371. Fischer T, Balakrishnan N, Cramer E (2008) Mixture representation for order statistics from INID progressive censoring and its applications. J Multivariate Anal 99:1999–2015
372. Fisher RA (1930) Inverse probability. Math Proc Cambridge Phil Soc 26:528–535
373. Fisher RA (1934) Two new properties of mathematical likelihood. Proc Roy Soc Lond Ser A 144:285–307
374. Fisher RA (1935) The fiducial argument in statistical inference. Ann Eugenics 6:391–398
375. Fisher RA, Tippett LHC (1928) Limiting forms of the frequency distribution of the largest or smallest member of a sample. Math Proc Cambridge Phil Soc 24:180–190
376. Fleming TR, Harrington DP (2005) Counting processes and survival analysis. Wiley, New York
377. Franco M, Ruiz JM (1995) On characterization of continuous distributions with adjacent order statistics. Statistics 26:375–385
378. Franco M, Ruiz JM (1996) On characterization of continuous distributions by conditional expectation of record values. Sankhyā 58:135–141
379. Franco M, Ruiz JM (1998) Characterization of discrete populations through conditional expectations of order statistics. Stat Papers 39:249–262
380. Franco M, Ruiz JM (1999) Characterization based on conditional expectations of adjacent order statistics: a unified approach. Proc Am Math Soc 127:861–874
381. Franco M, Ruiz JM, Ruiz MC (2002) Stochastic orderings between spacings of generalized order statistics. Probab Eng Inform Sci 16:471–484
382. Fraser DAS (1979) Inference and linear models. McGraw-Hill, New York
383. Freund JE (1961) A bivariate extension of the exponential distribution. J Am Stat Assoc 56:971–977
384. Fu J, Xu A, Tang Y (2012) Objective Bayesian analysis of Pareto distribution under progressive Type-II censoring. Stat Probab Lett 82:1829–1836
385. Gajek L, Gather U (1989) Characterizations of the exponential distribution by failure rate and moment properties of order statistics. In: Hüsler J, Reiss RD (eds) Extreme value theory. Springer, Berlin, pp 114–124
386. Gajek L, Rychlik T (1998) Projection method for moment bounds on order statistics from restricted families: II. independent case. J Multivariate Anal 64:156–182
387. Gajjar A, Khatri C (1969) Progressively censored samples from log-normal and logistic distributions. Technometrics 11:793–803

388. Galambos J (1975) Characterizations of probability distributions by properties of order statistics II (discrete distributions). In: Patil GP, Kotz S, Ord JK (eds) Statistical distributions in scientific work, vol 3. D. Reidel, Boston, pp 89–102
389. Galambos J (1978) The asymptotic theory of extreme order statistics. Wiley, New York
390. Galambos J (1987) The asymptotic theory of extreme order statistics, 2nd edn. Krieger, Malabar
391. Gan G, Bain LJ (1995) Distribution of order statistics for discrete parents with applications to censored sampling. J Stat Plan Infer 44:37–46
392. Ganguly A, Mitra S, Samanta D, Kundu D (2012) Exact inference for the two-parameter exponential distribution under Type-II hybrid censoring. J Stat Plan Infer 142:613–625
393. Gather U, Kamps U, Schweitzer N (1998) Characterizations of distributions via identically distributed functions of order statistics. In: Balakrishnan N, Rao CR (eds) Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 257–290
394. Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the bayesian restauration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741
395. Gerber L (1981) The volume cut off a simplex by a half-space. Pacific J Math 94:311–314
396. Gertsbakh I, Kagan A (1999) Characterization of the Weibull distribution by properties of the Fisher information under type-I censoring. Stat Probab Lett 42:99–105
397. Ghafoori S, Habibi Rad A, Doostparast M (2011) Bayesian two-sample prediction with progressively Type-II censored data for some lifetime models. J Iran Stat Soc (JIRSS) 10:63–86
398. Ghitany ME, Tuan VK, Balakrishnan N (2014) Likelihood estimation for a general class of inverse exponentiated distributions based on complete and progressively censored data. J Stat Comput Simul 84:96–106
399. Ghitany ME, Al-Jarallah RA, Balakrishnan N (2013) On the existence and uniqueness of the MLEs of the parameters of a general class of exponentiated distributions. Statistics 47:605–612
400. Ghitany ME, Alqallaf F, Balakrishnan N (2014) On the likelihood estimation of the parameters of Gompertz distribution based on complete and progressively Type-II censored samples. J Stat Comput Simul 84:1803–1812
401. Ghosh M, Razmpour A (1984) Estimation of the common location parameter of several exponentials. Sankhyā 46:383–394
402. Gibbons J, Chakraborti S (2003) Nonparametric statistical inference. Taylor & Francis, Philadelphia
403. Gibbons DI, Vance LC (1983) Estimators for the 2-parameter Weibull distribution with progressively censored samples. IEEE Trans Reliab 32:95–99
404. Gnedenko BV (1943) Sur la distribution limite du terme maximum d'une série aléatoire. Ann Math 44:423–453
405. Goldberger AS (1962) Best linear unbiased prediction in the generalized linear regression model. J Am Stat Assoc 57:369–375
406. Górny J, Cramer E (2014) Exact likelihood inference for exponential distributions under generalized progressive hybrid censoring schemes. Technical report. RWTH Aachen University, Aachen
407. Gouno E, Balakrishnan N (2001) Step-stress accelerated life test. In: Balakrishnan N, Rao CR (eds) Handbook of statistics. Advances in reliability, vol 20. North Holland, Amsterdam, pp 623–639
408. Gouno E, Sen A, Balakrishnan N (2004) Optimal step-stress test under progressive Type-I censoring. IEEE Trans Reliab 53:388–393
409. Govindarajulu Z (1963) On moments of order statistics and quasi-ranges from normal populations. Ann Math Stat 34:633–651
410. Govindarajulu Z (1963) Relationships among moments of order statistics in samples from two related populations. Technometrics 5:514–518

411. Govindarajulu Z (1966) Best linear estimates under symmetric censoring of the parameters of double exponential population. J Am Stat Assoc 61:248–258
412. Govindarajulu Z (1980) Characterization of the geometric distribution using properties of order statistics. J Stat Plan Infer 4:237–247
413. Graybill FA (1983) Matrices with applications in statistics, 2nd edn. Wadsworth, Belmont
414. Greenwood M (1946) The statistical study of infectious diseases. J Roy Stat Soc 109: 85–110
415. Grudzień Z, Szynal D (1985) On the expected values of kth record values and associated characterizations of distributions. In: Konecny F, Mogyoródi J, Wertz W (eds) Probability and statistical decision theory, vol A. Reidel, Dordrecht, pp 119–127
416. Guenther WC, Patil SA, Uppuluri VRR (1976) One-sided β-content tolerance factors for the two parameter exponential distribution. Technometrics 18:333–340
417. Guilbaud O (1982) Functions of non-iid random vectors expressed as functions of iid random vectors. Scand J Stat 9:229–233
418. Guilbaud O (2001) Exact non-parametric confidence intervals for quantiles with progressive type-II censoring. Scand J Stat 28:699–713
419. Guilbaud O (2004) Exact non-parametric confidence, prediction and tolerance intervals with progressive type-II censoring. Scand J Stat 31:265–281
420. Guilbaud O (2007) Comments on: progressive censoring methodology: an appraisal. TEST 16:279–281
421. Gumbel EJ (1954) The maxima of the mean largest value of the range. Ann Math Stat 25:76–84
422. Gumbel EJ, Mustafi CK (1967) Some analytical properties of bivariate extremal distributions. J Am Stat Assoc 62:569–588
423. Gupta RD, Kundu D (2006) On the comparison of Fisher information of the Weibull and GE distributions. J Stat Plan Infer 136:3130–3144
424. Gupta PL, Gupta RC, Lvin SJ (1996) Analysis of failure time data by Burr distribution. Comm Stat Theory Meth 25:2013–2024
425. Gupta PL, Gupta RC, Tripathi RC (1997) On the monotonic properties of discrete failure rates. J Stat Plan Infer 65:255–268
426. Hahn GJ, Meeker WQ (1991) Statistical intervals: a guide for practitioners. Wiley, New York
427. Haj Ahmad H, Awad A (2009) Optimal two-stage progressive censoring with exponential life-times. Dirasat (Science) 36:134–144
428. Haj Ahmad H, Awad A (2009) Optimality criterion for progressive Type-II right censoring based on Awad sup-entropy measures. J Stat 16:12–27
429. Hájek J, Šidák Z (1967) Theory of rank tests. Academic, New York
430. Han D, Balakrishnan N, Sen A, Gouno E (2006) Corrections on 'Optimal step-stress test under progressive Type-I censoring'. IEEE Trans Reliab 55:613–614
431. Harter HL, Moore AH (1966) Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples. J Am Stat Assoc 61:842–851
432. Hartley HO, David HA (1954) Universal bounds for mean range and extreme observation. Ann Math Stat 25:85–99
433. Hashemi M, Asadi M (2007) Some characterization results on generalized Pareto distribution based on progressive Type-II right censoring. J Iran Stat Soc (JIRSS) 6:99–110
434. Hashemi M, Tavangar M, Asadi M (2010) Some properties of the residual lifetime of progressively Type-II right censored order statistics. Stat Probab Lett 80:848–859
435. Hassanein KM (1969) Estimation of the parameters of the extreme value distribution by use of two or three order statistics. Biometrika 56:429–436
436. Hassanein KM (1971) Percentile estimators for the parameters of the Weibull distribution. Biometrika 58:673–676
437. Hegazy YAS, Green JR (1975) Some new goodness-of-fit tests using order statistics. Appl Stat 24:299–308

438. Helu A, Samawi H, Raqab MZ (2013) Estimation on Lomax progressive censoring using the EM algorithm. J Stat Comput Simul (to appear)
439. Hemmati F, Khorram E (2013) Statistical analysis of the log-normal distribution under Type-II progressive hybrid censoring schemes. Comm Stat Simul Comput 42:52–75
440. Herd RG (1956) Estimation of parameters of a population from a multi-censored sample. Ph.D. thesis, Iowa State College, Ames, Iowa
441. Hiriart-Urruty JB, Lemaréchal C (1993) Convex analysis and minimization algorithms I. Springer, Berlin
442. Hoadley B (1971) Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case. Ann Math Stat 42:1977–1991
443. Hoel DG (1972) A representation of mortality data by competing risks. Biometrics 28: 475–488
444. Hofmann G, Cramer E, Balakrishnan N, Kunert G (2005) An asymptotic approach to progressive censoring. J Stat Plan Infer 130:207–227
445. Hofmann G, Balakrishnan N, Ahmadi J (2005) Characterization of hazard function factorization by Fisher information in minima and upper record values. Stat Probab Lett 72:51–57
446. Hogg RV, Craig AT (1978) Introduction to mathematical statistics, 4th edn. Macmillian Publishing Co., New York
447. Hong CW, Wu JW, Cheng CH (2007) Computational procedure of performance assessment of lifetime index of businesses for the Pareto lifetime model with the right type II censored sample. Appl Comput Math 184:336–350
448. Hong CW, Wu JW, Cheng CH (2008) Computational procedure of performance assessment of lifetime index of Pareto lifetime businesses based on confidence interval. Appl Soft Comput 8:698–705
449. Hong CW, Lee WC, Wu JW (2012) Computational procedure of performance assessment of lifetime index of products for the Weibull distribution with the progressive first-failure-censored sampling plan. J Appl Math 2012:13
450. Horn PS, Schlipf JS (1986) Generating subsets of order statistics with applications to trimmed means and means of trimmings. J Stat Comput Simul 24:83–97
451. Hosono YOH, Kase S (1981) Design of single sampling plans for doubly exponential characteristics. In: Lenz HJ, Wetherill GB, Wilrich PT (eds) Frontiers in quality control. Physica, Vienna, pp 94–112
452. Houben A (1998) Verallgemeinerte Ordnungsstatistiken bei zufälligem Stichprobenumfang. Ph.D. thesis, RWTH Aachen University, Aachen, Germany
453. Hu T (1994) Statistical dependence of multivariate distributions and stationary Markov chains with applications. Ph.D. thesis, Department of Mathematics, University of Science and Technology of China
454. Hu T (1995) Monotone coupling and stochastic ordering of order statistics. Syst. Sci. Math. Sci. (English Series) 8:209–214
455. Hu T, Wei Y (2001) Stochastic comparisons of spacings from restricted families of distributions. Stat Probab Lett 53:91–99
456. Hu T, Zhuang W (2003) Stochastic comparisons of m-spacings. Technical report. Department of Statistics and Finance, University of Science and Technology of China, Hefei
457. Hu T, Zhuang W (2005) A note on stochastic comparisons of generalized order statistics. Stat Probab Lett 72:163–170
458. Huang JS (1975) A note on order statistics from Pareto distribution. Scand Actuar J 1975:187–190
459. Huang JS (1997) Sharp bounds for the expected value of order statistics. Stat Probab Lett 33:105–107
460. Huang JS, Ghosh M (1982) A note on strong unimodality of order statistics. J Am Stat Assoc 77:929–930
461. Huang JS, Ghosh M (1983) Corrigenda: "a note on strong unimodality of order statistics" [J Am Stat Assoc (1982) 77(380):929–930]. J Am Stat Assoc 78:1008

462. Huang SR, Wu SJ (2008) Reliability sampling plans under progressive Type-I interval censoring using cost functions. IEEE Trans Reliab 57:445–451
463. Huang SR, Wu SJ (2012) Bayesian estimation and prediction for Weibull model with progressive censoring. J Stat Comput Simul 82:1607–1620
464. Huber PJ (1964) Robust estimation of a location parameter. Ann Math Stat 35:73–101
465. Huber PJ (1981) Robust statistics. Wiley, New York
466. Huffer FW, Lin CT (2001) Computing the joint distribution of general linear combinations of spacings or exponential variates. Stat Sinica 11:1141–1157
467. Hutson AD (1999) Calculating nonparametric confidence intervals for quantiles using fractional order statistics. J Appl Stat 26:343–353
468. Ibragimov IA (1956) On the composition of unimodal distributions. Teor Veroyatnost i Primenen 1:283–288
469. Iliopoulos G, Balakrishnan N (2009) Conditional independence of blocked ordered data. Stat Probab Lett 79:1008–1015
470. Iliopoulos G, Balakrishnan N (2011) Exact likelihood inference for Laplace distribution based on Type-II censored samples. J Stat Plan Infer 141:1224–1239
471. Ishii G (1978) Tokeiteki Yosoku (statistical prediction). In: Basic Sugaku, vol 7. Gendai-Sugakusha, Tokyo (in japanese)
472. Iwińska M (1986) On the characterizations of the exponential distribution by order statistics and record values. Fasc Math 16:101–107
473. Izadi M, Khaledi BE (2007) Progressive Type II censored order statistics and their concomitants: some stochastic comparisons results. J Iran Stat Soc (JIRSS) 6:111–124
474. Jaheen ZF (2003) Prediction of progressive censored data from the Gompertz model. Comm Stat Simul Comput 32:663–676
475. Jaheen ZF (2005) Estimation based on generalized order statistics from the Burr model. Comm Stat Theory Meth 34:785–794
476. Jalali A, Watkins AJ (2009) On maximum likelihood estimation for the two parameter Burr XII distribution. Comm Stat Theory Meth 38:1916–1926
477. Jambunathan MV (1954) Some properties of beta and gamma distributions. Ann Math Stat 25:401–405
478. James IR (1986) On estimating equations with censored data. Biometrika 73:35–42
479. Joarder A, Krishna H, Kundu D (2009) On Type-II progressive hybrid censoring. J Mod Appl Stat Meth 8:534–546
480. Jogdeo K (1978) On a probability bound of Marshall and Olkin. Ann Stat 6:232–234
481. Johnson LG (1964) The statistical treatment of fatigue experiments. Elsevier, Amsterdam
482. Johnson NL, Kotz S (1972) Distributions in statistics: continuous multivariate distributions. Wiley, New York
483. Johnson NL, Kotz S, Balakrishnan N (1994) Continuous univariate distributions, vol 1, 2nd edn. Wiley, New York
484. Johnson NL, Kotz S, Balakrishnan N (1995) Continuous univariate distributions, vol 2, 2nd edn. Wiley, New York
485. Joshi PC (1972) Efficient estimation of the mean of an exponential distribution when an outlier is present. Technometrics 14:137–143
486. Joshi PC (1978) Recurrence relations between moments of order statistics from exponential and truncated exponential distributions. Sankhyā Ser B 39:362–371
487. Joshi PC (1979) A note on the moments of order statistics from doubly truncated exponential distribution. Ann Inst Stat Math 31:321–324
488. Joshi PC (1982) A note on the mixed moments of order statistics from exponential and truncated exponential distributions. J Stat Plan Infer 6:13–16
489. Jun CH, Lee H, Lee SH, Balamurali S (2010) A variables repetitive group sampling plan under failure-censored reliability tests for Weibull distribution. J Appl Stat 37:453–460
490. Kalbfleisch JD, Prentice RL (2011) The statistical analysis of failure time data, 2nd edn. Wiley, Hoboken

491. Kale BK, Sinha SK (1971) Estimation of expected life in the presence of an outlier observation. Technometrics 13:755–759
492. Kałuszka M, Okolewski A (2004) Tsallis entropy bounds for generalized order statistics. Probab Math Stat 24:253–262
493. Kambo N (1978) Maximum likelihood estimators of the location and scale parameters of the exponential distribution from a censored sample. Comm Stat Theory Meth 7:1129–1132
494. Kaminskiy MP, Krivtsov VV (2005) A simple procedure for Bayesian estimation of the Weibull distribution. IEEE Trans Reliab 54:612–616
495. Kaminsky KS, Nelson PI (1975) Best linear unbiased prediction of order statistics in location and scale families. J Am Stat Assoc 70:145–150
496. Kaminsky KS, Rhodin L (1985) Maximum likelihood prediction. Ann Inst Stat Math 37:507–517
497. Kamps U (1990) Characterizations of the exponential distribution by weighted sums of iid random variables. Stat Papers 31:233–237
498. Kamps U (1995) A concept of generalized order statistics. Teubner, Stuttgart
499. Kamps U (1995) A concept of generalized order statistics. J Stat Plan Infer 48:1–23
500. Kamps U (1998) Characterizations of distributions by recurrence relations and identities for moments of order statistics. In: Balakrishnan N, Rao CR (eds) Order statistics: theory and methods. Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 291–311
501. Kamps U (1998) Subranges of generalized order statistics from exponential distributions. Fasc Math 28:63–70
502. Kamps U (1999) Order statistics, generalized. In: Kotz S, Read CB, Banks DL (eds) Encyclopedia of statistical sciences, update vol 3. Wiley, New York, pp 553–557
503. Kamps U, Cramer E (2001) On distributions of generalized order statistics. Statistics 35:269–280
504. Kamps U, Keseling C (2003) A theorem of Rossberg for generalized order statistics. Sankhyā Ser A 65:259–270
505. Kane VE (1986) Process capability indices. J Qual Tech 18:41–52
506. Kang SB, Seo JI (2011) Estimation in an exponentiated half logistic distribution under progressively Type-II censoring. Commun Korean Stat Soc 18:657–666
507. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. J Am Stat Assoc 53:457–481
508. Kappenman RF (1975) Conditional confidence intervals for double exponential distribution parameters. Technometrics 17:233–235
509. Karlin S (1968) Total positivity, vol I. Stanford University Press, Stanford
510. Karlin S, Rinott Y (1980) Classes of orderings of measures and related correlation inequalities. I. Multivariate totally positive distributions. J Multivariate Anal 10:467–498
511. Kateri M, Balakrishnan N (2008) Inference for a simple step-stress model with Type-II censoring, and Weibull distributed lifetimes. IEEE Trans Reliab 57:616–626
512. Keating JP, Mason RL, Sen PK (1993) Pitman's measure of closeness: a comparison of statistical estimators. SIAM, Society for Industrial and Applied Mathematics, Philadelphia
513. Keilson J, Sumita U (1982) Uniform stochastic ordering and related inequalities. Can J Stat 10:181–198
514. Kemp CD, Kemp W (1987) Rapid generation of frequency tables. Appl Stat 36:277–282
515. Kendall MG, Stuart A, Ord JK (1987) Kendall's advanced theory of statistics, vol I, 5th edn. Griffin, London
516. Keseling C (1999) Charakterisierungen von Wahrscheinlichkeitsverteilungen durch verallgemeinerte Ordnungstatistiken. Ph.D. thesis, RWTH Aachen University, Aachen, Germany
517. Keseling C (1999) Conditional distributions of generalized order statistics and some characterizations. Metrika 49:27–40
518. Khaledi BE (2005) Some new results on stochastic orderings between generalized order statistics. J Iran Stat Soc (JIRSS) 4:35–49

519. Khaledi BE, Kochar SC (1999) Stochastic orderings between distributions and their sample spacings II. Stat Probab Lett 44:161–166
520. Khaledi BE, Kochar SC (2002) Stochastic ordering among order statistics and sample spacings. In: Misra JC (ed) Uncertainty and optimality: probability, statistics and operations research. World Scientific Publications, Singapore, pp 167–203
521. Khaledi BE, Kochar SC (2004) Ordering convolutions of gamma random variables. Sankhyā 66:466–473
522. Khaledi BE, Kochar SC (2005) Dependence orderings for generalized order statistics. Stat Probab Lett 73:357–367
523. Khaledi BE, Kochar SC (2006) Weibull distribution: some stochastic comparisons results. J Stat Plan Infer 136:3121–3129
524. Khamis I, Higgins J (1998) A new model for step-stress testing. IEEE Trans Reliab 47: 131–134
525. Khan AH, Abu-Salih MS (1988) Characterization of the Weibull and the inverse Weibull distributions through conditional moments. J Inform Optim Sci 9:355–362
526. Khan AH, Abu-Salih MS (1989) Characterizations of probability distributions by conditional expectations of order statistics. Metron 47:171–181
527. Khan AH, Khan IA (1986) Characterization of the Pareto and the power function distributions. J Stat Res 20:71–79
528. Khan AH, Khan IA (1987) Moments of order statistics from Burr distribution and its characterization. Metron 45:21–29
529. Kim C, Han K (2009) Estimation of the scale parameter of the Rayleigh distribution under general progressive censoring. J Korean Stat Soc 38:239–246
530. Kim C, Han K (2010) Estimation of the scale parameter of the half-logistic distribution under progressively type II censored sample. Stat Papers 51:375–387
531. Kim C, Jung J, Chung Y (2011) Bayesian estimation for the exponentiated Weibull model under Type II progressive censoring. Stat Papers 52:53–70
532. Kinaci I (2013) A generalization of flexible progressive censoring. Pakistan J Stat 29: 377–387
533. King JR (1971) Probability charts for decision making. Industrial Press, New York
534. Klakattawi HS, Baharith LA, Al-Dayian GR (2011) Bayesian and non bayesian estimations on the exponentiated modified Weibull distribution for progressive censored sample. Comm Stat Simul Comput 40:1291–1309
535. Klakattawi HS, Baharith LA, AL-Dayian GR (2012) Bayesian predictions of progressive censored data from the exponentiated modified Weibull distribution. Can J Comput Math Nat Sci Eng Med 3:247–255
536. Klein JP, Moeschberger ML (2003) Survival analysis: techniques for censored and truncated data, 2nd edn. Springer, New York
537. Kochar SC (1981) A new distribution-free test for the equality of two failure rates. Biometrika 68:423–426
538. Kochar SC (1996) Dispersive ordering of order statistics. Stat Probab Lett 27:271–274
539. Kochar SC (2006) Lorenz ordering of order statistics. Stat Probab Lett 76:1855–1860
540. Kochar SC, Korwar R (1996) Stochastic orders for spacings of heterogeneous exponential random variables. J Multivariate Anal 57:69–83
541. Kochar SC, Rojo J (1996) Some new results on stochastic comparisons of spacings from heterogeneous exponential distributions. J Multivariate Anal 59:272–281
542. Kocherlakota S, Balakrishnan N (1986) One- and two-sided sampling plans based on the exponential distribution. Naval Res Logist Quart 33:513–522
543. Kong F (1998) Parameter estimation under multiply type-II censoring. In: Balakrishnan N, Rao CR (eds) Order statistics: applications, vol 17. North-Holland, Amsterdam, pp 315–335
544. Korwar R (2003) On the likelihood ratio order for progressive type II censored order statistics. Sankhyā Ser A 65:793–798

545. Kotlarski I (1962) On groups of n independent random variables whose product follows a beta distribution. Colloq Math 9:325–332
546. Kotz S, Kozubowski TJ, Podgórski K (2001) The Laplace distribution and generalizations. A revisit with applications to communications, economics, engineering, and finance. Birkhäuser, Boston
547. Kotz S, Lumelskii Y, Pensky M (2003) The stress-strength model and its generalizations: theory and applications. World Scientific Publications, Singapore
548. Kozubowski TJ, Panorska AK, Qeadan F, Gershunov A, Rominger D (2009) Testing exponentiality versus Pareto distribution via likelihood ratio. Comm Stat Simul Comput 38:118–139
549. Kräuter AR (1983) Permanents: a concise survey (Permanenten: Ein kurzer Überblick). Sémin Lothar Comb 9:1–34 (in German)
550. Krishna H, Kumar K (2011) Reliability estimation in Lindley distribution with progressively type II right censored sample. Math Comput Simul 82:281–294
551. Krishna H, Kumar K (2013) Reliability estimation in generalized inverted exponential distribution with progressively type II censored sample. J Stat Comput Simul 83:1007–1019
552. Krishna H, Malik M (2012) Reliability estimation in Maxwell distribution with progressively Type-II censored data. J Stat Comput Simul 82:623–641
553. Krishnamoorthy K, Mathew T (2009) Statistical tolerance regions: theory, applications, and computation. Wiley, Hoboken
554. Kullback S (1959) Information theory and statistics. Wiley, New York
555. Kulldorff G (1961) Contributions to the theory of estimation from grouped and partially grouped samples. Wiley, New York
556. Kulldorff G, Vännman K (1973) Estimation of the location and scale parameters of a Pareto distribution by linear functions of order statistics. J Am Stat Assoc 68:218–227
557. Kundu D (2008) Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. Technometrics 50:144–154
558. Kundu D, Balakrishnan N (2009) Point and interval estimation for a simple step-stress model with random stress-change time. J Probab Stat Sci 7:113–126
559. Kundu D, Gupta RD (2006) Estimation of $P[Y < X]$ for Weibull distributions. IEEE Trans Reliab 55:270–280
560. Kundu D, Joarder A (2006) Analysis of Type-II progressively hybrid censored competing risks data. J Mod Appl Stat Meth 5:152–170
561. Kundu D, Joarder A (2006) Analysis of Type-II progressively hybrid censored data. Comput Stat Data Anal 50:2509–2528
562. Kundu D, Pradhan B (2009) Bayesian inference and life testing plans for generalized exponential distribution. Sci China Ser A 52:1373–1388
563. Kundu D, Pradhan B (2011) Bayesian analysis of progressively censored competing risks data. Sankhyā Ser B 73:276–296
564. Kundu D, Kannan N, Balakrishnan N (2004) Analysis of progressively censored competing risks data. In: Balakrishnan N, Rao CR (eds) Advances in survival analysis. Handbook of statistics, vol 23. Elsevier, Amsterdam, pp 331–348
565. Kundu D, Samanta D, Ganguly A, Mitra S (2013) Bayesian analysis of different hybrid and progressive life tests. Comm Stat Simul Comput 42:2160–2173
566. Kuş C, Kaya MF (2007) Estimation for the parameters of the Pareto distribution under progressive censoring. Comm Stat Theory Meth 36:1359–1365
567. Lai CD, Xie M, Murthy DNP (2003) A modified Weibull distribution. IEEE Trans Reliab 52:33–37
568. Laumen B, Cramer E (2013) Inference for the lifetime performance index with progressively Type-II censored samples from gamma distributions. Technical report. RWTH Aachen University, Aachen
569. Lawless JF (1971) A prediction problem concerning samples from the exponential distribution, with application in life testing. Technometrics 13:725–730

570. Lawless JF (1973) Conditional versus unconditional confidence intervals for the parameters of the Weibull distribution. J Am Stat Assoc 68:665–669
571. Lawless JF (1973) On the estimation of safe life when the underlying life distribution is Weibull. Technometrics 15:857–865
572. Lawless JF (1978) Confidence interval estimation for the Weibull and extreme value distributions. Technometrics 20:355–368
573. Lawless JF (1980) Inference in the generalized gamma and log gamma distributions. Technometrics 22:409–419
574. Lawless JF (1982) Statistical models and methods for lifetime data. Wiley, New York
575. Lawless JF (2003) Statistical models and methods for lifetime data, 2nd edn. Wiley, New York
576. Leadbetter MR, Lindgren G, Rootzén H (1983) Extremes and related properties of random sequences and processes. Springer, New York
577. Lee WC, Wu JW, Hong CW (2009) Assessing the lifetime performance index of products from progressively type II right censored data using Burr XII model. Math Comput Simul. 79:2167–2179
578. Lee WC, Wu JW, Hong CW (2009) Assessing the lifetime performance index of products with the exponential distribution under progressively type II right censored samples. J Comput Appl Math 231:648–656
579. Lee WC, Wu JW, Hong ML, Lin LS, Chan RL (2011) Assessing the lifetime performance index of Rayleigh products based on the Bayesian estimation under progressive Type II right censored samples. J Comput Appl Math 235:1676–1688
580. Lee HM, Lee WC, Lei CL, Wu JW (2011) Computational procedure of assessing lifetime performance index of Weibull lifetime products with the upper record values. Math Comput Simul 81:1177–1189
581. Lee HM, Wu JW, Lei CL, Hung WL (2011) Implementing lifetime performance index of products with two-parameter exponential distribution. Inter J Syst Sci 42:1305–1321
582. Lehmann EL, Casella G (1998) Theory of point estimation, 2nd edn. Springer, New York
583. Lemon GH (1975) Maximum likelihood estimation for the three parameter Weibull distribution based on censored samples. Technometrics 17:247–254
584. Lewis T, Thompson J (1981) Dispersive distributions and the connection between dispersivity and strong unimodality. J Appl Probab 18:76–90
585. Li Y (1999) A note on the number of records near the maximum. Stat Probab Lett 43: 153–158
586. Li H, Shaked M (1993) Stochastic majorization of stochastically monotone families of random variables. Adv Appl Probab 25:895–913
587. Li X, Shi Y, Wei J, Chai J (2007) Empirical Bayes estimators of reliability performances using LINEX loss under progressively Type-II censored samples. Math Comput Simul 73:320–326
588. Li F, Shi Y, Tian Y (2008) Bayesian estimation of a Weibull model based on progressively censored data. Chinese J Eng Math 25:641–650
589. Liang T (2014) Comment on 'Exact Bayesian variable sampling plans for the exponential distribution with progressive hybrid censoring'. J Stat Comput Simul 84:468–476
590. Liang H, Huang S, Bai F (2006) A hybrid algorithm for computing permanents of sparse matrices. Appl Math Comput 172:708–716
591. Libby DL, Novick MR (1982) Multivariate generalized beta distributions with applications to utility assessment. J Educ Stat 7:271–294
592. Lieberman GJ, Resnikoff GJ (1955) Sampling plans for inspection by variables. J Am Stat Assoc 50:457–516
593. Lieblein J (1953) On the exact evaluation of the variances and covariances of order statistics in samples from the extreme-value distribution. Ann Math Stat 24:282–287
594. Lieblein J (1955) On moments of order statistics from the Weibull distribution. Ann Math Stat 26:330–333
595. Lieblein J, Zelen M (1956) Statistical investigation of the fatigue life of deep groove ball bearings. J Res Nat Bur Stand 57:273–316

596. Lihong S, Xinsheng Z (2005) Stochastic comparisons of order statistics from gamma distributions. J Multivariate Anal 93:112–121
597. Likeš J (1967) Distributions of some statistics in samples from exponential and power-function populations. J Am Stat Assoc 62:259–271
598. Lillo RE, Nanda AK, Shaked M (2001) Preservation of some likelihood ratio stochastic orders by order statistics. Stat Probab Lett 51:111–119
599. Lin CT, Balakrishnan N (2011) Asymptotic properties of maximum likelihood estimators based on progressive Type-II censoring. Metrika 74:349–360
600. Lin CT, Huang YL (2012) On progressive hybrid censored exponential distribution. J Stat Comput Simul 82:689–709
601. Lin YJ, Lio YL (2012) Bayesian inference under progressive type-I interval censoring. J Appl Stat 39:1811–1824
602. Lin CT, Wu SJS, Balakrishnan N (2004) Interval estimation of parameters of log-gamma distribution based on progressively censored data. Comm Stat Theory Meth 33:2595–2626
603. Lin CT, Wu SJS, Balakrishnan N (2006) Inference for log-gamma distribution based on progressively Type-II censored data. Comm Stat Theory Meth 35:1271–1292
604. Lin CT, Wu SJS, Balakrishnan N (2006) Monte Carlo methods for Bayesian inference on the linear hazard rate distribution. Comm Stat Simul Comput 35:575–590
605. Lin CT, Huang YL, Balakrishnan N (2008) A new method for goodness-of-fit testing based on Type-II right censored samples. IEEE Trans Reliab 57:633–642
606. Lin CT, Wu SJS, Balakrishnan N (2009) Planning life tests with progressively Type-I interval censored data from the lognormal distribution. J Stat Plan Infer 139:54–61
607. Lin CT, Ng HKT, Chan PS (2009) Statistical inference of Type-II progressively hybrid censored data with Weibull lifetimes. Comm Stat Theory Meth 38:1710–1729
608. Lin CT, Huang YL, Balakrishnan N (2011) Exact Bayesian variable sampling plans for the exponential distribution with progressive hybrid censoring. J Stat Comput Simul 81: 873–882
609. Lin CT, Balakrishnan N, Wu SJS (2011) Planning life tests based on progressively Type-I grouped censored data from the Weibull distribution. Comm Stat Simul Comput 40: 574–595
610. Lin CT, Chou CC, Huang YL (2012) Inference for the Weibull distribution with progressive hybrid censoring. Comput Stat Data Anal 56:451–467
611. Lin CT, Huang YL, Balakrishnan N (2013) Correction on 'Exact Bayesian variable sampling plans for the exponential distribution with progressive hybrid censoring, J Stat Comput Simul 81:873–882 (2011)'. J Stat Comput Simul 83:402–404
612. Lindley D (1980) Approximate Bayesian methods. Trabajos de Estadistica Y de Investigacion Operativa 31:223–245
613. Lio YL, Tsai TR (2012) Estimation of $\delta = P(X < Y)$ for Burr XII distribution based on the progressively first failure-censored samples. J Appl Stat 39:309–322
614. Lio YL, Chen DG, Tsai TR (2011) Parameter estimations for generalized Rayleigh distribution under progressively Type-I interval censored data. Open J Stat 1:46–57
615. Lio YL, Tsai TR, Chiang JY (2012) Estimation on the lower confidence limit of the breaking strength percentiles under progressive type-II censoring. J Chin Inst Ind Eng 29:16–29
616. López-Blázquez F, Miño B (1998) A characterization of the geometric distribution. Stat Papers 39:231–236
617. López Blázquez F, Moreno Rebollo JL (1997) A characterization of distributions based on linear regression of order statistics and record values. Sankhyā 59:311–323
618. Louis TA (1982) Finding the observed information matrix when using the EM algorithm. J Roy Stat Soc Ser B 44:226–233
619. Lu Y, Storer B (2001) A tampered Brownian motion process model for partial step-stress accelerated life testing. J Stat Plan Infer 94:15–24
620. Ludwig O (1959) Ungleichungen für Extremwerte und andere Ranggrößen in Anwendung auf biometrische Probleme. Biom. Z 1:203–209

621. Lurie D, Hartley HO (1972) Machine-generation of order statistics for Monte Carlo computations. Am Stat 26:26–27
622. Lurie D, Mason RL (1973) Empirical investigation of several techniques for computer generation of order statistics. Comm Stat 2:363–371
623. Lwin T (1973) Estimation of the tail of the Paretian law. Skand Aktuarietidskr 1972: 170–178
624. Ma C (1997) A note on stochastic ordering of order statistics. J Appl Probab 34:785–789
625. Madi MT (1993) Multiple step-stress accelerated life test: the tampered failure rate model. Comm Stat Theory Meth 22:295–306
626. Madi MT, Raqab MZ (2007) Bayesian prediction of rainfall records using the generalized exponential distribution. Environmetrics 18:541–549
627. Madi MT, Raqab MZ (2009) Bayesian inference for the generalized exponential distribution based on progressively censored data. Comm Stat Theory Meth 38:2016–2029
628. Mahalanobis PC, Bose SS, Ray PR, Banerji SK (1934) Tables of random samples from a normal population. Sankhyā 1:289–328
629. Mahmoud MAW, Rostom NKA, Yhieha NM (2008) Recurrence relations for the single and product moments of progressively type-II right censored order statistics from the generalized Pareto distribution. J Egypt Math Soc 16:61–74
630. Mahmoud MAW, Soliman AA, Ellah AHA, El-Sagheer RM (2013) Bayesian estimation using MCMC approach based on progressive first-failure censoring from generalized Pareto distribution. Am J Theor Appl Stat 2:128–141
631. Majumdar H, Sen PK (1978) Nonparametric tests for multiple regression under progressive censoring. J Multivariate Anal 8:73–95
632. Malik J (1970) A characterization of the Pareto distribution. Skand Aktuarietidskr 53: 115–117
633. Malmquist S (1950) On a property of order statistics from a rectangular distribution. Skand Aktuarietidskr 33:214–222
634. Mann NR (1969) Exact three-order-statistic confidence bounds on reliable life for a Weibull model with progressive censoring. J Am Stat Assoc 64:306–315
635. Mann NR (1969) Optimum estimators for linear functions of location and scale parameters. Ann Math Stat 40:2149–2155
636. Mann NR (1971) Best linear invariant estimation for Weibull parameters under progressive censoring. Technometrics 13:521–533
637. Mao T, Hu T (2010) Stochastic properties of INID progressive Type-II censored order statistics. J Multivariate Anal 101:1493–1500
638. Marohn F (2002) A characterization of generalized Pareto distributions by progressive censoring schemes and goodness-of-fit tests. Comm Stat Theory Meth 31:1055–1065
639. Marohn F (2005) Exponential inequalities for tail probabilities of generalized order statistics. Metrika 61:251–260
640. Marshall AW, Olkin I (2007) Life distributions: structure of nonparametric, semiparametric, and parametric families. Springer, New York
641. Maswadah M (2003) Conditional confidence interval estimation for the inverse Weibull distribution based on censored generalized order statistics. J Stat Comput Simul 73: 887–898
642. Maturi TA, Coolen-Schrijner P, Coolen FP (2010) Nonparametric predictive comparison of lifetime data under progressive censoring. J Stat Plan Infer 140:515–525
643. McNeil AJ, Nešlehová J (2009) Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions. Ann Stat 37:3059–3097
644. McNichols D, Padgett W (1988) Inference for step-stress accelerated life tests under arbitrary right-censorship. J Stat Plan Infer 20:169–179
645. Meeker WQ, Escobar LA (1998) Statistical methods for reliability data. Wiley, New York
646. Meeker WQ, Hahn GJ (1985) How to Plan accelerated life tests. ASQC basic references in quality control: statistical techniques, vol 10. The American Society for Quality Control, Milwaukee

647. Michael JR, Schucany WR (1979) A new approach to testing goodness of fit for censored samples. Technometrics 21:435–441
648. MIL-STD-781-C (1977) Reliability design qualification and production acceptance tests: exponential distribution. U.S. Government Printing Office, Washington
649. Miller R, Nelson W (1983) Optimum simple step-stress plans for accelerated life testing. IEEE Trans Reliab R-32:59–65
650. Minc H (1978) Permanents. Addison-Wesley, Reading
651. Misra N, van der Meulen EC (2003) On stochastic properties of m-spacings. J Stat Plan Infer 115:683–697
652. Mohie El-Din MM, Shafay AR (2013) One- and two-sample Bayesian prediction intervals based on progressively Type-II censored data. Stat Papers 54:287–307
653. Mohie El-Din MM, Mahmoud MAW, Abo Youssef SE (1991) Moments of order statistics from parabolic and skewed distributions and a characterization of Weibull distribution. Comm Stat Simul Comput 20:639–645
654. Mokhtari EB, Rad AH, Yousefzadeh F (2011) Inference for Weibull distribution based on progressively Type-II hybrid censored data. J Stat Plan Infer 141:2824–2838
655. Montanari GC, Cacciari M (1988) Progressively-censored aging tests on XPLE-insulated cable models. IEEE Trans Electr Insul 23:365–372
656. Montgomery DC (2013) Statistical quality control: A modern introduction, 7th edn. Wiley, Hoboken
657. Moriguti S (1953) A modification of Schwarz's inequality with applications to distributions. Ann Math Stat 24:107–113
658. Mosteller F (1946) On some useful "inefficient" statistics. Ann Math Stat 17:377–408
659. Müller A, Stoyan D (2002) Comparison methods for stochastic models and risks. Wiley, Chichester
660. Nagaraja HN (1977) On a characterization based on record values. Aust J Stat 19:70–73
661. Nagaraja HN (1982) On the non-Markovian structure of discrete order statistics. J Stat Plan Infer 7:29–33
662. Nagaraja HN (1986) A note on the conditional Markov property of discrete order statistics. J Stat Plan Infer 13:37–43
663. Nagaraja HN (1988) Some characterizations of continuous distributions based on regressions of adjacent order statistics and record values. Sankhyā Ser A 50:70–73
664. Nagaraja HN (1988) Some characterizations of discrete distributions based on linear regressions of adjacent order statistics. J Stat Plan Infer 20:65–75
665. Nagaraja HN (1992) Order statistics from discrete distributions. Statistics 23:189–216. With discussion and a rejoinder by the author.
666. Nagaraja HN (1994) Tukey's linear sensitivity and order statistics. Ann Inst Stat Math 46:757–768
667. Nagaraja HN (2007) Comments on: progressive censoring methodology: an appraisal. TEST 16:260–261
668. Nagaraja HN, Abo-Eleneen ZA (2003) Fisher information in order statistics. Pakistan J Stat 19:161–173
669. Nagaraja HN, Srivastava RC (1987) Some characterizations of geometric type distributions based on order statistics. J Stat Plan Infer 17:181–191
670. Nasri-Roudsari D (1996) Extreme value theory of generalized order statistics. J Stat Plan Infer 55:281–297
671. Nasri-Roudsari D, Cramer E (1999) On the convergence rates of extreme generalized order statistics. Extremes 2:421–447
672. Nelson LS (1963) Tables for a precedence life test. Technometrics 5:491–499
673. Nelson LS (1986) Precedence life test. In: Kotz S, Johnson NL (eds) Encyclopedia of statistical science, vol 7. Wiley, New York, pp 134–136
674. Nelson W (1969) Hazard plotting for incomplete failure data. J Qual Tech 1:27–52
675. Nelson W (1980) Accelerated life testing: step-stress models and data analyses. IEEE Trans Reliab R-29:103–108

676. Nelson W (1982) Applied life data analysis. Wiley, New York
677. Nelson W (1990) Accelerated testing: statistical models, test plans, and data analyses. Wiley, Hoboken
678. Nelson W, Meeker WQ (1978) Theory for optimum accelerated censored life tests for Weibull and extreme value distributions. Technometrics 20:171–177
679. Neumaier A (2001) Introduction to numerical analysis. Cambridge University Press, Cambridge
680. Nevzorov VB (2001) Records: mathematical theory. Translations of mathematical monographs, vol 194. American Mathematical Society, Providence, Rhode Island
681. Ng HKT (2005) Parameter estimation for a modified Weibull distribution, for progressively Type-II censored samples. IEEE Trans Reliab 54:374–381
682. Ng HKT, Balakrishnan N (2002) Wilcoxon-type rank-sum precedence tests: large-sample approximation and evaluation. Appl Stoch Models Bus Ind 18:271–286
683. Ng HKT, Balakrishnan N (2004) Wilcoxon-type rank-sum precedence tests. Aust N Z J Stat 46:631–648
684. Ng HKT, Balakrishnan N (2005) Weighted precedence and maximal precedence tests and an extension to progressive censoring. J Stat Plan Infer 135:197–221
685. Ng HKT, Balakrishnan N (2006) Precedence testing. In: Kotz S, Balakrishnan N, Read CB, Vidakovic B (eds) Encyclopedia of statistical science, vol 9, 2nd edn. Wiley, Hoboken, pp 6317–6323
686. Ng HKT, Balakrishnan N (2010) Precedence-type test based on Kaplan-Meier estimator of cumulative distribution function. J Stat Plan Infer 140:2295–2311
687. Ng HKT, Wang Z (2009) Statistical estimation for the parameters of Weibull distribution based on progressively type-I interval censored sample. J Stat Comput Simul 79:145–159
688. Ng HKT, Chan PS, Balakrishnan N (2002) Estimation of parameters from progressively censored data using EM algorithm. Comput Stat Data Anal 39:371–386
689. Ng HKT, Chan PS, Balakrishnan N (2004) Optimal progressive censoring plans for the Weibull distribution. Technometrics 46:470–481
690. Ng HKT, Kundu D, Chan PS (2009) Statistical analysis of exponential lifetimes under an adaptive Type-II progressive censoring scheme. Naval Res Logist 56:687–698
691. Ng HKT, Luo L, Hu Y, Duan F (2012) Parameter estimation of three-parameter Weibull distribution based on progressively Type-II censored samples. J Stat Comput Simul 82:1661–1678
692. Nigm EM, Abo-Eleneen ZA (2008) Estimation of the parameters of the inverse Weibull distribution with progressive censoring data. In: Ahsanullah M (ed) Applied statistics research progress. Nova Science Publisher, New York, pp 39–47
693. Oakes D, Dasu T (1990) A note on residual life. Biometrika 77:409–410
694. Ogawa J (1951) Contributions to the theory of systematic statistics. I. Osaka Math J 3: 175–213
695. Oja H (1981) On location, scale, skewness and kurtosis of univariate distributions. Scand J Stat 8:154–168
696. Ouyang LY (1993) Characterizations of the uniform distribution by conditional expectations. Inter J Inform Manag Sci 4:107–111
697. Ouyang LY (1995) Characterization through the conditional expectation of a function of one order statistic relative to an adjacent one. Sankhyā Ser A 57:500–503
698. Pakes AG, Li Y (1998) Limit laws for the number of near maxima via the Poisson approximation. Stat Probab Lett 40:395–401
699. Pakes AG, Steutel FW (1997) On the number of records near the maximum. Aust N Z J Stat 39:179–192
700. Pakyari R, Balakrishnan N (2012) A general purpose approximate goodness-of-fit test for progressively Type-II censored data. IEEE Trans Reliab 61:238–244
701. Pakyari R, Balakrishnan N (2013) Goodness-of-fit tests for progressively Type-II censored data from location-scale distributions. J Stat Comput Simul 83:167–178

702. Papadatos N (1997) Exact bounds for the expectation of order statistics from non-negative populations. Ann Inst Stat Math 49:727–736
703. Pareek B, Kundu D, Kumar S (2009) On progressively censored competing risks data for Weibull distributions. Comput Stat Data Anal 53:4083–4094
704. Park S (1996) Fisher information in order statistics. J Am Stat Assoc 91:385–390
705. Park S (2003) On the asymptotic Fisher information in order statistics. Metrika 57:71–80
706. Park S (2005) Testing exponentiality based on Kullback-Leibler information with the Type-II censored data. IEEE Trans Reliab 54:22–26
707. Park S (2013) On Kullback-Leibler information of order statistics in terms of the relative risk. Metrika (to appear)
708. Park S, Ng HKT (2012) Missing information and an optimal one-step plan in a Type II progressive censoring scheme. Stat Probab Lett 82:396–402
709. Park S, Zheng G (2004) Equal Fisher information in order statistics. Sankhyā 66:20–34
710. Park S, Balakrishnan N, Zheng G (2008) Fisher information in hybrid censored data. Stat Probab Lett 78:2781–2786
711. Park S, Balakrishnan N, Kim SW (2011) Fisher information in progressive hybrid censoring schemes. Statistics 45:623–631
712. Parsi S, Ganjali M, Farsipour NS (2009) Simultaneous confidence intervals for the parameters of Pareto distribution under progressive censoring. Comm Stat Theory Meth 39:94–106
713. Parsi S, Ganjali M, Farsipour NS (2011) Conditional maximum likelihood and interval estimation for two Weibull populations under joint Type-II progressive censoring. Comm Stat Theory Meth 40:2117–2135
714. Patnaik PB (1949) The non-central χ^2- and F-distribution and their applications. Biometrika 36:202–232
715. Pellerey F, Shaked M, Zinn J (2000) Nonhomogeneous Poisson processes and logconcavity. Probab Eng Inform Sci 14:353–373
716. Peng XY, Yan ZZ (2013) Bayesian estimation for generalized exponential distribution based on progressive type-I interval censoring. Acta Math Appl Sin Engl Ser 29:391–402
717. Pérez-González C, Fernández AJ (2009) Accuracy of approximate progressively censored reliability sampling plans for exponential models. Stat Papers 50:161–170
718. Pfeifer D (1989) Einführung in die Extremwertstatistik. Teubner, Stuttgart (in german)
719. Pham-Gia T, Duong QP (1989) The generalized beta and F distributions in statistical modelling. Math Comput Modelling 12:1613–1625
720. Pickands J (1975) Statistical inference using extreme order statistics. Ann Stat 3:119–131
721. Pintilie M (2006) Competing risks: a practical perspective. Wiley, Hoboken
722. Pitman EJG (1937) The closest estimates of statistical parameters. Math Proc Cambridge Philos Soc 33:212–222
723. Pledger G, Proschan F (1971) Comparisons of order statistics and of spacings from heterogeneous distributions. In: Rustagi JS (ed) Optimizing Methods Statistics. Academic, New York, pp 89–113
724. Pollard D (1984) Convergence of stochastic processes. Springer, New York
725. Potdar KG, Shirke DT (2014) Inference for the scale parameter of lifetime distribution of k-unit parallel system based on progressively censored data. J Stat Comput Simul 84:171–185
726. Pradhan B (2007) Point and interval estimation for the lifetime distribution of a k-unit parallel system based on progressively Type-II censored data. Econom Qual Control 22:175–186
727. Pradhan B, Kundu D (2009) On progressively censored generalized exponential distribution. TEST 18:497–515
728. Pradhan B, Kundu D (2013) Inference and optimal censoring schemes for progressively censored Birnbaum-Saunders distribution. J Stat Plan Infer 143:1098–1108
729. Prékopa A (1973) On logarithmic concave measures and functions. Acta Sci Math 34: 335–343
730. Proschan F (1963) Theoretical explanation of observed decreasing failure rate. Technometrics 5:375–383

731. Proschan F, Sethuraman J (1976) Stochastic comparisons of order statistics from heterogeneous populations, with applications in reliability. J Multivariate Anal 6:608–616
732. Pudeg A (1991) Charakterisierungen von Wahrscheinlichkeitsverteilungen durch Verteilungseigenschaften der Ordnungsstatistiken und Rekorde. Ph.D. thesis, RWTH Aachen University, Aachen, Germany
733. Pukelsheim F (1993) Optimal design of experiments. Wiley, New York
734. Puri PS, Rubin H (1970) A characterization based on the absolute difference of two i.i.d. random variables. Ann Math Stat 41:2113–2122
735. Quesenberry C, Miller F (1977) Power studies of some tests for uniformity. J Stat Comput Simul 5:169–191
736. Rad AH, Yousefzadeh F, Balakrishnan N (2011) Goodness-of-fit test based on Kullback-Leibler information for progressively Type-II censored data. IEEE Trans Reliab 60:570–579
737. Ramberg JS, Tadikamalla PR (1978) On the generation of subsets of order statistics. J Stat Comput Simul 7:239–241
738. Rao CR (1949) On some problems arising out of discrimination with multiple characters. Sankhyā 9:343–366
739. Rao CR, Shanbhag DN (1986) Recent results on characterization of probability distributions: a unified approach through extensions of Deny's theorem. Adv Appl Probab 18:660–678
740. Rao CR, Shanbhag DN (1994) Choquet-Deny type functional equations with applications to stochastic models. Wiley, New York
741. Rao CR, Shanbhag DN (1998) Recent approaches to characterizations based on order statistics and record values. In: Balakrishnan N, Rao CR (eds) Order statistics: theory and methods. Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 231–256
742. Raqab MZ (2003) P-norm bounds for moments of progressive type II censored order statistics. Stat Probab Lett 64:393–402
743. Raqab MZ, Amin WA (1996) Some ordering results on order statistics and record values. IAPQR Trans 21:1–8
744. Raqab MZ, Madi MT (2002) Bayesian prediction of the total time on test using doubly censored Rayleigh data. J Stat Comput Simul 72:781–789
745. Raqab MZ, Madi MT (2011) Inference for the generalized Rayleigh distribution based on progressively censored data. J Stat Plan Infer 141:3313–3322
746. Raqab MZ, Asgharzadeh A, Valiollahi R (2010) Prediction for Pareto distribution based on progressively Type-II censored samples. Comput Stat Data Anal 54:1732–1743
747. Rasouli A, Balakrishnan N (2010) Exact likelihood inference for two exponential populations under joint progressive Type-II censoring. Comm Stat Theory Meth 39:2172–2191
748. Rastogi MK, Tripathi YM (2014) Parameter and reliability estimation for an exponentiated half-logistic distribution under progressive type II censoring. J Stat Comput Simul 84:1711–1727
749. Rastogi MK, Tripathi YM, Wu SJ (2012) Estimating the parameters of a bathtub-shaped distribution under progressive type-II censoring. J Appl Stat 39:2389–2411
750. Reiss RD (1989) Approximate distributions of order statistics. Springer, New York
751. Resnick SI (1987) Extreme values, regular variation, and point processes. Springer, New York
752. Rezapour M, Alamatsaz MH, Balakrishnan N (2013) On properties of dependent progressively Type-II censored order statistics. Metrika 76:909–917
753. Rezapour M, Alamatsaz MH, Balakrishnan N, Cramer E (2013) On properties of progressively Type-II censored order statistics arising from dependent and non-identical random variables. Stat Meth 10:58–71
754. Ringer LJ, Sprinkle EE (1972) Estimation of the parameters of the Weibull distribution from multicensored samples. IEEE Trans Reliab 21:46–51
755. Robertson T, Wright F, Dykstra R (1988) Order restricted statistical inference. Wiley, New York
756. Rogers GS (1959) A note on the stochastic independence of functions of order statistics. Ann Math Stat 30:1263–1264

757. Rogers GS (1963) An alternative proof of the characterization of the density Ax^B. Am Math Monthly 70:857–858
758. Rojo J, He GZ (1991) New properties and characterizations of the dispersive ordering. Stat Probab Lett 11:365–372
759. Ross SM (1996) Stochastic processes, 2nd edn. Wiley, New York
760. Rossberg HJ (1972) Characterization of distribution functions by the independence of certain functions of order statistics. Sankhyā Ser A 34:111–120
761. Rüschendorf L (1985) Two remarks on order statistics. J Stat Plan Infer 11:71–74
762. Ryan T (2011) Statistical methods for quality improvement, 3rd edn. Wiley, Hoboken
763. Rychlik T (1993) Sharp bounds for expectations of L-estimates from dependent samples. Comm Stat Theory Meth 22:1053–1068. Correction: Comm Stat Theory Meth 23:305–306
764. Rychlik T (2001) Projecting statistical functionals. Lecture notes in statistics, vol 160. Springer, New York
765. Rychlik T (2010) Evaluations of generalized order statistics from bounded populations. Stat Papers 51:165–177
766. Ryser HJ (1973) Combinatorial mathematics, 4th edn. The Carus mathematical monographs, vol 14. Mathematical Association of America, Washington
767. Saleh AKME, Ali MM (1966) Asymptotic optimum quantiles for the estimation of the parameters of the negative exponential distribution. Ann Math Stat 37:143–151
768. Sanjel D, Balakrishnan N (2008) A Laguerre polynomial approximation for a goodness-of-fit test for exponential distribution based on progressively censored data. J Stat Comput Simul 78:503–513
769. Saraçoğlu B, Kinaci I, Kundu D (2012) On estimation of $R = P(Y < X)$ for exponential distribution under progressive type-II censoring. J Stat Comput Simul 82:729–744
770. Saran J, Pande V (2012) Recurrence relations for moments of progressively Type-II right censored order statistics from half logistic distribution. J Stat Theory Appl 11:87–96
771. Saran J, Pushkarna N (2001) Recurrence relations for moments of progressive type-II right censored order statistics from Burr distribution. Statistics 35:495–507
772. Sarhan AE (1954) Estimation of the mean and standard deviation by order statistics. Ann Math Stat 25:317–328
773. Sarhan AE (1955) Estimation of the mean and standard deviation by order statistics, part III. Ann Math Stat 26:576–592
774. Sarhan AM, Al-Ruzaizaa A (2010) Statistical inference in connection with the Weibull model using Type-II progressively censored data with random scheme. Pakistan J Stat 26:267–279
775. Sarhan AE, Greenberg BG (1959) Estimation of location and scale parameters for the rectangular population from censored samples. J Roy Stat Soc Ser B 21:356–363
776. Sarhan AE, Greenberg BG (1962) Contributions to order statistics. Wiley, New York
777. Sarhan AE, Greenberg BG, Ogawa J (1963) Simplified estimates for the exponential distribution. Ann Math Stat 34:102–116
778. Sarhan AM, Hamilton D, Smith B (2012) Parameter estimation for a two-parameter bathtub-shaped lifetime distribution. Appl Math Model 36:5380–5392
779. Sarkar SK (1983) Some results on χ^2 and a related distribution. Sankhyā 45:253–255
780. Saunders D (1984) Dispersive ordering of distributions. Adv Appl Probab 16:693–694
781. Scheffé H, Tukey JW (1945) Non-parametric estimation. I. Validation of order statistics. Ann Math Stat 16:187–192
782. Schenk N (2001) Point estimation with sequential order statistics from exponential distributions. Ph.D. thesis, University of Oldenburg, Oldenburg, Germany
783. Schenk N, Burkschat M, Cramer E, Kamps U (2011) Bayesian estimation and prediction with multiply Type-II censored samples of sequential order statistics from one- and two-parameter exponential distributions. J Stat Plan Infer 141:1575–1587
784. Schneider H (1989) Failure-censored variables-sampling plans for lognormal and Weibull distributions. Technometrics 31:199–206

785. Schoenberg IJ (1946) Contributions to the problem of approximation of eqidistant data by analytic functions, part A: on the problem of smoothing or graduation, a first class of analytic approximation formulas. Quart Appl Math 4:45–99
786. Schoenberg IJ (1959) On variation diminishing approximation methods. In: Langer RE (ed) On numerical approximation. Proceedings of the symposium, Madison, 1958. University of Wisconsin Press, Madison, pp 249–274
787. Schucany WR (1972) Order statistics in simulation. J Stat Comput Simul 1:281–286
788. Sedyakin NM (1966) On one physical principle in reliability theory (in Russian). Tech Cybern 3:80–87
789. Sen PK (1959) On the moments of the sample quantiles. Calcutta Stat Assoc Bull 9:1–19
790. Sen PK (1970) A note on order statistics for heterogeneous distributions. Ann Math Stat 41:2137–2139
791. Sen PK, Singer JM (1993) Large sample methods in statistics. Chapman & Hall, New York
792. Sen A, Kannan N, Kundu D (2013) Bayesian planning and inference of a progressively censored sample from linear hazard rate distribution. Comput Stat Data Anal 62:108–121
793. Serfling RJ (1980) Approximation theorems of mathematical statistics. Wiley, New York
794. Shafay AR, Balakrishnan N, Sultan KS (2014) Two-sample Bayesian prediction for sequential order statistics from exponential distribution based on multiply Type-II censored samples. J Stat Comput Simul 84:526–544
795. Shah BK (1966) On the bivariate moments of order statistics from a logistic distribution. Ann Math Stat 37:1002–1010
796. Shah BK (1970) Note on moments of a logistic order statistics. Ann Math Stat 41: 2150–2152
797. Shaked M (1982) Dispersive ordering of distributions. J Appl Probab 19:310–320
798. Shaked M (1985) Ordering distributions by dispersion. In: Kotz S, Johnson NL (eds) Encyclopedia of statistical sciences, vol 6. Wiley, New York
799. Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, Berlin
800. Shaked M, Singpurwalla ND (1983) Inference for step-stress accelerated life tests. J Stat Plan Infer 7:295–306
801. Shapiro SS, Wilk MB (1972) An analysis of variance test for the exponential distribution (complete samples). Technometrics 14:355–370
802. Sharafi M, Balakrishnan N, Khaledi BE (2013) Distribution-free comparison of hazard rates of two distributions under progressive type-II censoring. J Stat Comput Simul 83: 1527–1542
803. Shen KF, Shen YJ, Leu LY (2011) Design of optimal step-stress accelerated life tests under progressive type I censoring with random removals. Qual Quant 45:587–597
804. Singh S, Tripathi YM, Wu SJ (2014) On estimating parameters of a progressively censored lognormal distribution. J Stat Comput Simul (to appear)
805. Sinha AN, Sen PK (1982) Tests based on empirical processes for progressive censoring schemes with staggering entry and random withdrawal. Sankhyā, Ser B 44:1–18
806. Sinha AN, Sen PK (1983) Staggering entry, random withdrawal and progressive censoring schemes: Some nonparametric procedures. In: Ghosh JK, Roy J (eds) Statistics applications and new directions: proceedings of the Indian statistical institute golden jubilee conference. Indian Statistical Institute, Calcutta, pp 531–547
807. Smirnov NV (1952) Limit distributions for the terms of a variational series. Am Math Soc Transl Ser 1:82–143
808. Smirnov NV (1967) Some remarks on limit laws for order statistics. Theory Probab Appl 12:337–339
809. Smith RL (1985) Maximum likelihood estimation in a class of nonregular cases. Biometrika 72:67–90
810. Soland RM (1969) Bayesian analysis of the Weibull process with unknown scale and shape parameters. IEEE Trans Reliab R-18:181–184

811. Soliman AA (2005) Estimation of parameters of life from progressively censored data using Burr-XII model. IEEE Trans Reliab 54:34–42
812. Soliman AA (2006) Estimators for the finite mixture of Rayleigh model based on progressively censored data. Comm Stat Theory Meth 35:803–820
813. Soliman AA (2008) Estimations for Pareto model using general progressive censored data and asymmetric loss. Comm Stat Theory Meth 37:1353–1370
814. Soliman AA, Abd Ellah AH, Abou-Elheggag NA, Abd-Elmougod GA (2011) A simulation-based approach to the study of coefficient of variation of Gompertz distribution under progressive first-failure censoring. Indian J Pure Appl Math 42:335–356
815. Soliman AA, Ellah A, Abou-Elheggag NA, Modhesh AA (2011) Bayesian inference and prediction of Burr Type XII distribution for progressive first failure censored sampling. Intell Inform Manag 3:175–185
816. Soliman AA, Al-Hossain AY, Al-Harbi MM (2011) Predicting observables from Weibull model based on general progressive censored data with asymmetric loss. Stat Meth 8: 451–461
817. Soliman AA, Abd-Ellah AH, Abou-Elheggag NA, Abd-Elmougod GA (2012) Estimation of the parameters of life for Gompertz distribution using progressive first-failure censored data. Comput Stat Data Anal 56:2471–2485
818. Soliman AA, Abd-Ellah AH, Abou-Elheggag NA, Ahmed EA (2012) Modified Weibull model: a Bayes study using MCMC approach based on progressive censoring data. Reliab Eng Syst Safety 100:48–57
819. Soliman AA, Abd Ellah AH, Abou-Elheggag NA, Modhesh AA (2013) Estimation from Burr type XII distribution using progressive first-failure censored data. J Stat Comput Simul 83:2270–2290
820. Soofi ES (2000) Principal information theoretic approaches. J Am Stat Assoc 95:1349–1353
821. Spinelli JJ, Stephens MA (1987) Tests for exponentiality when origin and scale parameters are unknown. Technometrics 29:471–476
822. Srikantan KS (1962) Recurrence relations between the pdf's of order statistics, and some applications. Ann Math Stat 33:169–177
823. Stacy EW (1962) A generalization of the gamma distribution. Ann Math Stat 33:1187–1192
824. Stephens MA (1969) Results from the relation between two statistics of the Kolmogorov-Smirnov type. Ann Math Stat 40:1833–1837
825. Stigler SM (1977) Fractional order statistics, with applications. J Am Stat Assoc 72: 544–550
826. Sukhatme PV (1937) Test of significance for samples of the χ^2-population with two degrees of freedom. Ann Eugenics 8:52–56
827. Sultan KS, Alsadat NH, Kundu D (2014) Bayesian and maximum likelihood estimations of the inverse Weibull parameters under progressive type-II censoring. J Stat Comput Simul 84:2248–2265
828. Sun T, Shi Y, Dang S (2012) Inference for Burr-XII with progressive Type-II censoring with random removals in step-stress partially accelerated life test. In: 2012 international conference on quality, reliability, risk, maintenance, and safety engineering (ICQR2MSE) , pp 889–894
829. Szekli R (1995) Stochastic ordering and dependence in applied probability. Lecture notes in statistics, vol 97. Springer, New York
830. Tadikamalla PR (1980) A look at the Burr and related distributions. Int Stat Rev 48:337–344
831. Tadikamalla PR, Balakrishnan N (1998) Computer simulation of order statistics. In: Balakrishnan N, Rao CR (eds) Order statistics: theory and methods. Handbook of statistics, vol 16. Elsevier, Amsterdam, pp 65–74
832. Takada Y (1981) Relation of the best invariant predictor and the best unbiased predictor in location and scale families. Ann Stat 9:917–921
833. Takada Y (1982) A comment on best invariant predictors. Ann Stat 10:971–978
834. Takada Y (1991) Median unbiasedness in an invariant prediction problem. Stat Probab Lett 12:281–283

835. Takahasi K (1988) A note on hazard rates of order statistics. Comm Stat Theory Meth 17:4133–4136
836. Tan Z (2009) A new approach to MLE of Weibull distribution with interval data. Reliab Eng Syst Safety 94:394–403
837. Tang LC (2003) Multiple steps step-stress accelerated. In: Pham H (ed) Handbook of reliability engineering. Springer, New York, pp 441–455
838. Tanner MA (1993) Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions, 2nd edn. Springer, New York
839. Tavangar M, Asadi M (2011) On stochastic and aging properties of generalized order statistics. Probab Eng Inform Sci 25:187–204
840. Tavangar M, Hashemi M (2013) On characterizations of the generalized Pareto distributions based on progressively censored order statistics. Stat Papers 54:381–390
841. Teimouri M, Gupta AK (2012) Estimation methods for the Gompertz-Makeham distribution under progressively Type-I interval censoring scheme. Natl Acad Sci Lett 35:227–235
842. Teimouri M, Nadarajah S, Hsing Shih S (2014) EM algorithms for beta kernel distributions. J Stat Comput Simul 84:451–467
843. Thomas DR, Wilson WM (1972) Linear order statistic estimation for the two-parameter Weibull and extreme value distribution from Type-II progressively censored samples. Technometrics 14:679–691
844. Thompson WR (1936) On confidence ranges for the median and other expectation distributions for populations of unknown distribution form. Ann Math Stat 7:122–128
845. Tierney L, Kadane JB (1986) Accurate approximations for posterior moments and marginal densities. J Am Stat Assoc 81:82–86
846. Tietjen GL, Kahaner DK, Beckman RJ (1977) Variances and covariances of the normal order statistics for sample sizes 2 to 50. In: Selected tables in mathematical statistics, vol 5. American Mathematical Society, Providence, pp 1–73
847. Tiku M (1980) Goodness of fit statistics based on the spacings of complete or censored samples. Aust J Stat 22:260–275
848. Tomer SK, Panwar MS (2014) Estimation procedures for Maxwell distribution under type-I progressive hybrid censoring scheme. J Stat Comput Simul (to appear)
849. Tong H (1974) A note on the estimation of $Pr(Y < X)$ in the exponential case. Technometrics 16:625. Correction (1975) Technometrics 17:395
850. Tong H (1975) Errata: a note on the estimation of $Pr\{y < x\}$ in the exponential case. Technometrics 17:395
851. Tong H (1975) Letter to the editor. Technometrics 17:393
852. Tong LI, Chen K, Chen H (2002) Statistical testing for assessing the performance of lifetime index of electronic components with exponential distribution. Int J Qual Reliab Manag 19:812–824
853. Torrado N, Lillo R, Wiper M (2012) Sequential order statistics: ageing and stochastic orderings. Method Comput Appl Probab 14:579–596
854. Tran TTH (2006) Discrete generalized order statistics. Ph.D. thesis, University of Oldenburg, Oldenburg, Germany
855. Tsai TR, Lin CW (2010) Acceptance sampling plans under progressive interval censoring with likelihood ratio. Stat Papers 51:259–271
856. Tsallis C (1988) Possible generalization of Boltzmann-Gibbs statistics. J Stat Phys 52: 479–487
857. Tse SK, Xiang L (2003) Interval estimation for Weibull-distributed life data under Type II progressive censoring with random removals. J Biopharm Stat 13:1–16
858. Tse SK, Yang C (2003) Reliability sampling plans for the Weibull distribution under Type II progressive censoring with binomial removals. J Appl Stat 30:709–718
859. Tse SK, Yuen HK (1998) Expected experiment times for the Weibull distribution under progressive censoring with random removals. J Appl Stat 25:75–83

860. Tse SK, Yang C, Yuen HK (2000) Statistical analysis of Weibull distributed lifetime data under Type II progressive censoring with binomial removals. J Appl Stat 27:1033–1043
861. Tse SK, Ding C, Yang C (2008) Optimal accelerated life tests under interval censoring with random removals: the case of Weibull failure distribution. Statistics 42:435–451
862. Tsiatis A (1975) A nonidentifiability aspect of the problem of competing risks. Proc Natl Acad Sci USA 72:20–22
863. Tukey JW (1965) Which part of the sample contains the information? Proc Natl Acad Sci USA 53:127–134
864. Vajda I (1989) Theory of statistical inference and information. Kluwer Academic Publishers, Dordrecht
865. Valiollahi R, Asgharzadeh A, Raqab MZ (2013) Estimation of $P(Y < X)$ for Weibull distribution under progressive Type-II censoring. Comm Stat Theory Meth 42:4476–4498
866. Van Dorp JR, Mazzuchi TA (2004) A general Bayes exponential inference model for accelerated life testing. J Stat Plan Infer 119:55–74
867. Vännman K (1976) Estimators based on order statistics from a Pareto distribution. J Am Stat Assoc 71:704–708
868. Vännman K (1998) Families of capability indices for one-sided specification limits. Statistics 31:43–66
869. Varde SD (1969) Life testing and reliability estimation for the two parameter exponential distribution. J Am Stat Assoc 64:621–631
870. Varian HR (1974) A Bayesian approach to real estate assessment. In: Fienberg S, Zellner A (eds) Studies in Bayesian econometrics and statistics. North-Holland, Amsterdam, pp 195–208
871. Varian HR (2000) Variants in economic theory: selected works of Hal R. Varian. Edward Elgar Publ Ltd., Northampton
872. Vasicek O (1976) A test for normality based on sample entropy. J Roy Stat Soc Ser B 38: 54–59
873. Vasudeva Rao A, Dattatreya Rao A, Narasimham V (1994) Asymptotically optimal grouping for maximum likelihood estimation of Weibull parameters. Comm Stat Simul Comput 23:1077–1096
874. Vaughan RJ, Venables WN (1972) Permanent expressions for order statistic densities. J Roy Stat Soc B 34:308–310
875. Viveros R, Balakrishnan N (1994) Interval estimation of parameters of life from progressively censored data. Technometrics 36:84–91
876. Volterman W, Balakrishnan N (2010) Exact nonparametric confidence, prediction and tolerance intervals based on multi-sample Type-II right censored data. J Stat Plan Infer 140:3306–3316
877. Volterman W, Davies KF, Balakrishnan N (2012) Pitman closeness as a criterion for the determination of the optimal progressive censoring scheme. Stat Meth 9:563–572
878. Volterman W, Balakrishnan N, Cramer E (2014) Exact meta-analysis of several independent progressively Type-II censored data. Appl Math Model 38:949–960
879. Volterman W, Davies KF, Balakrishnan N (2013) Simultaneous Pitman closeness of progressively Type-II right-censored order statistics to population quantiles. Statistics 47:439–452
880. Volterman W, Davies KF, Balakrishnan N (2013) Two-sample Pitman closeness comparison under progressive Type-II censoring. Statistics 47:1305–1320
881. Vuong QN, Cramer E (2010) Computation of optimal progressive Type-II censoring plans in large samples via genetic algorithms. Technical report. RWTH Aachen University, Aachen
882. Wang BX (2008) Goodness-of-fit test for the exponential distribution based on progressively Type-II censored sample. J Stat Comput Simul 78:125–132
883. Wang BX (2009) Interval estimation for the scaled half-logistic distribution under progressive Type-II censoring. Comm Stat Theory Meth 38:364–371
884. Wang BX (2010) Interval estimation for exponential progressive Type-II censored step-stress accelerated life-testing. J Stat Plan Infer 140:2706–2718

885. Wang BX (2012) Exact interval estimation for the scale family under general progressive Type-II censoring. Comm Stat Theory Meth 41:4444–4452
886. Wang FK, Cheng YF (2010) EM algorithm for estimating the Burr XII parameters with multiple censored data. Qual Reliab Eng Int 26:615–630
887. Wang Y, He S (2005) Fisher information in censored data. Stat Probab Lett 73:199–206
888. Wang L, Shi Y (2012) Reliability analysis based on progressively first-failure-censored samples for the proportional hazard rate model. Math Comput Simul 82:1383–1395
889. Wang BX, Yu K (2009) Optimum plan for step-stress model with progressive Type-II censoring. TEST 18:115–135
890. Wang BX, Yu K, Jones M (2010) Inference under progressively Type II right censored sampling for certain lifetime distributions. Technometrics 52:453–460
891. Watkins AJ (2001) Commentary: inference in simple step-stress models. IEEE Trans Reliab 50:36–37
892. Watson GS, Leadbetter MR (1964) Hazard analysis. I. Biometrika 51:175–184
893. Watson GS, Leadbetter MR (1964) Hazard analysis II. Sankhyā Ser A 26:101–116
894. Weerahandi S (1993) Generalized confidence intervals. J Am Stat Assoc 88:899–905
895. Weerahandi S (2004) Generalized inference in repeated measures. Exact methods in MANOVA and mixed models. Wiley, Hoboken
896. Weinman DG (1966) A multivariate extension of the exponential distribution. Ph.D. thesis, Arizona State University
897. Wilson EB, Hilferty MM (1931) The distribution of chi-square. Proc Natl Acad Sci 17: 684–688
898. Wingo DR (1973) Solution of the three-parameter Weibull equations by constrained modified quasilinearization (progressively censored samples). IEEE Trans Reliab R-22: 96–102
899. Wingo DR (1983) Maximum likelihood methods for fitting the Burr type XII distribution to life test data. Biom J 25:77–84
900. Wingo DR (1993) Maximum likelihood estimation of Burr XII distribution parameters under Type II censoring. Microelectron Reliab 33:1251–1257
901. Wingo DR (1993) Maximum likelihood methods for fitting the Burr type XII distribution to multiply progressively censored life test data. Metrika 40:203–210
902. Wong JY (1993) Simultaneously estimating the three Weibull parameters from progressively censored samples. Microelectron Reliab 3:2217–2224
903. Wu CY (1966) The types of limit distributions for some terms of variational series. Sci Sinica 15:749–762
904. Wu SJ (2002) Estimations of the parameters of the Weibull distribution with progressively censored data. J Jpn Stat Soc 32:155–163
905. Wu SJ (2003) Estimation for the two-parameter Pareto distribution under progressive censoring with uniform removals. J Stat Comput Simul 73:125–134
906. Wu SJ (2008) Estimation of the two-parameter bathtub-shaped lifetime distribution with progressive censoring. J Appl Stat 35:1139–1150
907. Wu SF (2008) Interval estimation for a Pareto distribution based on a doubly Type II censored sample. Comput Stat Data Anal 52:3779–3788
908. Wu SF (2010) Interval estimation for the Pareto distribution based on the progressive Type II censored sample. J Stat Comput Simul 80:463–474
909. Wu SF (2010) Interval estimation for the two-parameter exponential distribution under progressive censoring. Qual Quant 44:181–189
910. Wu SJ, Chang CT (2002) Parameter estimations based on exponential progressive type II censored data with binomial removals. Int J Inform Manag Sci 13:37–46
911. Wu SJ, Huang SR (2010) Optimal progressive group-censoring plans for exponential distribution in presence of cost constraint. Stat Papers 51:431–443
912. Wu SJ, Huang SR (2012) Progressively first-failure censored reliability sampling plans with cost constraint. Comput Stat Data Anal 56:2018–2030

913. Wu SJ, Kuş C (2009) On estimation based on progressive first-failure-censored sampling. Comput Stat Data Anal 53:3659–3670
914. Wu SJ, Chang CT, Tsai TR (2003) Point and interval estimations for the Gompertz distribution under progressive type-II censoring. Metron 61(3):403–418
915. Wu SJ, Chen DH, Chen ST (2006) Bayesian inference for Rayleigh distribution under progressive censored sample. Appl Stoch Models Bus Ind 22:269–279
916. Wu CC, Wu SF, Chan HY (2006) MLE and the estimated expected test time for the two-parameter Gompertz distribution under progressive censoring with binomial removals. Appl Math Comput 181:1657–1670
917. Wu SJ, Lin YP, Chen YJ (2006) Planning step-stress life test with progressively type I group-censored exponential data. Stat Neerl 60:46–56
918. Wu JW, Wu SF, Yu CM (2007) One-sample Bayesian predictive interval of future ordered observations for the Pareto distribution. Qual Quant 41:251–263
919. Wu SJ, Chen YJ, Chang CT (2007) Statistical inference based on progressively censored samples with random removals from the Burr type XII distribution. J Stat Comput Simul 77:19–27
920. Wu SJ, Lin YP, Chen ST (2008) Optimal step-stress test under type I progressive group-censoring with random removals. J Stat Plan Infer 138:817–826
921. Wu SJ, Chang CT, Liao KJ, Huang SR (2008) Planning of progressive group-censoring life tests with cost considerations. J Appl Stat 35:1293–1304
922. Wu SF, Wu CC, Chen YL, Yu YR, Lin YP (2010) Interval estimation of a two-parameter Burr-XII distribution under progressive censoring. Statistics 44:77–88
923. Wu SF, Wu CC, Chou CH, Lin HM (2011) Statistical inferences of a two-parameter distribution with the bathtub shape based on progressive censored sample. J Stat Comput Simul 81:315–329
924. Xiang L, Tse SK (2005) Maximum likelihood estimation in survival studies under progressive interval censoring with random removals. J Biopharm Stat 15:981–991
925. Xie H, Hu T (2010) Some new results on multivariate dispersive ordering of generalized order statistics. J Multivariate Anal 101:964–970
926. Xie H, Zhuang W (2011) Some new results on ordering of simple spacings of generalized order statistics. Probab Eng Inform Sci 25:71–81
927. Xie Q, Balakrishnan N, Han D (2008) Exact inference and optimal censoring scheme for a simple step-stress model under progressive Type-II censoring. In: Balakrishnan N (ed) Advances in mathematical and statistical modeling. Birkhäuser, Boston, pp 107–137
928. Xiong C (1998) Inferences on a simple step-stress model with type-II censored exponential data. IEEE Trans Reliab 47:142–146
929. Xiong C, Ji M (2004) Analysis of grouped and censored data from step-stress life test. IEEE Trans Reliab 53:22–28
930. Xiong C, Milliken G (1999) Step-stress life-testing with random stress-change times for exponential data. IEEE Trans Reliab 48:141–148
931. Xiuyun P, Zaizai Y (2011) Parameter estimations with gamma distribution based on progressive Type-I interval censoring. In: IEEE international conference on computer science and automation engineering (CSAE), 2011, pp 449–453
932. Yang SS (1977) General distribution theory of the concomitants of order statistics. Ann Stat 5:996–1002
933. Yang C, Tse SK (2005) Planning accelerated life tests under progressive Type I interval censoring with random removals. Comm Stat Simul Comput 34:1001–1025
934. Yeh L, Balakrishnan N (1995) Tolerance limits and acceptance sampling plans. In: Balakrishnan N, Basu AP (eds) The exponential distribution: theory, methods, and applications. Gordon and Breach, Newark, pp 119–138
935. Yue HB, Shi YM (2013) Optimal sample size allocation for multi-level stress testing under progressive hybrid interval censoring. Appl Mech Mater 423–426:2423–2426

936. Yuen HK, Tse SK (1996) Parameters estimation for Weibull distributed lifetimes under progressive censoring with random removals. J Stat Comput Simul 55:57–71
937. Zacks S (1971) The theory of statistical inference. Wiley, New York
938. Zellner A (1986) Bayesian estimation and prediction using asymmetric loss functions. J Am Stat Assoc 81:446–451
939. Zhang J (2013) Simplification of joint confidence regions for the parameters of Pareto distribution. Comput Stat 28:1453–1462
940. Zhang Y, Meeker WQ (2005) Bayesian life test planning for the Weibull distribution with given shape parameter. Metrika 61:237–249
941. Zheng G (2001) A characterization of the factorization of hazard function by the Fisher information under Type-II censoring with application to the Weibull family. Stat Probab Lett 52:249–253
942. Zheng G, Gastwirth JL (2000) Where is the Fisher information in an ordered sample? Stat Sinica 10:1267–1280
943. Zheng G, Park S (2004) On the Fisher information in multiply censored and progressively censored data. Comm Stat Theory Meth 33:1821–1835
944. Zheng G, Park S (2005) Another look at life testing. J Stat Plan Infer 127:103–117
945. Zheng G, Balakrishnan N, Park S (2009) Fisher information in ordered data: a review. Stat Interface 2:101–113
946. Zhuang W, Hu T (2009) Multivariate dispersive ordering of spacings of generalized order statistics. Appl Math Lett 22:968–974
947. Zieliński R, Zieliński W (2005) Best exact nonparametric confidence intervals for quantiles. Statistics 39:67–71
948. Zijlstra M (1983) Characterizations of the geometric distribution by distributional properties. J Appl Probab 20:843–850

Author Index

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7,

C

D

E

Index

N. Balakrishnan and E. Cramer, *The Art of Progressive Censoring*, Statistics for Industry and Technology, DOI 10.1007/978-0-8176-4807-7,

Q

R

S

T

U

V

W