Lecture Notes in Statistics 208
Proceedings

Haijun Li
Xiaohu Li
Editors

Stochastic Orders in Reliability and Risk

In Honor of Professor Moshe Shaked

Springer

Lecture Notes in Statistics 208

Edited by
P. Bickel, P. Diggle, S. Fienberg, U. Gather, I. Olkin, S. Zeger

For further volumes:
http://www.springer.com/series/694

Haijun Li • Xiaohu Li
Editors

Stochastic Orders in Reliability and Risk

In Honor of Professor Moshe Shaked

Springer

Editors
Haijun Li
Department of Mathematics
Washington State University
Pullman, Washington, USA

Xiaohu Li
School of Mathematical Sciences
Xiamen University
Xiamen, China, People's Republic

ISSN 1869-7240
ISBN 978-1-4614-6891-2 ISBN 978-1-4614-6892-9 (eBook)
DOI 10.1007/978-1-4614-6892-9
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013939034

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Introduction

In summer of 2010, the first author (HL) visited the second author (XL) at Lanzhou University, China, and chaired the dissertation defense for XL's two graduating doctoral students. During the visit, we discussed that a large reliability meeting (MMR2011) was scheduled to be held in Beijing in the summer of 2011 and that the meeting would attract some stochastic inequality people, including Professor Moshe Shaked, to visit Beijing. XL then initiated the idea of organizing a small academic gathering for these people at Xiamen University, China, focusing specifically on stochastic inequalities in honor of Moshe Shaked—our common academic mentor, our coauthor, and our good friend. A stochastic order workshop was immediately planned to promote close collaboration in honor of Moshe. The people whom we have contacted with were overwhelmingly enthusiastic about the idea. Some people couldn't come but sent us their suggestions about the workshop. The funding for this workshop was provided by XL's NNSF research funds with support from the School of Mathematical Sciences and Center for Actuarial Studies at Xiamen University.

Xiamen is situated on the southeast coast of China, to the west of Taiwan Strait. Known as a "Garden on the Sea," Xiamen is surrounded by ocean on three sides. The International Workshop on Stochastic Orders in Reliability and Risk Management, or SORR2011, was held in Xiamen City Hotel from June 27 to June 29, 2011. SORR2011 featured 11 invited speeches and nine contributed talks, covering a wide range of topics from theory of stochastic orders to applications in reliability and risk/ruin analysis. Professor Moshe Shaked delivered the opening keynote speech. A social highlight of SORR2011 was a surprise banquet party for Professor Moshe Shaked and Ms Edith Shaked.

This volume is based on the talks presented at the workshop and the invited contributions to this special occasion to honor Professor Moshe Shaked, who has made fundamental and widespread contributions to theory of stochastic orders and its applications in reliability, queueing modeling, operations research, economics, and risk analysis. All the papers submitted were subjected to reviewing, and all the accepted papers have been edited to standardize notations and terminologies. The volume consists of 19 contributions that are organized along the following five categories:

Part I: Theory of Stochastic Orders

- "A Global Dependence Stochastic Order Based on the Presence of Noise" by Moshe Shaked, Miguel A. Sordo, and Alfonso Suárez-Llorens
- "Duality Theory and Transfers for Stochastic Order Relations" by Alfred Müller
- "Reversing Conditional Orderings" by Rachele Foschi and Fabio Spizzichino

Part II: Stochastic Comparison of Order Statistics

- "Multivariate Comparisons of Ordered Data" by Félix Belzunce
- "On Stochastic Properties of Spacings with Applications in Multiple-Outlier Models" by Nuria Torrado and Rosa E. Lillo
- "On Sample Range from Two Heterogeneous Exponential Variables" by Peng Zhao and Xiaohu Li

Part III: Stochastic Orders in Reliability

- "On Bivariate Signatures for Systems with Independent Modules" by Gaofeng Da and Taizhong Hu
- "Stochastic Comparisons of Cumulative Entropies" by Antonio Di Crescenzo and Maria Longobardi
- "Decreasing Percentile Residual Life Aging Notion: Properties and Estimation" by Alba M. Franco-Pereira, Jacobo de Uña, Rosa E. Lillo, and Moshe Shaked

- “A Review on Convolutions of Gamma Random Variables” by Baha-Eldin Khaledi and Subhash Kochar
- “Allocation of Active Redundancies to Coherent Systems: A Brief Review” by Xiaohu Li and Weiyong Ding
- “On Used Systems and Systems with Used Components” by Xiaohu Li, Franco Pellerey, and Yinping You

Part IV: Stochastic Orders in Risk Analysis

- “Dynamic Risk Measures Within Discrete-Time Risk Models” by Hélène Cossette and Etienne Marceau
- “Excess Wealth Transform with Applications” by Subhash Kochar and Maochao Xu

Part V: Applications

- “Intermediate Tail Dependence: A Review and Some New Results” by Lei Hua and Harry Joe
- “Second-Order Conditions of Regular Variation and Drees Type Inequalities” by Tiantian Mao
- “Individual and Moving Ratio Charts for Weibull Processes” by Francis Pascual
- “On a Slow Server Problem” by Vladimir Rykov
- “Dependence Comparison of Multivariate Extremes via Stochastic Tail Orders” by Haijun Li

We thank all the authors and workshop participants for their contributions. This volume is dedicated to Professor Moshe Shaked to celebrate his academic achievements and also intended to stimulate further research on stochastic orders and their applications.

Professor Moshe Shaked

Moshe Shaked has been for the past 31 years a professor of mathematics at the University of Arizona, Tucson, AZ. He received his B.A. and M.A. degrees from Hebrew University of Jerusalem in 1967 and 1971, respectively. Moshe pursued his graduate studies in mathematics and statistics under Albert W. Marshall at the University of Rochester from 1971 to 1975. Moshe received his Ph.D. in 1975 and his dissertation was entitled “On Concepts of Positive Dependence.”

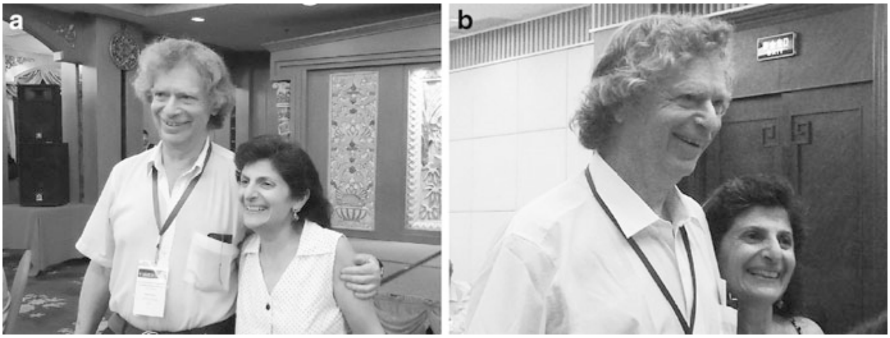


Figure 1: Moshe Shaked and Edith Shaked. (**a**) Beijing, China, June 2011. (**b**) Xiamen, China, June 2011

After short stays at the University of New Mexico, University of British Columbia, and Indiana University, Moshe became an associate professor of mathematics at the University of Arizona in 1981. Since 1986, he has been a full professor at Arizona (Fig. 1).

Moshe has made fundamental contributions in various areas of probability, statistics, and operations research. He has published over 180 papers and many of his papers appeared in the top journals in probability, statistics, and operations research. Coauthored with George Shanthikumar, Moshe published one of the two popular books on stochastic orders [426] (the other book was written by Alfred Müller and Dietrich Stoyan [335]). Moshe's contribution is extremely broad; for example, Moshe made seminal contributions to the following areas:

- Dependence analysis, positive and negative dependence notions, dependence by mixture of distributions, distributions with fixed marginals, and global dependence
- Comparison of stochastic processes, aging properties of stochastic processes, and aging first passage times
- Stochastic variability orders, dispersive ordering of distributions, and excess wealth order
- Accelerated life tests—inference, nonparametric approach, and goodness of fit
- Multivariate phase-type distributions

- Multivariate aging notions and multivariate life distributions
- Multivariate conditional hazard rate functions
- Linkages as a tool for construction of multivariate distributions
- Inventory centralization costs and games
- Stochastic convexity and concavity and stochastic majorization
- Stochastic comparisons of order statistics
- Total time on test transform order
- Use of antithetic variables in simulation
- Scientific activity and truth acquisition in social epistemology

In recognition of his many contributions, Moshe Shaked was elected as a Fellow of the Institute of Mathematical Statistics in 1986. He has been serving in editorial boards of various probability, statistics, and operations research journals and book series.

Moshe enjoys collaborations and has been working with more than 60 collaborators worldwide. Moshe is a stimulating, accommodating, and generous collaborator with colleagues and students alike. Moshe and Edith travel a lot professionally, so the concepts of "vacation" and "conference" often have the same meaning for them. Changing a routine in Tucson, visiting different places in other parts of the world, and meeting new friends (potential collaborators?) are all both relaxing and rewarding for Moshe and Edith. In coffee breaks of several conferences, we have witnessed that Moshe still worked on problems with collaborators one by one. It seems to us that Moshe values collaborating itself as much as he values possible products (i.e., papers) resulting from collaboration. This reminds us of Paul Erdős, a great mathematician, who strongly believed in scientific collaboration and practiced mathematics research as a social activity.

On the personal side, it was Moshe who helped HL get his academic job in the USA and it was Moshe who mentored XL in launching his academic career. Collaborating with Moshe has been a real treat for both of us, and by working with Moshe, we learned and became greatly appreciative to the true value of professionalism.

Stochastic Orders: A Historical Perspective

Stochastic ordering refers to comparing random elements in some stochastic sense and has evolved into a deep field of enormous breadth with ample structures of its own, establishing strong ties with numerous striking applications in economics, finance, insurance, management science, operations research, statistics, and other fields in engineering, natural, and social sciences. Stochastic ordering is a fundamental guide for decision making under uncertainty and an essential tool in the study of structural properties of complex stochastic systems.

Take two random variables X and Y, for example. One way to compare them is to compare their survival functions; that is, if

$$\mathrm{P}\{X > t\} \le \mathrm{P}\{Y > t\}, \text{ for all real } t, \tag{1}$$

then Y is more likely to "survive" beyond t than X does, and we say X is stochastically smaller than Y and denote this by $X \le_{\mathrm{st}} Y$. Using approximations, the path-wise ordering Eq. (1) can be showed to be equivalent to

$$\mathrm{E}[\phi(X)] \le \mathrm{E}[\phi(Y)], \text{ for all nondecreasing functions } \phi : \mathbb{R} \to \mathbb{R}, \tag{2}$$

provided that the expectations exist. That is, $X \le_{\mathrm{st}} Y$ is equivalent to the comparisons with respect to a class of increasing functionals of random variables. If a system performance measure can be written as an increasing functional $\mathrm{E}[\phi(\mathrm{X})]$, where $\phi(\cdot)$ is increasing, then the system performance comparison boils down to the stochastic order Eq. (1).

The stochastic order $\le_{\mathrm{st}}$ enjoys nice operational properties (see [335, 426]), and its utility can be greatly enhanced via coupling [444]. For any two random variables X and Y, $X \le_{\mathrm{st}} Y$ if and only if there exist two random variables $\hat{X}$ and $\widehat{Y}$, defined on the same probability space $(\Omega, \mathcal{F}, \mathrm{P})$, such that $\widehat{X}$ and X have the same (marginal) distribution, $\hat{Y}$ and Y have the same (marginal) distribution, and

$$\mathrm{P}\{\hat{X} \le \hat{Y}\} = 1. \tag{3}$$

That is, one can work with almost-sure inequalities on the coupling space $(\Omega, \mathcal{F}, \mathrm{P})$ and move back to the original random variables using marginal distributional equivalence.

The stochastic order $\le_{\mathrm{st}}$ is also mathematically robust; namely, the order $\le_{\mathrm{st}}$, as described in Eqs. (1)–(3), can be extended to probability measures defined on a partially ordered Polish space [220] (i.e., a complete separable metric space endowed with a closed partial ordering). For example, the stochastic order $\leq_{\rm st}$ on $\mathbb{R}^\infty$ can be applied to comparing two discrete-time stochastic processes. The stochastic order $\leq_{\rm st}$ is also extended to nonadditive measures [138]. The models that involve nonadditive probability measures have been used in decision theory to cope with observed violations of expected utility [412] (e.g., the Keynes–Ellsberg paradox). These models describe such distortions using different transforms of usual probabilities and have been applied to insurance premium pricing [116, 465, 466].

The stochastic order $\leq_{\rm st}$ is just one example that illustrates the deep stochastic comparison theory with widespread applications [335, 426]. The stochastic order $\leq_{\rm st}$, however, is one of *strong orderings*, and many stochastic systems can only be compared using *weak orders*. One example of weak integral stochastic orders is the increasing and convex order $\leq_{\rm icx}$ that uses the set of all increasing and convex functions in Eq. (2). The idea of seeking various weaker versions of a problem solution has been used throughout mathematics (e.g., in the theory of partial differential equations), and indeed various weak stochastic orders and their applications add enormous breadth to the field of stochastic orders.

The studies on stochastic orders have a long and colorful history. To the best of our knowledge, the studies on inequalities of type (2) for convex functions $\phi(\cdot)$ can be traced back to Karamata [223]. Known as the dilation order, the comparison Eq. (2) for all continuous convex functions $\phi(\cdot)$ is closely related to the notion of majorization. The theory of stochastic inequalities based on majorization is summarized in Marshall and Olkin [308] and its updated version [312].

Historically, stochastic orders have been used to define and study multivariate dependence. Some strongest dependence notions can be defined in terms of *total positivity* [224]. Earlier studies have been focused on dependence structures of multivariate normal distributions and multivariate distributions of elliptical type (see Tong [449]). For analyzing dependence structures of non-normal multivariate distributions, stochastic orders have been substantially used in Joe [211] and Nelsen [355], in which dependence structures of copulas, especially extreme value copulas, have been systematically investigated using orthant and supermodular orders.

Stochastic orders have been applied to various domain fields and especially to reliability theory. Both of us first learned stochastic orders from the 1975 seminal book on reliability and life testing by Barlow and Proschan [38], where Erich L. Lehmann's earlier contributions to the field are highlighted. To show how stochastic orders can be used in reliability contexts, let us consider the following example.

There are a few aging notions and three of them, IFR (increasing failure rate), IFRA (increasing failure rate average), and NBU (new better than used), are particularly useful. IFR implies IFRA, which in turn implies NBU. We now illustrate how the IFRA and NBU can naturally arise from Markov chains with stochastically monotone structures. We consider only the discrete case to ease the notations and a more complete survey can be found in [237].

Let $\{X_n, n \geq 0\}$ be a discrete-time, homogenous Markov chain on $\mathbb{R}_+$. The chain is said to be stochastically monotone if

$$[X_n | X_{n-1} = x] \leq_{\rm st} [X_n | X_{n-1} = x'], \text{ whenever } x \leq x'. \tag{4}$$

Consider the discrete first passage time $T_x := \inf\{n : X_n > x\}$. In such a discrete setting,

1. T_x is IFRA if either $\mathrm{P}\{T_x = 0\} = 1$ or $\mathrm{P}\{T_x = 0\} = 0$ and $[\mathrm{P}\{T_x > n\}]^{1/n}$ is decreasing in $n \geq 1$
2. T_x is NBU if $[T_x - m | T_x > m] \leq_{\rm st} T_x$ for all $m \geq 0$.

Theorem. *Assume that $\{X_n, n \geq 0\}$ is stochastically monotone.*

1. (Brown and Chaganty [79]) *T_x is NBU for any x.*

2. (Shaked and Shanthikumar [419]) *If, in addition, $\{X_n, n \geq 0\}$ has increasing sample paths, then T_x is IFRA for any x.*

That is, the aging properties NBU and IFRA emerge from Markov chains with stochastic order relation (4). The continuous-time version of this theorem can also be obtained. The comparison method used here is again robust and this theorem can be extended to a Markov chain with general partially ordered Polish state space.

It is well known that an IFRA life distribution arises from a weak limit of a sequence of coherent systems of independent, exponentially distributed components. The method used to establish such a result, however, is restricted to the continuous case (see, e.g., [38], page 87). In contrast, this result can be reestablished using a sequence of stochastically monotone Markov chains along the lines of the above theorem. More importantly, the stochastic order approach used in this theorem sheds structural insight on the fact that aging properties arise in a very natural way from stochastically monotone systems.

Many stochastic systems used in reliability and queueing modeling are indeed stochastically monotone in the sense of Eq. (4). The English edition of Dietrich Stoyan's book ([443], 1977 version in German, 1979 version in Russian) attracted quite a few queueing theorists in the 1980s and early 1990s to apply stochastic comparison methods to queueing modeling and analysis. The 1994 book by Moshe Shaked and George Shanthikumar included several chapters (written by some leading queueing and reliability theorists) that highlight research on stochastic orders in queueing and reliability contexts.

The comparison methods of stochastic processes have been discussed in detail in Szekli [446]. The studies on dependence and aging via stochastic orders are presented in Spizzichino [440]. An early study of stochastic orders in risk contexts is documented in Mosler [329] and more recent applications of stochastic orders to analyzing actuarial risks are discussed in Denuit et al. [117].

The most up-to- date, comprehensive treatments of stochastic orders are given by Müller and Stoyan [335] and Shaked and Shanthikumar [426].

Looking Forward

In the late 1980s and early 1990s, there were several international workshops focusing exclusively on stochastic orders and dependence. We mention some of them below.

- Symposium on Dependence in Probability and Statistics [62], Hidden Valley Conference Center, Pennsylvania, August 1–5, 1987. Organizers: H.W. Block, A.R. Sampson, and T.H. Savits

- Stochastic Orders and Decision Under Risk [330], Hamburg, Germany, May 16–20, 1989. Organizers: K. Mosler and M. Scarsini

- Stochastic Inequalities [431], Seattle, WA, July 1991. Organizers: Moshe Shaked and Y. L. Tong

- Distribution with Fixed Marginals and Related Topics [398], Seattle, WA, August, 1993. Organizers: L. Rüschendorf, B. Schweizer, and M. D. Taylor

These workshops and their proceedings enhanced communication and collaboration between scholars working in different fields and simulated research on stochastic orders and dependence. It is our hope that at the time we honor Professor Moshe Shaked, the Xiamen Workshop and this volume will revive the community workshop tradition on stochastic orders and dependence and strengthen research collaboration.

Last but not least, we would like to thank the School of Mathematical Sciences of Xiamen University for the support to the SORR2011. We would also like to express our sincere thanks to XL's graduate students Jianhua Lin, Jintang Wu, Yinping You, Rui Fang and Chen Li. Without their effort in organizing the Xiamen workshop, we would not have had such a wonderful academic meeting. Our special thanks go to Mr. Rui Fang, who helped us edit and revise the Latex source files of all submitted papers. Due to his enthusiasm and quiet efficiency, we finally present this nice volume (Fig. 2).

Pullman, WA, USA Haijun Li
Xiamen, China Xiaohu Li

List of Contributors

Félix Belzunce
Departamento Estadstica e Investigación Operativa, Universidad de Murcia, Espinardo, Spain

Hélène Cossette
École d'Actuariat, Université Laval, Québec, Canada

Gaofeng Da
Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei, China

Jacobo de Uña
Department of Statistics and Operations Research, Universidad de Vigo, Vigo, Spain

Antonio Di Crescenzo
Dipartimento di Matematica, Università di Salerno, Salerno, Italy

Weiyong Ding
College of Science, Hebei United University, Tangshan, China

Rachele Foschi
Laboratory of Innovation Management and Economics, IMT Advanced Studies Lucca, Lucca, Italy

Alba M. Franco-Pereira
Department of Statistics and Operations Research, Universidad de Vigo, Vigo, Spain

Taizhong Hu
Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei, China

Lei Hua
Division of Statistics, Northern Illinois University, DeKalb, IL, USA

Harry Joe
Department of Statistics, University of British Columbia, Vancouver, BC, Canada

Baha-Eldin Khaledi
Department of Statistics, Razi University, Kermanshah, Iran

Subhash Kochar
Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland, OR, USA

Haijun Li
Department of Mathematics, Washington State University, Pullman, WA, USA

Xiaohu Li
School of Mathematical Sciences, Xiamen University, Xiamen, China

Rosa E. Lillo
Department of Statistics, Universidad Carlos III de Madrid, Madrid, Spain

Maria Longobardi
Dipartimento di Matematica e Applicazioni, Università di Napoli Federico II, Napoli, Italy

Tiantian Mao
Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei, China

Etienne Marceau
École d'Actuariat, Université Laval, Québec, Canada

Alfred Müller
Department Mathematik, Universität Siegen, Siegen, Germany

Moshe Shaked
Department of Mathematics, University of Arizona, Tuscon, AZ, USA

Miguel A. Sordo
Departamento de Estadística e I. O., Universidad de Cádiz, Cádiz, Spain

Fabio Spizzichino
Dipartimento di Matematica G. Castelnuovo, Università degli Studi di Roma La Sapienza, Roma, Italy

Alfonso Suárez-Llorens
Departamento de Estadística e I. O., Universidad de Cádiz, Cádiz, Spain

Francis Pascual
Department of Mathematics, Washington State University, Pullman, WA, USA

Franco Pellerey
Dipartimento di Scienze Matematica, Politecnico di Torino, Torino, Italy

Vladimir Rykov
Department of Applied Mathematics and Computational Modeling, Gubkin Russian State University of Oil and Gas, Moscow, Russia

Nuria Torrado
Department of Statistics and Operations Research, Universidad Pública de Navarra, Pamplona, Spain

Maochao Xu
Department of Mathematics, Illinois State University, Normal, IL, USA

Yinping You
School of Mathematical Sciences, Xiamen University, Xiamen, China

Peng Zhao
School of Mathematical Sciences, Jiangsu Normal University, Xuzhou, China

Figure 2: SORR 2011, Xiamen, China, June 27–29, 2011

Contents

List of Figures

List of Tables

Part I

Theory of Stochastic Orders

Chapter 1
A Global Dependence Stochastic Order Based on the Presence of Noise

Moshe Shaked, Miguel A. Sordo, and Alfonso Suárez-Llorens

Abstract: Two basic ideas that give rise to global dependence stochastic orders were introduced and studied in Shaked et al. (Methodology and Computing in Applied Probability 14:617–648, 2012). Here these are reviewed, and two new ideas that give rise to new global dependence orders are then brought out and discussed. Two particular global dependence orders that come up from the two new general ideas are studied in detail. It is shown, among other things, how these orders can be identified and verified. In particular, conditions on the underlying copulas that yield these global dependence orders are given. The theory is illustrated by some examples. It is shown that some global dependence measures are preserved by the new global dependence orders. An application in reliability theory illustrates the usefulness of the new global dependence orders.

M. Shaked (✉)
Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA
e-mail: shaked@math.arizona.edu

M.A. Sordo • A. Suárez-Llorens
Departamento de Estadística e I. O., Universidad de Cádiz,
C/ Duque de Nájera 8, CP: 11002, Cádiz, Spain
e-mail: mangel.sordo@uca.es; alfonso.suarez@uca.es

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_1,

1.1 Introduction

Since the late 1960s a number of researchers have introduced and studied stochastic orders that compare the strength of the positive dependence of the compared random vectors. Some important contributions in this area can be found in the papers by Yanagimoto and Okamoto [482], Kimeldorf and Sampson [238], Fang and Joe [157], Bäuerle [46], Shaked and Shanthikumar [424], Avérous et al. [23], Colangelo et al. [92], Dolati et al. [134], as well as in the monograph by Joe [211]. Further references and an extensive overview of positive dependence stochastic orders can be found in Chap. 9 of Shaked and Shanthikumar [426].

The terminology of "positive dependence orders" follows from the fact that such orders compare random vectors according to the strength of their *positive dependence*. Thus, a "positive dependent" random vector is larger, with respect to such orders, than a vector of independent random variables, and the latter is larger than a "negative dependent" random vector (here the exact definitions of positive and negative dependence depend on the context). However, a vector of random variables that strongly depend on each other may not be comparable to a vector of independent random variables, with respect to a positive dependent order, if the strongly dependent random variables are not positively (or negatively) dependent.

In order to avoid this drawback of positive dependence stochastic orders, one needs to define and study notions of global dependence stochastic orders (GDO). A step in this direction has been made recently by Dette et al. [122]. They introduced some novel notions of GDO based on *conditional distributions*. A drawback of their notions is that, according to these notions, vectors of independent random variables may not be comparable to (explicitly, need not be "smaller" than) other vectors that have the same marginal distributions.

Another recent work in this area is the paper of Shaked et al. [428]. They described two other ideas, based on *conditional distributions*, that give rise to some GDO. Some of the resulting orders have the desirable properties that every random vector is larger, with respect to the these global dependence orders, than the vector of independent random variables with the same marginal distributions. Also, every random vector is smaller, with respect to the these global dependence orders, than the vector of totally dependent random variables with the same marginal distributions.

In this paper we introduce and discuss further new ideas that yield some GDO.

It is worth mentioning that a completely different approach to non-monotone dependence orders was taken by Silvey [433], Ali and Silvey [7, 8], Joe [209, 210], and Scarsini [411]. This other approach is based on the variation of the values of the copulas that are associated with the compared random vectors and not on conditional distributions. Thus it has an entirely different flavor and applications than the orders that are studied in Dette et al. [122], in Shaked et al. [428], and in this paper.

In the sequel, "increasing" and "decreasing" stand, respectively, for "nondecreasing" and "nonincreasing." The symbol "$\stackrel{\text{st}}{=}$" denotes equality in law. For any distribution function F, we let $F^{-1}(p) = \sup\{x : F(x) \le p\}$, $0 < p < 1$, denote the corresponding right-continuous quantile function. For any random vector $\boldsymbol{Z}$ and an event A we denote by $[\boldsymbol{Z}|A]$ any random vector that is distributed according to the conditional distribution of $\boldsymbol{Z}$ given A.

1.2 Two Previous Ideas

In order to motivate and intuitively grasp the new ideas of this paper, we first describe the ideas of Shaked et al. [428].

Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be two random vectors such that $X \stackrel{\text{st}}{=} \tilde{X}$ and $Y \stackrel{\text{st}}{=} \tilde{Y}$. Let us consider the conditional random variables $\big[Y\big|X = x\big]$ and $\big[\tilde{Y}\big|\tilde{X} = x\big]$ for $x \in \text{support}(X)$. Define the regression curves

$$m(x) = \mathrm{E}\big[Y\big|X = x\big] \quad \text{and} \quad \tilde{m}(x) = \mathrm{E}\big[\tilde{Y}\big|\tilde{X} = x\big], \quad x \in \text{support}(X), \tag{1.2.1}$$

and the error curves

$$e(x) = \mathrm{Var}\big[Y\big|X = x\big] \quad \text{and} \quad \tilde{e}(x) = \mathrm{Var}\big[\tilde{Y}\big|\tilde{X} = x\big], \quad x \in \text{support}(X), \tag{1.2.2}$$

provided that they exist.

First let us consider $m(X)$ and $\tilde{m}(\tilde{X})$. Intuitively, if X and Y are "close to independence," then by observing $X = x$ we learn "almost nothing" about Y; that is, Y does not vary much with X, or, in other words, "Y does not inherit much of the variability of X." As a result, $m(X)$ has a small variability. In the extreme case of independence we have that $m(X) = \mathrm{E}Y$, that is, $m(X)$ is a constant that has no variability at all. On the other hand, intuitively, if $\tilde{X}$ and $\tilde{Y}$ are "close to total dependence," then by observing $\tilde{X} = x$ we learn "a lot" about $\tilde{Y}$; that is, $\tilde{Y}$ varies much with $\tilde{X}$, or, in other words, "$\tilde{Y}$ inherits much of the variability of $\tilde{X}$." Therefore it makes sense to define (X, Y) as smaller than $(\tilde{X}, \tilde{Y})$, with respect to some corresponding GDO, if

$$m(X) \leq_{\text{variability}} \tilde{m}(\tilde{X}), \tag{1.2.3}$$

where $\leq_{\text{variability}}$ is a univariate variability stochastic order (such as $\leq_{\text{cx}}$, or $\leq_{\text{disp}}$, or $\leq_{\text{ew}}$; for the exact definition of these orders see the sequel and/or Shaked and Shanthikumar [426]). Formally, Shaked et al. [428] defined a family of GDO, indexed by the univariate order $\leq_{\text{variability}}$, by postulating that

$$(X, Y) \leq_{\text{GDO}_1\text{-variability}} (\tilde{X}, \tilde{Y})$$

if

$$\mathrm{E}[Y \big| X] \leq_{\text{variability}} \mathrm{E}\big[\tilde{Y} \big| \tilde{X}\big] \tag{1.2.4}$$

(the subscript 1 under GDO above indicates that this order arises from Idea 1). Note that the condition (1.2.4) is a rewrite of the condition (1.2.3).

Note that the order $\leq_{\text{GDO}_1\text{-variability}}$ is based on the comparison of the *variability of the conditional expectations* $\mathrm{E}[Y \big| X]$ and $\mathrm{E}\big[\tilde{Y} \big| \tilde{X}\big]$. An alternative idea is to define a GDO through a comparison of the *magnitude of the conditional variances* $e(X) = \mathrm{Var}[Y \big| X]$ and $\tilde{e}(\tilde{X}) = \mathrm{Var}\big[\tilde{Y} \big| \tilde{X}\big]$.

So let us now consider $e(X)$ and $\tilde{e}(\tilde{X})$. Intuitively, again, if X and Y are "close to independence," then by observing $X = x$ we learn "almost nothing" about Y. One way to interpret this is to note that the "uncertainty about Y" then, given X, is relatively "large"; namely, $\mathrm{Var}[Y \big| X]$ is "large." In the extreme case of independence we have that the conditional variance $\mathrm{Var}[Y \big| X]$ is as large as possible; that is, it satisfies $\mathrm{Var}[Y \big| X] = \mathrm{Var}[Y]$. On the other hand, intuitively, again, if $\tilde{X}$ and $\tilde{Y}$ are "close to total dependence," then by observing $\tilde{X} = x$ we learn "a lot" about $\tilde{Y}$. One way to interpret this is to note that the "uncertainty about $\tilde{Y}$" then, given $\tilde{X}$, is relatively "small"; namely, $\mathrm{Var}\big[\tilde{Y}\big|\tilde{X}\big]$ is "small." In the extreme case of total dependence (t.e., when $\big[\tilde{Y}\big|\tilde{X}=x\big]$ is degenerate for each $x \in \mathrm{support}(X)$) we see that $\mathrm{Var}\big[\tilde{Y}\big|\tilde{X}\big]$ is as small as possible; that is, it satisfies $\mathrm{Var}\big[\tilde{Y}\big|\tilde{X}\big]=0$. Therefore it makes sense to define (X, Y) as smaller than $(\tilde{X}, \tilde{Y})$, with respect to some corresponding GDO, if

$$e(X) \geq_{\text{magnitude}} \tilde{e}(\tilde{X}), \tag{1.2.5}$$

where $\leq_{\text{magnitude}}$ is a univariate stochastic order of magnitude (such as $\leq_{\text{st}}$; for the exact definition of this order see the sequel and/or Shaked and Shanthikumar [426]). Formally, Shaked et al. [428] defined a family of GDO, indexed by the univariate order $\leq_{\text{magnitude}}$, by postulating that

$$(X, Y) \leq_{\text{GDO}_2\text{-magnitude}} (\tilde{X}, \tilde{Y})$$

if

$$\mathrm{Var}\big[Y\big|X\big] \geq_{\text{magnitude}} \mathrm{Var}\big[\tilde{Y}\big|\tilde{X}\big] \tag{1.2.6}$$

(the subscript 2 under GDO above indicates that this order arises from Idea 2). Note that the condition (1.2.6) is a rewrite of the condition (1.2.5).

Shaked et al. [428] obtained a better insight into the ideas that underlie the global dependence orders $\leq_{\text{GDO}_1\text{-variability}}$ and $\leq_{\text{GDO}_2\text{-magnitude}}$ as follows. Consider a random vector (X, Y) and its regression and error curves m and e and define the following random variable (which is a function of X and Y):

$$Z = \frac{Y - m(X)}{\sqrt{e(X)}}. \tag{1.2.7}$$

Shaked et al. [428] noticed that the random variable Z is standard (i.e., it has mean 0 and variance 1). Although in general Z is not independent of X, it satisfies $\mathrm{Cov}(h(X), Z) = 0$ for all functions h. Using Z in Eq. (1.2.7), Y can be expressed as

$$Y = m(X) + \sqrt{e(X)} \cdot Z, \tag{1.2.8}$$

where $\mathrm{Cov}(\sqrt{e(X)}, Z) = 0$ and $\mathrm{Cov}(\sqrt{m(X)}, \sqrt{e(X)} \cdot Z) = 0$.

For any conditional random variable below, say, $[Y\big|X = x]$, we denote the corresponding conditional distribution function by $F_{Y|X=x}$, and the corresponding inverse by $F^{-1}_{Y|X=x}$. When x in the expression $F^{-1}_{Y|X=x}$ is substituted by X, we simply denote it as $F^{-1}_{Y|X}$.

From Eq. (1.2.7) it is clear that

$$F^{-1}_{Y|X=x}(u) = m(x) + \sqrt{e(x)} \cdot F^{-1}_{Z|X=x}(u), \quad u \in (0,1).$$

Therefore

$$Y \stackrel{\text{st}}{=} F^{-1}_{Y|X}(U) = m(X) + \sqrt{e(X)} \cdot F^{-1}_{Z|X}(U), \tag{1.2.9}$$

where U is a uniform random variable on $[0,1]$, which is independent of X. Using the notation in Eq. (1.2.9), we note that $F^{-1}_{Z|X}(U) \stackrel{\text{st}}{=} Z$, so that $F^{-1}_{Z|X}(U)$ has mean 0 and variance 1. From Eq. (1.2.9) it is clear that $m(X)$ and $e(X)$ have different effects when we predict Y. If we are looking for definitions of global dependence orders, we may try to order $m(X)$ in variability and $e(X)$ in magnitude. These are the ideas that underlie the definitions of $\leq_{\text{GDO}_1\text{-variability}}$ and $\leq_{\text{GDO}_2\text{-magnitude}}$. The expression (1.2.9) will be used in the sequel.

Based on some ideas in Dabrowska [101] and in Dette et al. [122], we listed in Shaked et al. [428] some desirable properties for any global dependence order. In two of these properties we use the following notation. Let X and Y be two random variables with a general joint distribution. Let $X^{\perp}$ and $Y^{\perp}$ be two independent random variables such that

$$X^{\perp} \stackrel{\text{st}}{=} X \quad \text{and} \quad Y^{\perp} \stackrel{\text{st}}{=} Y. \tag{1.2.10}$$

Furthermore, let h be a function such that

$$h(X) \stackrel{\text{st}}{=} Y,$$

and consider the random variables $X^{\top}$ and $Y^{\top}$ defined by

$$X^{\top} = X \quad \text{and} \quad Y^{\top} = h(X). \tag{1.2.11}$$

That is, $(X^{\perp}, Y^{\perp})$ is the "independent version" of (X, Y), whereas $(X^{\top}, Y^{\top})$ is a "totally dependent" version of (X, Y).

The following desirable properties were studied in Shaked et al. [428]:

(O1) The order $\le_{\mathrm{GDO}}$ is reflexive and transitive.

(O2) If $(X,Y) \le_{\mathrm{GDO}} (\tilde{X},\tilde{Y})$, and if (V,W) and $(\tilde{V},\tilde{W})$ satisfy $V \overset{\mathrm{st}}{=} \tilde{V} \overset{\mathrm{st}}{=} X$ and $W \overset{\mathrm{st}}{=} \tilde{W} \overset{\mathrm{st}}{=} Y$, as well as $\mathrm{E}[W|V] \overset{\mathrm{st}}{=} \mathrm{E}[Y|X]$ and $\mathrm{E}[\tilde{W}|\tilde{V}] \overset{\mathrm{st}}{=} \mathrm{E}[\tilde{Y}|\tilde{X}]$, then $(V,W) \le_{\mathrm{GDO}} (\tilde{V},\tilde{W})$.

(O3) If $(X,Y) \le_{\mathrm{GDO}} (\tilde{X},\tilde{Y})$ then $(\phi(X), l(Y)) \le_{\mathrm{GDO}} (\phi(\tilde{X}), l(\tilde{Y}))$, where ϕ is a one-to-one measurable function, and l is a linear function.

(O4) Let (X,Y) be any random vector, then $(X,Y) \ge_{\mathrm{GDO}} (X^{\perp}, Y^{\perp})$.

(O5) Let (X,Y) be any random vector, then $(X,Y) \le_{\mathrm{GDO}} (X^{\top}, Y^{\top})$.

Remark 1.2.1. Shaked et al. [428] noticed that property (O2) can be interpreted as follows: It says that all bivariate random vectors, that have the same margins and the same distribution for $m(X)$, form an equivalence class for global dependence orders that satisfy property (O2). Thus, property (O2) is a desirable property when the global dependence between X and Y is measured by a quantity that is based on the distribution or the expectation of the error curve $e(X)$. Such measures are natural for orders that arise from Idea 1, and therefore property (O2) may be a desirable one for orders of the type $\le_{\mathrm{GDO_1\text{-}variability}}$. But it may be too strong a property for other global dependence orders.

1.3 Two New Ideas

In order to motivate the new ideas for GDOs we find it handy to use the following terminology (see, e.g., Mizuno [328], Ganuza and Penalva [172], and Wu and Mielniczuk [476]). Consider the random vector (X,Y). Let us think about X as the unobservable or unknown state of the world or, alternatively, as the input into some physical system. Furthermore, let us think about Y as the noisy signal or, alternatively, the output that corresponds to X. Similarly, when we consider $(\tilde{X},\tilde{Y})$, we can think about $\tilde{X}$ as the input into and about $\tilde{Y}$ as the output from, some, conceptually other, physical system.

Again, let (X,Y) be a random vector. If U is a uniform (0,1) random variable that is independent of X, then

$$(X,Y) \overset{\mathrm{st}}{=} \big(X, F^{-1}_{Y|X}(U)\big). \tag{1.3.1}$$

Based on Eq. (1.3.1), we can view Y as the outcome of the bivariate function

$$Y \overset{\text{st}}{=} F_{Y|X}^{-1}(U).$$

To avoid confusion we follow the notation of Wu and Mielniczuk [476] and denote

$$G(x,u) = F_{Y|X=x}^{-1}(u), \tag{1.3.2}$$

and thus we can view Y as the outcome of the bivariate function

$$Y \overset{\text{st}}{=} G(X,U). \tag{1.3.3}$$

In the words of Wu and Mielniczuk [476] "Y is viewed as the output from a random physical system with X and U being the input and the noise or error, respectively." They interpret dependence "as how the output Y depends on the input X in *the presence of the noise* U." Note that U summarily describes all the factors (others than X) that influence Y.

Note that an expression for $G(X,U)$ that uses the random variable Z from Eq. (1.2.7) is given in Eq. (1.2.9). This expression will be used later in Example 1.4.13.

Remark 1.3.1. The effects on Y, of the random variables X and U in Eq. (1.3.3), contradict and complement each other. Intuitively it is seen that the "more dependent" Y is on X, the "less dependent" it is on U. Similarly, the "less dependent" Y is on X, the "more dependent" it is on U. This intuition leads us to definitions of GDOs, which are presented in Eqs. (1.3.12) and (1.3.13) below.

Remark 1.3.2. It is worthwhile to point out that the expression (1.3.1) is essentially equivalent to the classical standard construction transformation (see, for instance, Rüschendorf and de Valk [397] or Shaked and Shanthikumar [426, Sect. 6.B.3]). In order to see it, consider, in addition to the uniform (0,1) random variable U, another uniform (0,1) random variable V that is independent of U. Denote the distribution function of X by F_X, and let F_X^{-1} be the corresponding inverse. Then the classical standard construction expands Eq. (1.3.1) as follows:

$$(X,Y) \overset{\text{st}}{=} (X, F_{Y|X}^{-1}(U)) \overset{\text{st}}{=} \left(F_X^{-1}(V), F_{Y|X=F_X^{-1}(V)}^{-1}(U)\right).$$

Furthermore, analogously to Eq. (1.3.3), Y can be expressed as

$$Y \stackrel{\text{st}}{=} G(X,U) \stackrel{\text{st}}{=} F^{-1}_{Y|X=F_X^{-1}(V)}(U) \stackrel{\text{st}}{=} H(V,U), \tag{1.3.4}$$

where H is defined by $H(v,u) = F^{-1}_{Y|X=F_X^{-1}(v)}(u))$, $(u,v) \in (0,1)^2$. It is apparent from Eq. (1.3.4) that we need two independent uniform random variables V and U to explain Y. In Shaked et al. [428] we studied global dependence between X and Y through the effect of V. The definitions of the new GDOs that are introduced in Eqs. (1.3.12) and (1.3.13) below study the global dependence between X and Y through the effect of U.

We can also write Eq. (1.2.9) in terms of V and U as

$$\begin{aligned} Y \stackrel{\text{st}}{=} F^{-1}_{Y|X}(U) &= m(X) + \sqrt{e(X)} \cdot F^{-1}_{Z|X}(U) \\ &\stackrel{\text{st}}{=} m(F_X^{-1}(V)) + \sqrt{e(F_X^{-1}(V))} \cdot F^{-1}_{Z|X=F_X^{-1}(V)}(U). \end{aligned}$$

It is clear that $\mathrm{E}(Y\big|V=u)$ and $\mathrm{E}(Y\big|U=u)$ have different interpretations.

Before we proceed to the definitions of the new GDOs, we derive two interesting relationships between the copulas of (X,Y) and (U,Y) that will be useful in the sequel. So let X, Y, and U be as in Eq. (1.3.3). To recall the definition of a copula that is needed here, let (X,Y) have the continuous joint distribution function $F_{X,Y}$ with the marginal distribution functions F_X and F_Y. Then the function $C_{(X,Y)}$, defined by

$$C_{(X,Y)}(u,v) = F_{X,Y}\big(F_X^{-1}(u), F_Y^{-1}(v)\big), \quad (u,v) \in (0,1)^2,$$

is a bivariate distribution function with marginal distributions that are uniform on (0,1), and it is called the copula that is associated with (X,Y). From Nelsen [355, Theorem 2.2.7] we know that the partial derivative $\partial C_{(X,Y)}(u,v)/\partial v$ exists for almost all $v \in (0,1)$. The copula $C_{(U,Y)}$ that is associated with (U,Y) is similarly defined.

Proposition 1.3.3. *Let X, Y, and U be as in Eq. (1.3.3), and assume that (X,Y) and (U,Y) have continuous distribution functions. Then*

$$C_{(U,Y)}(u,v) = \int_0^1 \min\Big\{u, \frac{\partial}{\partial p} C_{(X,Y)}(p,v)\Big\} \mathrm{d}p.$$

Proof: From Eq. (1.3.4) we see that

$$(U, Y) \stackrel{\text{st}}{=} \left(U, F^{-1}_{Y|X=F_X^{-1}(V)}(U)\right),$$

where V is a uniform (0,1) random variable, independent of U. Therefore $C_{(U,Y)}$ is the distribution function of

$$(U, F_Y(Y)) \stackrel{\text{st}}{=} \left(U, F_Y(F^{-1}_{Y|X=F_X^{-1}(V)}(U))\right).$$

Hence

$$\begin{aligned}
C_{(U,Y)}(u,v) &= \mathrm{P}\left\{U \le u, F_Y\left(F^{-1}_{Y|X=F_X^{-1}(V)}(U)\right) \le v\right\} \\
&= \mathrm{P}\left\{U \le u, F^{-1}_{Y|X=F_X^{-1}(V)}(U) \le F_Y^{-1}(v)\right\} &\quad (1.3.5)\\
&= \int_0^1 \mathrm{P}\left\{U \le u, F^{-1}_{Y|X=F_X^{-1}(p)}(U) \le F_Y^{-1}(v)\right\} \mathrm{d}p \\
&= \int_0^1 \mathrm{P}\left\{U \le u, U \le F_{Y|X=F_X^{-1}(p)}\left(F_Y^{-1}(v)\right)\right\} \mathrm{d}p \\
&= \int_0^1 \mathrm{P}\left\{U \le u, U \le \frac{\partial}{\partial p} C_{(X,Y)}(p,v)\right\} \mathrm{d}p &\quad (1.3.6)\\
&= \int_0^1 \min\left\{u, \frac{\partial}{\partial p} C_{(X,Y)}(p,v)\right\} \mathrm{d}p, &\quad (1.3.7)
\end{aligned}$$

where Eq. (1.3.6) follows from (2.9.1) in Nelsen [355] and Eq. (1.3.7) from the fact that U is a uniform (0,1) random variable. ■

The relationship between $C_{(X,Y)}$ and $C_{(U,Y)}$ in Proposition 1.3.3 goes along nicely with the intuition. For instance, if X and Y are independent, that is, $C_{(X,Y)}(u,v) = uv$, then $C_{(U,Y)}(u,v) = \min\{u,v\}$, which means that U and Y are comonotone. On the other hand, if X and Y are comonotone, then it is easy to see that $C_{(U,Y)}(u,v) = uv$; that is, U and Y are independent.

Recall from Lehmann [278] that a random vector (X, Y) is called positive regression dependent (PRD) if $[Y \big| X = x]$ stochastically increases in x (in such a case (X, Y) is also often called stochastically increasing in sequence).

Proposition 1.3.4. *Let X, Y, and U be as in Eq. (1.3.3), and assume that (X,Y) and (U,Y) have continuous distribution functions. Furthermore, assume that (X,Y) is* PRD. *Then*

$$C_{(U,Y)}(u,v) = \int_0^u z(w,v)\,\mathrm{d}w,$$

where

$$z(w,v) = \begin{cases} \sup\left\{p : \frac{\partial}{\partial p} C_{(X,Y)}(p,v) \geq w\right\}, & \textit{if the supremum is well defined;} \\ 0, & \textit{otherwise.} \end{cases}$$

Proof: Starting from Eq. (1.3.5) in the proof of Proposition 1.3.3 we have

$$\begin{aligned} C_{(U,Y)}(u,v) &= \mathrm{P}\left\{U \leq u, F^{-1}_{Y|X=F_X^{-1}(V)}(U) \leq F_Y^{-1}(v)\right\} \\ &= \int_0^u \mathrm{P}\left\{F^{-1}_{Y|X=F_X^{-1}(V)}(w) \leq F_Y^{-1}(v)\right\} \mathrm{d}w \\ &= \int_0^u \mathrm{P}\left\{F_{Y|X=F_X^{-1}(V)}(F_Y^{-1}(v)) \geq w\right\} \mathrm{d}w \\ &= \int_0^u \mathrm{P}\left\{V \in \{p : F_{Y|X=F_X^{-1}(p)}(F_Y^{-1}(v)) \geq w\}\right\} \mathrm{d}w \\ &= \int_0^u z(w,v)\,\mathrm{d}w, \end{aligned}$$

where the last equality follows from the PRD assumption and (2.9.1) in Nelsen [355]. ■

We now proceed to the definitions of the new GDOs. Let (X,Y) and $(\tilde{X},\tilde{Y})$ be two random vectors such that $X \overset{\text{st}}{=} \tilde{X}$ and $Y \overset{\text{st}}{=} \tilde{Y}$. Write

$$Y = G(X,U) \quad \text{and} \quad \tilde{Y} = \tilde{G}(\tilde{X},\tilde{U}), \tag{1.3.8}$$

where G is defined in Eq. (1.3.2), U is a uniform (0,1) random variable that is independent of X, $\tilde{G}$ is defined by $\tilde{G}(x,u) = F^{-1}_{\tilde{Y}|\tilde{X}=x}(u)$, and $\tilde{U}$ is a uniform (0,1) random variable that is independent of $\tilde{X}$. Note that in Eq. (1.3.8) we assume, without loss of generality, almost sure equality rather than stochastic equality as in Eq. (1.3.3).

Referring to the representation (1.3.8), define the conditional expectations

$$k(u) = \mathrm{E}\big[Y\big|U = u\big] \quad \text{and} \quad \tilde{k}(u) = \mathrm{E}\big[\tilde{Y}\big|\tilde{U} = u\big], \quad u \in (0,1), \tag{1.3.9}$$

and the conditional variances

$$d(u) = \mathrm{Var}\big[Y\big|U = u\big] \quad \text{and} \quad \tilde{d}(u) = \mathrm{Var}\big[\tilde{Y}\big|\tilde{U} = u\big], \quad u \in (0,1), \tag{1.3.10}$$

provided that they exist.

We will now describe intuitive reasonings that will lead us to two new approaches that yield definitions of families of GDOs. The reasonings parallel those of Shaked et al. [428], as was recounted earlier in Sect. 1.2.

To describe the first approach, let Y and $\tilde{Y}$ be as in Eq. (1.3.8) and consider $k(U)$ and $\tilde{k}(\tilde{U})$. First let us intuitively examine the case when $\tilde{U}$ and $\tilde{Y}$ are "close to independence" (note that this means that $\tilde{X}$ and $\tilde{Y}$ are "close to total dependence"). Then the event $\tilde{U} = u$ tells us "almost nothing" about $\tilde{Y}$; that is, $\tilde{Y}$ does not vary much with $\tilde{U}$. As a result, $\tilde{k}(\tilde{U})$ has a small variability. In the extreme case when $\tilde{U}$ and $\tilde{Y}$ are independent we have that $\tilde{Y} = \tilde{G}(\tilde{X}, \tilde{U})$ does not depend on $\tilde{U}$ and then $\tilde{k}(\tilde{U})$ is a constant that has no variability at all; note that in this case we can write

$$\tilde{Y} = \tilde{h}(\tilde{X}) \tag{1.3.11}$$

for some function $\tilde{h}$, and therefore $\tilde{Y}$ then is totally dependent on $\tilde{X}$ (although the dependence need not be one-to-one). On the other hand, let us now examine the case when U and Y are "close to total dependence" (note that this means that X and Y are "close to independence"). In this case the event $U = u$ tells us "a lot" about Y. That is, Y varies much with U. As a result, $k(U)$ has a large variability. Therefore, according to the approach that is motivated by the intuition above, it makes sense to define (X, Y) as smaller than $(\tilde{X}, \tilde{Y})$, with respect to some corresponding GDO, if

$$k(U) \geq_{\text{variability}} \tilde{k}(\tilde{U}),$$

or, equivalently, if

$$\mathrm{E}[Y\big|U] \geq_{\text{variability}} \mathrm{E}\big[\tilde{Y}\big|\tilde{U}\big],$$

where $\leq_{\text{variability}}$ is some univariate variability stochastic order. If either one of the above equivalent inequalities holds we denote it by

$$(X, Y) \leq_{\text{GDO}_3\text{-variability}} (\tilde{X}, \tilde{Y}) \tag{1.3.12}$$

(the subscript 3 under GDO above indicates that this order arises from the approach that is described in the present paragraph).

To describe the second approach, let, again, Y and $\tilde{Y}$ be as in Eq. (1.3.8), but now we consider $d(U)$ and $\tilde{d}(\tilde{U})$. First let us intuitively examine the case when $\tilde{U}$ and $\tilde{Y}$ are "close to independence" (as before, this means that $\tilde{X}$ and $\tilde{Y}$ are "close to total dependence"). Then, again, the event $\tilde{U} = u$ tells us "almost nothing" about $\tilde{Y}$; that is, given $\tilde{U}$, the uncertainty about $\tilde{Y}$ is large. As a result, $\tilde{d}(\tilde{U})$ is large. In the extreme case when $\tilde{U}$ and $\tilde{Y}$ are independent we have that $\tilde{d}(\tilde{U})$ is as large as possible, that is, $\tilde{d}(\tilde{U}) = \text{Var}[\tilde{Y}]$; we have noticed earlier that in this case $\tilde{Y}$ then is totally dependent on $\tilde{X}$ (although the dependence need not be one-to-one). On the other hand, let us now examine the case when U and Y are "close to total dependence" (as before, this means that X and Y are "close to independence"). In this case the event $U = u$ tells us "a lot" about Y; that is, given U, the uncertainty about Y is small. As a result, $d(U)$ is small. Therefore, according to this second approach that is motivated by the intuition above, it makes sense to define (X, Y) as smaller than $(\tilde{X}, \tilde{Y})$, with respect to some corresponding GDO, if

$$d(U) \leq_{\text{magnitude}} \tilde{d}(\tilde{U}),$$

or, equivalently, if

$$\text{Var}[Y \big| U] \leq_{\text{magnitude}} \text{Var}\big[\tilde{Y} \big| \tilde{U}\big],$$

where $\leq_{\text{magnitude}}$ is some univariate stochastic order of magnitude. If either one of the above equivalent inequalities holds we denote it by

$$(X, Y) \leq_{\text{GDO}_4\text{-magnitude}} (\tilde{X}, \tilde{Y}) \tag{1.3.13}$$

(the subscript 4 under GDO above indicates that this order arises from the approach that is described in the present paragraph).

We emphasize that the GDO_3 and GDO_4 orders depend on the capacity of prediction of Y (respectively, $\tilde{Y}$) from X (respectively, $\tilde{X}$). In other words, the GDO_3 and GDO_4 orders are not symmetric in the sense that $(X, Y) \le_{\text{GDO}_3\text{-magnitude}} (\tilde{X}, \tilde{Y})$ (respectively, $(X, Y) \le_{\text{GDO}_4\text{-magnitude}} (\tilde{X}, \tilde{Y})$) may hold, whereas $(Y, X) \le_{\text{GDO}_3\text{-magnitude}} (\tilde{Y}, \tilde{X})$ (respectively, $(Y, X) \le_{\text{GDO}_4\text{-magnitude}} (\tilde{Y}, \tilde{X})$) need not hold at the same time. That is, the GDO_3 and GDO_4 orders have the same drawback that the GDO_1 and GDO_2 orders have, as discussed at the end of Sect. 2.4 in Shaked et al. [428]. Even in the extreme case of total dependence of $\tilde{Y}$ on $\tilde{X}$ [i.e., when Eq. (1.3.11) holds], if $\tilde{h}$ in Eq. (1.3.11) is not one-to-one then we do not have total dependence of $\tilde{X}$ on $\tilde{Y}$ in the sense that $\tilde{X}$ may not be predicted from $\tilde{Y}$ with certainty.

Before we close this section, a comment on the uniform random variables U and V of Remark 1.3.2 is in place. The random variable V there can be thought of as the "generator" of the input X into some physical system, whereas the random variable U is the "generator" of the corresponding noise. From Remark 1.3.1 we see that V and U "contradict and complement each other." That is, the "more dependent" Y is on V, the "less dependent" it is on U, and vice versa. We may wonder how useful is the input X for the prediction of Y. One way of telling that is to compare the global dependence of (U, Y) with the global dependence of (V, Y). For example, the orders $\le_{\text{GDO}_1\text{-cx}}$ or $\le_{\text{GDO}_1\text{-disp}}$ (which are formally defined as special cases of Eq. (1.2.4) and which are studied in detail in Shaked et al. [428]) or the order $\le_{\text{GDO}_2\text{-st}}$ (which is formally defined as a special case of Eq. (1.2.6) and which is also studied in detail in [428]) can be used for such comparisons. Specifically, if it happens that $(V, Y) \le_{\text{GDO}_1\text{-cx}} (U, Y)$, or $(V, Y) \le_{\text{GDO}_1\text{-disp}} (U, Y)$, or $(V, Y) \le_{\text{GDO}_2\text{-st}} (U, Y)$, then we could say that prediction of Y from X is not recommended. Here is a particular example.

Example 1.3.5. Let (X, Y) be a bivariate normal random vector. For simplicity we assume that X and Y are standard normal random variables with correlation coefficient ρ. A straightforward computation leads to the function H in Eq. (1.3.4) as follows:

$$H(v, u) = F^{-1}_{Y|X=\phi^{-1}(v)}(u) = \rho\phi^{-1}(v) + \phi^{-1}(u)\sqrt{1-\rho^2},$$

where ϕ^{-1} denotes the inverse of the standard normal distribution function. Then

$$[Y \big| U = u] \stackrel{\text{st}}{=} \rho\phi^{-1}(V) + \phi^{-1}(u)\sqrt{1-\rho^2}, \tag{1.3.14}$$

and

$$[Y \big| V = v] \stackrel{\text{st}}{=} \rho\phi^{-1}(v) + \phi^{-1}(U)\sqrt{1-\rho^2}. \tag{1.3.15}$$

From Eq. (1.3.14) it follows that $\mathrm{E}[Y \big| U]$ has a normal distribution with mean 0 and variance ρ^2, whereas from Eq. (1.3.15) it follows that $\mathrm{E}[Y \big| V]$ has a normal distribution with mean 0 and variance $1 - \rho^2$. From these observations it follows that $(V, Y) \le_{\text{GDO}_1\text{-cx}} (U, Y) \iff \rho^2 \le 1 - \rho^2 \iff \rho^2 \le 1/2$ (see Example 3.4 in Shaked et al. [428]), and that $(V, Y) \le_{\text{GDO}_1\text{-disp}} (U, Y) \iff \rho^2 \le 1/2$ (see Example 4.5 in [428]), and that $(V, Y) \le_{\text{GDO}_2\text{-st}} (U, Y) \iff \rho^2 \le 1/2$ (see Example 5.5 in [428]). Thus, if $|\rho| \le \sqrt{2}/2$ (i.e., if the determination coefficient $R^2 = \rho^2 \cdot 100\,\% < 50\,\%$), it may be recommended not to use X to predict Y. ■

1.4 Some Properties of the New Orders

1.4.1 The Order $\le_{\mathbf{GDO_3\text{-cx}}}$

First we note that the univariate order $\le_{\text{cx}}$ is reflexive and transitive. Thus we have the following result.

Proposition 1.4.1. *The order* $\le_{\text{GDO}_3\text{-cx}}$ *satisfies property* (O1).

We do not know whether $\le_{\text{GDO}_3\text{-cx}}$ satisfies property (O2). But, following the comments in Remark 1.2.1 we conjecture that it does not. However, we have the following result.

Proposition 1.4.2. *The order* $\le_{\text{GDO}_3\text{-cx}}$ *satisfies property* (O3).

Proof: Let (X, Y) be a random vector. Furthermore, let ϕ be a one-to-one measurable function, and consider $(\phi(X), Y)$. Let $G(x, u)$ be defined as in Eq. (1.3.2), and define G_ϕ by

$$G_\phi(x, u) = G(\phi^{-1}(x), u).$$

Let U be a uniform random variable on $[0, 1]$, which is independent of X. Since, by Eq. (1.3.3), $Y \stackrel{\text{st}}{=} G(X, U)$, it follows that

$$Y \stackrel{\text{st}}{=} G_\phi(\phi(X), U).$$

Therefore the function $k(u)$ that is defined in Eq. (1.3.9) can be expressed both as

$$k(u) = \mathrm{E}[Y \big| U = u] = \mathrm{E}[G(X, u)]$$

and as

$$k(u) = \mathrm{E}[Y \big| U = u] = \mathrm{E}[G_\phi(\phi(X), u)].$$

In other words, both vectors (X, Y) and $(\phi(X), Y)$ share the same function $k(u)$. Similarly, if $(\tilde{X}, \tilde{Y})$ is another random vector then $(\tilde{X}, \tilde{Y})$ and $(\phi(\tilde{X}), \tilde{Y})$ share the same function $\tilde{k}(u)$ defined in Eq. (1.3.9). Thus, if $(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$ then $(\phi(X), Y) \leq_{\text{GDO}_3\text{-cx}} (\phi(\tilde{X}), \tilde{Y})$.

Next, it is easy to see that if $(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$ then $(X, l(Y)) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, l(\tilde{Y}))$ for every linear function l.

Property (O3) now follows from the above two observations. ■

We now proceed to obtain properties (O4) and (O5). Recall the notation $X^\perp$, $Y^\perp$, $X^\top$, and $Y^\top$ from Eqs. (1.2.10) and (1.2.11).

Proposition 1.4.3. *With the notation in Eqs.* (1.2.10) *and* (1.2.11), *for every random vector* (X, Y), *we have*

$$(X^\perp, Y^\perp) \leq_{\text{GDO}_3\text{-cx}} (X, Y) \leq_{\text{GDO}_3\text{-cx}} (X^\top, Y^\top). \tag{1.4.1}$$

Proof: Let U, $U^\perp$, and $U^\top$ be uniform (0,1) random variables such that U and X are independent, $U^\perp$ and $X^\perp$ are independent, and $U^\top$ and $X^\top$ are independent. Let G, $G^\perp$, and $G^\top$ be functions such that

$$Y = G(X, U), \quad Y^\perp = G^\perp(X^\perp, U^\perp), \quad \text{and} \quad Y^\top = G^\top(X^\top, U^\top).$$

In order to prove the first inequality in Eq. (1.4.1) we need to show that

$$\mathrm{E}[Y \big| U] \leq_{\text{cx}} \mathrm{E}[Y^\perp \big| U^\perp]. \tag{1.4.2}$$

First we show that the two random variables in Eq. (1.4.2) have the same expectation. Note that because of the independence of $X^\perp$ and $Y^\perp$ we see that $Y^\perp$ functionally depends on $U^\perp$, that is,

$$Y^\perp = G^\perp(X^\perp, U^\perp) = h(U^\perp) \tag{1.4.3}$$

for a certain function h. Then

$$[Y^\perp \big| U^\perp = u] = h(u), \quad u \in (0, 1),$$

and hence,

$$\mathrm{E}[Y^{\perp}\big|U^{\perp}] = h(U^{\perp}). \tag{1.4.4}$$

Thus

$$\mathrm{E}[\mathrm{E}[Y^{\perp}\big|U^{\perp}]] = \mathrm{E}[h(U^{\perp})] = \mathrm{E}[Y^{\perp}] = \mathrm{E}[Y] = \mathrm{E}[\mathrm{E}[Y\big|U]],$$

where the first equality follows from Eq. (1.4.4) and the second equality follows from Eq. (1.4.3). Thus, the two variables in Eq. (1.4.2) have the same expected value.

Now, let ϕ be a convex function. Then,

$$\mathrm{E}[\phi(\mathrm{E}[Y\big|U])] \le \mathrm{E}[\mathrm{E}[\phi(Y)\big|U]] = \mathrm{E}[\phi(Y)] = \mathrm{E}[\phi(Y^{\perp})] = \mathrm{E}[\phi(\mathrm{E}[Y^{\perp}\big|U^{\perp}])],$$

where the inequality follows from Jensen's Inequality.

In order to prove the second inequality in Eq. (1.4.1) we need to show that

$$\mathrm{E}[Y^{\top}\big|U^{\top}] \le_{\mathrm{cx}} \mathrm{E}[Y\big|U]. \tag{1.4.5}$$

Since $Y^{\top}$ functionally depends on $X^{\top}$, we have that $Y^{\top}$ and $U^{\top}$ are independent, and therefore $\mathrm{E}[Y^{\top}\big|U^{\top}]$ is degenerate at $\mathrm{E}[Y^{\top}]$. Since every random variable is larger, in the order $\le_{\mathrm{cx}}$, than a random variable that is degenerate at its mean, we obtain Eq. (1.4.5), and this gives the second inequality in Eq. (1.4.1). ■

The functions k and $\tilde{k}$ of Eq. (1.3.9) make up a basic tool in the study of GDO3s. So first we point out a useful property of these functions. Let U and Y be as in Eq. (1.3.3). Abusing notation we write it here as $Y = F^{-1}_{Y|X}(U)$ [actually the Y here is not the Y in Eq. (1.3.3), but it is a random variable that is stochastically equal to Y], where U and X are independent random variables. Therefore

$$[Y\big|U = u] \overset{\mathrm{st}}{=} G(X, u) = F^{-1}_{Y|X}(u), \quad u \in (0, 1). \tag{1.4.6}$$

Obviously, for every $x \in \mathrm{support}(X)$, we have that

$$F^{-1}_{Y|X=x}(u_1) \le F^{-1}_{Y|X=x}(u_2), \quad \text{for all } u_1 \le u_2 \text{ in } (0, 1).$$

Unconditioning this with respect to X we see that

$$[Y\big|U = u_1] \overset{\mathrm{st}}{=} F^{-1}_{Y|X}(u_1) \le_{\mathrm{st}} F^{-1}_{Y|X}(u_2) \overset{\mathrm{st}}{=} [Y\big|U = u_2], \quad \text{for all } u_1 \le u_2 \text{ in } (0, 1).$$

We thus have proved the following result. Recall the definition of PRD that was given before Proposition 1.3.4.

Proposition 1.4.4. *Let U and Y be as above. Then (U,Y) is* PRD.

If $[Y\big|U=u]$ stochastically increases in u then $k(u) = \mathrm{E}[Y\big|U=u]$ increases in u. Thus we have the following corollary.

Corollary 1.4.5. *The functions k and $\tilde{k}$ in Eq. (1.3.9) are increasing.*

Remark 1.4.6. At the first glance at Proposition 1.4.4 one may wonder why the "noise" U affects Y in a positive manner, that is, monotonically increasing. After all, the "noise" indeed should have an effect on Y, but the effect of a "noise" usually should be unpredictable and at least not monotone. Some reflection, however, clarifies this seemingly not intuitive observation. Note that we could replace U in Eqs. (1.3.1) and (1.3.3) by, say, $1-U$ [or, more generally, by $h(U)$, where h is any one-to-one function from (0,1) onto (0,1) such that $h(U)$ is a uniform (0,1) random variable], and this would not have affected the distributions of $k(U)$ and of $d(U)$. Similarly, we could replace $\tilde{U}$ in Eq. (1.3.9) and in Eq. (1.3.10) by $1-\tilde{U}$, etc., without affecting the distributions of $\tilde{k}(\tilde{U})$ and of $\tilde{d}(\tilde{U})$. So, the choice that we made [i.e., using U in Eqs. (1.3.1) and (1.3.3)] can be considered to be just a convenient choice that simplifies the presentation, but that does not affect the generality. And, with this choice, Proposition 1.4.4 shows that (U,Y) happens to be PRD—this is a nice and useful observation that does not restrict the generality or the intuition. A consequence of it (Corollary 1.4.5) is that k and $\tilde{k}$ in Eq. (1.3.9) are increasing, and that will be used in the sequel.

We proceed now to a lemma that gives a sufficient condition for the order $\leq_{\text{GDO}_3\text{-cx}}$. The condition (1.4.7) below is somewhat technical and may not be easy to verify. However, we use this lemma in the sequel in order to derive nicer conditions for the order $\leq_{\text{GDO}_3\text{-cx}}$ in terms of the corresponding copulas (see Theorems 1.4.8 and 1.4.9). For the purpose of stating the next result we recall that (U,Y) is said to be smaller than $(\tilde{U},\tilde{Y})$ in the positive quadrant dependence stochastic order (denoted as $(U,Y) \leq_{\text{PQD}} (\tilde{U},\tilde{Y})$) if (U,Y) and $(\tilde{U},\tilde{Y})$ have the same marginal distributions and if

$$\mathrm{P}\{U \leq u, Y \leq y\} \leq \mathrm{P}\{\tilde{U} \leq u, \tilde{Y} \leq y\} \quad \text{for any } u \text{ and } y.$$

The PQD order is a positive dependence stochastic order that is studied, for example, in Shaked and Shanthikumar [426, Sect. 9.A] and in references that are given there. The functions k and $\tilde{k}$ in the next proposition are defined in Eq. (1.3.9). From Corollary 1.4.5 we know that they are increasing. So the assumption in the next result that they are strictly increasing is not a particularly restrictive one. It is worthwhile to point out that the strict monotonicity of $k(u)$, or more explicitly of $F^{-1}_{Y|X=x}(u)$, in u, depends on both the marginal distributions of (X, Y) and on the dependence structure that is formalized by the corresponding copula; see the discussion in Example 1.4.15 below.

Lemma 1.4.7. *Let X and $\tilde{X}$ be two random variables, and let U and $\tilde{U}$ be two independent uniform* (0,1) *random variables. Let Y and $\tilde{Y}$ be as defined in Eq.* (1.3.8)*. Suppose that the corresponding functions k and $\tilde{k}$, defined in Eq.* (1.3.9)*, are strictly increasing. If*

$$(U, Y) \geq_{\mathrm{PQD}} (\tilde{U}, \tilde{Y}), \tag{1.4.7}$$

then

$$(X, Y) \leq_{\text{GDO3-cx}} (\tilde{X}, \tilde{Y}). \tag{1.4.8}$$

Proof: Let L denote the distribution function of $k(U)$. Then

$$L(t) = \mathrm{P}\{k(U) \leq t\} = \mathrm{P}\{U \leq k^{-1}(t)\} = k^{-1}(t), \quad \text{for all } t,$$

where the second equality follows from the strict monotonicity of k, and the last equality follows from the fact that U is uniform (0,1). Hence

$$L^{-1}(u) = k(u), \quad u \in (0, 1). \tag{1.4.9}$$

Similarly, if $\tilde{L}$ denotes the distribution function of $\tilde{k}(\tilde{U})$, then

$$\tilde{L}^{-1}(u) = \tilde{k}(u), \quad u \in (0, 1). \tag{1.4.10}$$

Now note (see Shaked and Shanthikumar [426], page 389) that assumption (1.4.7) implies that

$$\mathrm{E}[Y \big| U > u] \geq \mathrm{E}[\tilde{Y} \big| \tilde{U} > u] \quad \text{for all } u \in (0, 1);$$

that is,

$$\int_u^1 \mathrm{E}[Y \big| U = p]\,\mathrm{d}p \geq \int_u^1 \mathrm{E}[\tilde{Y} \big| \tilde{U} = p]\,\mathrm{d}p \quad \text{for all } u \in (0, 1),$$

that is,

$$\int_u^1 k(p)\,\mathrm{d}p \geq \int_u^1 \tilde{k}(p)\,\mathrm{d}p \quad \text{for all } u \in (0, 1);$$

which, by Eqs. (1.4.9) and (1.4.10), is equivalent to

$$\int_u^1 L^{-1}(p)\,\mathrm{d}p \geq \int_u^1 \tilde{L}^{-1}(p)\,\mathrm{d}p \quad \text{for all } u \in (0,1). \tag{1.4.11}$$

Now, by Lemma 2.1 of Fagiuoli et al. [154] (see also (3.A.15) in Shaked and Shanthikumar [426]), we see that the inequality (1.4.11) is equivalent to $k(U) \geq_{\mathrm{cx}} \tilde{k}(\tilde{U})$, which is just Eq. (1.4.8). ■

With the help of Lemma 1.4.7 we can now derive the following sufficient conditions for the relation $(X,Y) \leq_{\mathrm{GDO_3\text{-}cx}} (\tilde{X},\tilde{Y})$. As we argued earlier, the assumption in the results below that k and $\tilde{k}$ are strictly increasing is not a particularly restrictive one.

Theorem 1.4.8. *Let (X,Y) and $(\tilde{X},\tilde{Y})$ be two random vectors such that $X \overset{\mathrm{st}}{=} \tilde{X}$ and $Y \overset{\mathrm{st}}{=} \tilde{Y}$. Suppose that the corresponding functions k and $\tilde{k}$, defined in Eq. (1.3.9), are strictly increasing. Let $C_{(X,Y)}$ and $C_{(\tilde{X},\tilde{Y})}$ be the corresponding copulas. If*

$$\min\Big\{u, \frac{\partial}{\partial p} C_{(X,Y)}(p,v)\Big\} \geq \min\Big\{u, \frac{\partial}{\partial p} C_{(\tilde{X},\tilde{Y})}(p,v)\Big\}, \tag{1.4.12}$$

for all $p, v, u \in (0,1)$, then

$$(X,Y) \leq_{\mathrm{GDO_3\text{-}cx}} (\tilde{X},\tilde{Y}).$$

Proof: Let U and $\tilde{U}$ be as in Lemma 1.4.7. Using that lemma we only need to prove that $(U,Y) \geq_{\mathrm{PQD}} (\tilde{U},\tilde{Y})$. But, using Proposition 1.3.3, we see that

$$\begin{aligned} C_{(U,Y)}(u,v) &= \int_0^1 \min\Big\{u, \frac{\partial}{\partial p} C_{(X,Y)}(p,v)\Big\}\mathrm{d}p \\ &\geq \int_0^1 \min\Big\{u, \frac{\partial}{\partial p} C_{(\tilde{X},\tilde{Y})}(p,v)\Big\}\mathrm{d}p \\ &= C_{(\tilde{U},\tilde{Y})}(u,v), \end{aligned}$$

where the inequality follows from the assumption (1.4.12). ■

If in Theorem 1.4.8 we denote by (U_1,U_2) a random vector that is distributed according to $C_{(X,Y)}$ and by $(\tilde{U}_1,\tilde{U}_2)$ a random vector that is distributed according to $C_{(\tilde{X},\tilde{Y})}$ and if we denote by $f_{[U_1|U_2 \leq v]}$ (respectively, $f_{[\tilde{U}_1|\tilde{U}_2 \leq v]}$) the density function of $[U_1 \big| U_2 \leq v]$ (respectively, $[\tilde{U}_1 \big| U_2 \leq v]$), then the condition (1.4.12) can be written as

$$\min\left\{u, vf_{[U_1|U_2\le v]}(p)\right\} \ge \min\left\{u, vf_{[\tilde U_1|\tilde U_2\le v]}(p)\right\} \quad \text{for all } p,v,u\in(0,1).$$

If, in some applications, the condition (1.4.12) in Theorem 1.4.8 is not easy to verify, then the following result may be useful—see an application of it in Example 1.4.15 below.

Theorem 1.4.9. *Let* (X,Y) *and* $(\tilde X,\tilde Y)$ *be two* PRD *random vectors such that* $X \overset{\text{st}}{=} \tilde X$ *and* $Y \overset{\text{st}}{=} \tilde Y$*. Suppose that the corresponding functions* k *and* $\tilde k$*, defined in Eq.* (1.3.9)*, are strictly increasing. Let* $C_{(X,Y)}$ *and* $C_{(\tilde X,\tilde Y)}$ *be the corresponding copulas, and let* $z(w,v)$ *be as defined in Proposition* 1.3.4*. Similarly, let* $\tilde z(w,v)$ *be defined by*

$$\tilde z(w,v) = \begin{cases} \sup\{p : \frac{\partial}{\partial p} C_{(\tilde X,\tilde Y)}(p,v) \ge w\}, & \textit{if the supremum is well defined;}\\ 0, & \textit{otherwise.}\end{cases}$$

If for every $v\in[0,1]$ *the difference*

$$z(w,v)-\tilde z(w,v) \quad \textit{has only one sign change from + to − as } w \textit{ ranges from } 0 \textit{ to } 1, \tag{1.4.13}$$

then

$$(X,Y) \le_{\text{GDO}_3\text{-cx}} (\tilde X,\tilde Y).$$

Proof: Let U and $\tilde U$ be as in Lemma 1.4.7. Using that lemma we only need to prove that $(U,Y) \ge_{\text{PQD}} (\tilde U,\tilde Y)$. Using Proposition 1.3.4 we have

$$C_{(U,Y)}(u,v) - C_{(\tilde U,\tilde Y)}(u,v) = \int_0^u [z(w,v)-\tilde z(w,v)]\,\mathrm{d}w$$

Fix a $v\in(0,1)$. Under assumption (1.4.13), let c be a point of crossing. Denote $J_1=(-\infty,c]$ and $J_2=(c,\infty)$. Then $z(w,v)-\tilde z(w,v)\ge 0$ on J_1 and $z(w,v)-\tilde z(w,v)\le 0$ on J_2. Clearly

$$\lim_{u\to 1}\int_0^u [z(w,v)-\tilde z(w,v)]\,\mathrm{d}u = v-v=0,$$

and

$$\lim_{u\to 0}\int_0^u [z(w,v)-\tilde z(w,v)]\,\mathrm{d}u = 0.$$

Combining these observations shows that $C_{(U,Y)}(u,v) \ge C_{(\tilde U,\tilde Y)}(u,v)$ holds. ∎

In the next result we identify another sufficient condition for the order $\le_{\text{GDO}_3\text{-cx}}$.

Proposition 1.4.10. *Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be two random vectors such that $X \stackrel{\text{st}}{=} \tilde{X}$ and $Y \stackrel{\text{st}}{=} \tilde{Y}$. Suppose that the corresponding functions k and $\tilde{k}$, defined in Eq. (1.3.9), are strictly increasing. If*

$$[Y \big| X = x] \geq_{\text{cx}} \big[\tilde{Y} \big| \tilde{X} = x\big], \quad x \in \text{support}(X), \tag{1.4.14}$$

then $(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$.

Proof: Let U be a uniform (0,1) random variable that is independent of X, and express Y as in Eq. (1.3.8). Let L denote the distribution function of $k(U)$, where k is defined in Eq. (1.3.9). Then, by Eq. (1.4.9), we have $L^{-1}(u) = k(u)$ for $u \in (0, 1)$. Similarly, if $\tilde{L}$ denotes the distribution function of $\tilde{k}(\tilde{U})$, then $\tilde{L}^{-1}(u) = \tilde{k}(u)$for $u \in (0, 1)$.

Denote by F_X the marginal distribution function of X (which is also the marginal distribution function of $\tilde{X}$). From the characterization of the convex order given in Lemma 2.1 of Fagiuoli et al. [154] (or see (3.A.15) in Shaked and Shanthikumar [426]) we have

$$\begin{aligned}
\int_p^1 L^{-1}(u)\,\mathrm{d}u &= \int_p^1 k(u)\,\mathrm{d}u \\
&= \int_p^1 \mathrm{E}[Y \big| U = u]\,\mathrm{d}u \\
&= \int_p^1 \int_{x \in \text{support}(X)} F^{-1}_{Y|X=x}(u)\,\mathrm{d}F_X(x)\mathrm{d}u \\
&= \int_{x \in \text{support}(X)} \int_p^1 F^{-1}_{Y|X=x}(u)\,\mathrm{d}u\mathrm{d}F_X(x) \\
&\geq \int_{x \in \text{support}(X)} \int_p^1 F^{-1}_{\tilde{Y}|\tilde{X}=x}(u)\,\mathrm{d}u\mathrm{d}F_X(x) \\
&= \int_p^1 \mathrm{E}[\tilde{Y} \big| \tilde{U} = u]\,\mathrm{d}u \\
&= \int_p^1 \tilde{L}^{-1}(u)\,\mathrm{d}u
\end{aligned}$$

for every $p \in (0, 1)$, where the third equality follows from Eq. (1.4.6) and the inequality from Eq. (1.4.14). Using Lemma 2.1 of Fagiuoli et al. [154] once more we see that the above inequality yields $k(U) \geq_{\text{cx}} \tilde{k}(\tilde{U})$ which gives the stated result. ■

It is worthwhile to note that the condition (1.4.14) was shown in Shaked et al. [428] to imply $(X,Y) =_{\text{GDO}_1\text{-cx}} (\tilde{X},\tilde{Y})$ and $(X,Y) =_{\text{GDO}_2\text{-st}} (\tilde{X},\tilde{Y})$.

Although in this subsection we study the order $\le_{\text{GDO}_3\text{-cx}}$, we think that it is useful to make a short digression and show how the main tool in the proof of Proposition 1.4.10 can also yield a sufficient condition for the order $\le_{\text{GDO}_3\text{-disp}}$.

Proposition 1.4.11. *Let (X,Y) and $(\tilde{X},\tilde{Y})$ be two random vectors such that $X \overset{\text{st}}{=} \tilde{X}$ and $Y \overset{\text{st}}{=} \tilde{Y}$. Suppose that the corresponding functions k and $\tilde{k}$, defined in Eq. (1.3.9), are strictly increasing. If*

$$[Y\big|X = x] \ge_{\text{disp}} \big[\tilde{Y}\big|\tilde{X} = x\big], \quad x \in \text{support}(X), \tag{1.4.15}$$

then $(X,Y) \le_{\text{GDO}_3\text{-disp}} (\tilde{X},\tilde{Y})$.

Proof: We use here the notation in the proof of Proposition 1.4.10. For $0 < \alpha < \beta < 1$ we have

$$\begin{aligned}
L^{-1}(\beta) - L^{-1}(\alpha) &= k(\beta) - k(\alpha) \\
&= \mathrm{E}[Y\big|U = \beta] - \mathrm{E}[Y\big|U = \alpha] \\
&= \int_{x\in\text{support}(X)} [F^{-1}_{Y|X=x}(\beta) - F^{-1}_{Y|X=x}(\alpha)]\, \mathrm{d}F_X(x) \\
&\ge \int_{x\in\text{support}(X)} [F^{-1}_{\tilde{Y}|\tilde{X}=x}(\beta) - F^{-1}_{\tilde{Y}|\tilde{X}=x}(\alpha)]\, \mathrm{d}F_X(x) \\
&= \tilde{L}^{-1}(\beta) - \tilde{L}^{-1}(\alpha),
\end{aligned}$$

here the above inequality follows from Eq. (1.4.15). This gives the stated result. ■

It is worthwhile to note that the condition (1.4.15) was shown in Shaked et al. [428, Proposition 5.7] to imply $(X,Y) \le_{\text{GDO}_2\text{-st}} (\tilde{X},\tilde{Y})$. In fact, there is a mistype there—the stated conclusion there incorrectly says $(X,Y) \ge_{\text{GDO}_2\text{-st}} (\tilde{X},\tilde{Y})$ rather than $(X,Y) \le_{\text{GDO}_2\text{-st}} (\tilde{X},\tilde{Y})$.

We note that Propositions 1.4.10 and 1.4.11 go along with the intuition in the sense that if $[Y\big|X = x]$ is more variable than $[\tilde{Y}\big|\tilde{X} = x]$ for every $x \in \text{support}(X)$, then, intuitively, $\tilde{X}$ is a "better predictor" of $\tilde{Y}$ than X is of Y. Therefore we would expect Y to be "more independent" of X than $\tilde{Y}$ is of $\tilde{X}$.

As in Shaked et al. [428], we note that in the terminology of Belzunce et al. [56], condition (1.4.15) can be denoted as $(X, Y) \geq_{\text{c-disp}} (\tilde{X}, \tilde{Y})$. That is, if (X, Y) and $(\tilde{X}, \tilde{Y})$ have the same marginals, then Proposition 1.4.11 proves that

$$(X, Y) \geq_{\text{c-disp}} (\tilde{X}, \tilde{Y}) \Longrightarrow (X, Y) \leq_{\text{GDO}_3\text{-disp}} (\tilde{X}, \tilde{Y}).$$

The order $\geq_{\text{c-disp}}$ is called the multivariate conditional dispersion order in Belzunce et al. [56]. It is quite weak in the sense that it is implied by other common multivariate dispersion orders. Thus, Proposition 1.4.11 actually indicates many instances in which the order $\leq_{\text{GDO}_3\text{-disp}}$ holds.

Let (X, Y) be a random vector. Recall the definition of the function G in Eqs. (1.3.2) and (1.3.3). Wu and Mielniczuk [476, Sect. 2.1] defined the following global dependence measures. Let X' be an independent copy of X, and assume that both X and X' are independent of the uniform (0,1) random variable U. Now, let $Y = G(X, U) \in \mathscr{L}^p$, $p > 0$, and let $Y' = G(X', U)$. Define

$$\delta_p(X, Y) = \|Y - Y'\|_p \tag{1.4.16}$$

and, for $p \geq 1$, define

$$\tau_p(X, Y) = \|Y - \mathrm{E}[Y \big| U]\|_p. \tag{1.4.17}$$

A particular important case is $p = 2$. In this case,

$$\delta_2(X, Y) = \sqrt{2}\,\tau_2(X, Y) \tag{1.4.18}$$

(see Wu and Mielniczuk [476], page 128).

In relation to the measures $\delta_2(X, Y)$ and $\tau_2(X, Y)$, the order $\leq_{\text{GDO}_3\text{-cx}}$ has the following sensible property.

Theorem 1.4.12. *If* $(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$ *then*

$$\tau_2(X, Y) \leq \tau_2(\tilde{X}, \tilde{Y}) \quad \textit{and} \quad \delta_2(X, Y) \leq \delta_2(\tilde{X}, \tilde{Y}).$$

Proof: From Eq. (1.4.18) it is sufficient to prove the result for τ_2. By its definition,

$$\tau_2(X, Y) = \mathrm{E}\big[Y - \mathrm{E}\big[Y \big| U\big]\big]^2 = \mathrm{E}\Big[\mathrm{E}\big[Y - \mathrm{E}\big[Y \big| U\big]\big]^2 \Big| U\Big] = \mathrm{E}[\mathrm{Var}[Y \big| U]]. \tag{1.4.19}$$

Now, the assumption $(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$ means

$$\mathrm{E}[Y|U] \geq_{\mathrm{cx}} \mathrm{E}\big[\tilde{Y}|\tilde{U}\big],$$

which implies

$$\mathrm{Var}\big[\mathrm{E}[Y|U]\big] \geq \mathrm{Var}\big[\mathrm{E}\big[\tilde{Y}|\tilde{U}\big]\big]. \tag{1.4.20}$$

Since

$$\begin{aligned}\mathrm{Var}[Y] &= \mathrm{Var}[\mathrm{E}[Y|U]] + \mathrm{E}[\mathrm{Var}[Y|U]],\\ \mathrm{Var}[\tilde{Y}] &= \mathrm{Var}\big[\mathrm{E}\big[\tilde{Y}|\tilde{U}\big]\big] + \mathrm{E}\big[\mathrm{Var}\big[\tilde{Y}|\tilde{U}\big]\big],\end{aligned}$$

and

$$\mathrm{Var}[Y] = \mathrm{Var}[\tilde{Y}],$$

it follows from Eq. (1.4.20) that

$$\mathrm{E}[\mathrm{Var}[Y|U]] \leq \mathrm{E}\big[\mathrm{Var}\big[\tilde{Y}|\tilde{U}\big]\big],$$

which implies $\tau_2(X, Y) \leq \tau_2(\tilde{X}, \tilde{Y})$. ∎

Some examples of random vectors that are ordered with respect to the order $\leq_{\text{GDO}_3\text{-cx}}$ will now be described.

Example 1.4.13. In this example we compare two random vectors that are expressed in the form that is reminiscent of Eq. (1.2.8). But, we note that Eq. (1.2.8) describes any (i.e., general) bivariate random vector, whereas Eq. (1.4.21) below is a special case of Eq. (1.2.8). It is a special case in the sense that Z in Eq. (1.2.8) need not be independent of X (though they are uncorrelated), whereas Z in Eq. (1.4.21) is independent of X.

So, let (X, Y) and $(\tilde{X}, \tilde{Y})$ be two random vectors with the same marginals. Following Example 2.1 of Wu and Mielniczuk [476], consider the heteroscedastic regression model for these vectors:

$$Y = m(X) + \sigma(X)\,Z, \tag{1.4.21}$$

where X and Z are independent, $\mathrm{E}[Z] = 0$, and $\sigma(\cdot) > 0$, and

$$\tilde{Y} = \tilde{m}(\tilde{X}) + \sigma(\tilde{X})\,\tilde{Z}, \tag{1.4.22}$$

where $\tilde{X}$ and $\tilde{Z}$ are independent and $\mathrm{E}\big[\tilde{Z}\big] = 0$. Note that since $\mathrm{E}[Z] = 0$ and $\mathrm{E}\big[\tilde{Z}\big] = 0$ we have that m and $\tilde{m}$ in Eqs. (1.4.21) and

(1.4.22) are indeed the same m and $\tilde{m}$ as in Eq. (1.2.1). But σ in Eqs. (1.4.21) and (1.4.22) is neither e nor $\tilde{e}$ from Eq. (1.2.2). We also note that

$$\mathrm{E}[Y] = \mathrm{E}[m(X)] = \mathrm{E}\big[\tilde{m}(\tilde{X})\big] = \mathrm{E}\big[\tilde{Y}\big]. \tag{1.4.23}$$

Denoting the distribution function of Z by F_Z, we have [see Eq. (1.3.2)]

$$G(x,u) = F^{-1}_{Y|X=x}(u) = m(x) + \sigma(x)\, F_Z^{-1}(u), \quad u \in (0,1).$$

Thus,

$$G(X,U) = m(X) + \sigma(X)\, F_Z^{-1}(U),$$

where U is a uniform (0,1) random variable independent of X. We note that

$$\mathrm{E}[Y \big| U = u] = \mathrm{E}[G(X,U) \big| U = u] = \mathrm{E}[m(X)] + \mathrm{E}[\sigma(X)]\, F_Z^{-1}(u),$$

and therefore

$$\mathrm{E}[Y \big| U] = \mathrm{E}[m(X)] + \mathrm{E}[\sigma(X)]\, F_Z^{-1}(U). \tag{1.4.24}$$

Similarly, we get

$$\mathrm{E}\big[\tilde{Y} \big| \tilde{U}\big] = \mathrm{E}\big[\tilde{m}(\tilde{X})\big] + \mathrm{E}[\sigma(\tilde{X})]\, F_{\tilde{Z}}^{-1}(\tilde{U}), \tag{1.4.25}$$

where $\tilde{U}$ is a uniform (0,1) random variable independent of $\tilde{X}$.

Now, from Eqs. (1.4.23)–(1.4.25), we see that

$$(X,Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X},\tilde{Y}) \tag{1.4.26}$$

if, and only if,

$$\mathrm{E}[\sigma(X)]\, F_Z^{-1}(U) \geq_{\text{cx}} \mathrm{E}\big[\sigma(\tilde{X})\big]\, F_{\tilde{Z}}^{-1}(\tilde{U}).$$

Taking into account that $\mathrm{E}[\sigma(X)] = \mathrm{E}\big[\sigma(\tilde{X})\big]$, we see that Eq. (1.4.26) holds if, and only if,

$$F_Z^{-1}(U) \geq_{\text{cx}} F_{\tilde{Z}}^{-1}(\tilde{U}),$$

or, equivalently, if, and only if,

$$Z \geq_{\text{cx}} \tilde{Z}. \tag{1.4.27}$$

Indeed, under this regression model (1.4.21) and (1.4.22), condition (1.4.27) is quite intuitive for (X,Y) to be less globally dependent than $(\tilde{X},\tilde{Y})$. ■

Example 1.4.14. Let (X, Y) be a bivariate normal random vector with zero means, unit variances, and correlation coefficient ρ. Then the distribution of $[Y \big| X = x]$ is normal with mean ρx and variance $\sqrt{1-\rho^2}$. Explicitly it is given by

$$F_{Y|X=x}(y) = \Phi\left(\frac{y - \rho x}{\sqrt{1-\rho^2}}\right), \quad y \in \mathbb{R},$$

where Φ denotes the univariate standard normal distribution. It follows, using the notation in Eq. (1.3.2), that

$$G(x, u) = \rho x + \sqrt{1-\rho^2} \cdot \Phi^{-1}(u),$$

and thus, following Eq. (1.3.3),

$$Y \stackrel{\text{st}}{=} \rho X + \sqrt{1-\rho^2} \cdot \Phi^{-1}(U), \tag{1.4.28}$$

where X is a standard normal random variable and U is a uniform (0,1) random variable, independent of X.

Similarly, let $(\tilde{X}, \tilde{Y})$ be a bivariate normal random vector with zero means, unit variances, and correlation coefficient $\tilde{\rho}$. As above, we have that

$$\tilde{Y} \stackrel{\text{st}}{=} \tilde{\rho}\tilde{X} + \sqrt{1-\tilde{\rho}^2} \cdot \Phi^{-1}(\tilde{U}), \tag{1.4.29}$$

where $\tilde{X}$ is a standard normal random variable, and $\tilde{U}$ is a uniform (0,1) random variable, independent of $\tilde{X}$.

Now we compute

$$\mathrm{E}[Y \big| U] = \sqrt{1-\rho^2} \cdot \Phi^{-1}(U) \quad \text{and} \quad \mathrm{E}[\tilde{Y} \big| \tilde{U}] = \sqrt{1-\tilde{\rho}^2} \cdot \Phi^{-1}(\tilde{U}).$$

It is easy to see that

$$\mathrm{E}[Y \big| U] \geq_{\text{cx}} \mathrm{E}[\tilde{Y} \big| \tilde{U}] \Longleftrightarrow |\rho| \leq |\tilde{\rho}|,$$

that is,

$$(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y}) \Longleftrightarrow |\rho| \leq |\tilde{\rho}|.$$

It is of interest to note that in comparing two bivariate normal random vectors, the above equivalence is similar to the equivalences

$$(X, Y) \leq_{\text{GDO}_1\text{-cx}} (\tilde{X}, \tilde{Y}) \Longleftrightarrow |\rho| \leq |\tilde{\rho}|$$

and

$$(X, Y) \leq_{\text{GDO}_1\text{-disp}} (\tilde{X}, \tilde{Y}) \Longleftrightarrow |\rho| \leq |\tilde{\rho}|$$

that were obtained in Shaked, Sordo, and Suárez-Llorens [428]. ■

Example 1.4.15. Let (X, Y) be a random vector with absolutely continuous marginal distributions and with dependence structure according to the Farlie-Gumbel-Morgenstern copula given by

$$C_\theta(u, v) = uv + \theta uv(1-u)(1-v), \quad (u, v) \in [0, 1]^2,$$

where $\theta \in [-1, 1]$ is a parameter that governs the strength and the sign of the dependence. Suppose that $\theta \in (0, 1]$. A straightforward computation yields that (X, Y) is PRD. By Proposition 1.4.4 we know that (U, Y) is always PRD, and due to the fact that $F_{Y|X=x}(\cdot)$ is absolutely continuous we have that $k(u)$ is strictly increasing. We are now going to use Theorem 1.4.9. For this purpose we compute $\frac{\partial}{\partial p} C_\theta(p, v) = v + \theta v(1-v)(1-2p)$ and obtain

$$z_\theta(w, v) = \begin{cases} 1, & \text{if } w \le v - \theta v(1-v); \\ \frac{1}{2} - \frac{w-v}{2\theta v(1-v)}, & \text{if } v - \theta v(1-v) \le w \le v + \theta v(1-v); \\ 0, & \text{if } w \ge v + \theta v(1-v). \end{cases}$$

Let $(\tilde{X}, \tilde{Y})$ be another random vector with the same marginals as (X, Y) and having a Farlie-Gumbel-Morgenstern copula with parameter $\tilde{\theta}$. If $\theta < \tilde{\theta}$, fixing v, it is apparent, by just plotting the function $z_\theta(\cdot, v) - \tilde{z}_{\tilde{\theta}}(\cdot, v)$, that Eq. (1.4.13) holds. Hence it follows from Theorem 1.4.9 that if $\theta < \tilde{\theta}$ then $(X, Y) \le_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$. ■

The following example is inspired by Example 2.4 of Wu and Mielniczuk [476]. Unlike Wu and Mielniczuk, we use below the classical standard construction as discussed in Remark 1.3.2. As is shown in Proposition 1.4.4 and explained in Remark 1.4.6, this produces positive dependence between U and Y in Eq. (1.3.8). This leads to a technical dissimilarity between the following example and Example 2.4 in [476], that is, the U in Wu and Mielniczuk [476] is $1 - U$ below.

Example 1.4.16. Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be random vectors with the following discrete joint probability mass functions:

$X \backslash Y$	0	1	
0	p_{00}	p_{01}	$p_{0\cdot}$
1	p_{10}	p_{11}	$p_{1\cdot}$
	$p_{\cdot 0}$	$p_{\cdot 1}$	1

and

$\tilde{X} \backslash \tilde{Y}$	0	1	
0	$\tilde{p}_{00}$	$\tilde{p}_{01}$	$p_{0\cdot}$
1	$\tilde{p}_{10}$	$\tilde{p}_{11}$	$p_{1\cdot}$
	$p_{\cdot 0}$	$p_{\cdot 1}$	1

where the marginal probability mass function of X and $\tilde{X}$ is $(p_{0\cdot}, p_{1\cdot})$, the marginal probability mass function of Y and $\tilde{Y}$ is $(p_{\cdot 0}, p_{\cdot 1})$, and all the p_{ij}'s and $\tilde{p}_{ij}$'s are nonnegative and add up to the given marginal probabilities.

Let U be a uniform (0,1) random variable, independent of X. Then the function G in Eq. (1.3.8) is given by

$$Y = X \cdot I_{\{u \geq p_{11}/p_{1\cdot}\}}(U) + (1 - X) \cdot I_{\{u \geq p_{01}/p_{0\cdot}\}}(U), \tag{1.4.30}$$

where I denotes an indicator function. Similarly we can express $\tilde{Y}$ as

$$\tilde{Y} = \tilde{X} \cdot I_{\{u \geq p_{11}/p_{1\cdot}\}}(\tilde{U}) + (1 - \tilde{X}) \cdot I_{\{u \geq p_{01}/p_{0\cdot}\}}(\tilde{U}), \tag{1.4.31}$$

where $\tilde{X}$ and $\tilde{U}$ are independent. Consider now the following two cases:

Case 1: $\boxed{p_{00}p_{11} \geq p_{01}p_{10} \text{ and } \tilde{p}_{00}\tilde{p}_{11} \geq \tilde{p}_{01}\tilde{p}_{10}.}$ In this case X and Y are *positively dependent* (for instance, $\mathrm{Cov}(X, Y) \geq 0$) and also $\tilde{X}$ and $\tilde{Y}$ are positively dependent. Note that from the inequalities $p_{00}p_{11} \geq p_{01}p_{10}$ and $\tilde{p}_{00}\tilde{p}_{11} \geq \tilde{p}_{01}\tilde{p}_{10}$ it follows that $\frac{p_{01}}{p_{0\cdot}} \leq \frac{p_{11}}{p_{1\cdot}}$ and $\frac{\tilde{p}_{01}}{p_{0\cdot}} \leq \frac{\tilde{p}_{11}}{p_{1\cdot}}$. Then from Eq. (1.4.30) we have

$$\mathrm{E}[Y \big| U = u] = \begin{cases} 0, & \text{if } u \in [0, \frac{p_{01}}{p_{0\cdot}}]; \\ \mathrm{E}[1 - X], & \text{if } u \in [\frac{p_{01}}{p_{0\cdot}}, \frac{p_{11}}{p_{1\cdot}}]; \\ 1, & \text{if } u \in [\frac{p_{11}}{p_{1\cdot}}, 1]. \end{cases}$$

Hence

$$\mathrm{E}[Y \big| U] = \begin{cases} 0, & \text{with probability } \frac{p_{01}}{p_{0\cdot}}; \\ 1 - p_{1\cdot}, & \text{with probability } \frac{p_{11}}{p_{1\cdot}} - \frac{p_{01}}{p_{0\cdot}}; \\ 1, & \text{with probability } 1 - \frac{p_{11}}{p_{1\cdot}}. \end{cases}$$

Similarly,

$$\mathrm{E}[\tilde{Y} \big| \tilde{U}] = \begin{cases} 0, & \text{with probability } \frac{\tilde{p}_{01}}{p_{0\cdot}}; \\ 1 - p_{1\cdot}, & \text{with probability } \frac{\tilde{p}_{11}}{p_{1\cdot}} - \frac{\tilde{p}_{01}}{p_{0\cdot}}; \\ 1, & \text{with probability } 1 - \frac{\tilde{p}_{11}}{p_{1\cdot}}. \end{cases}$$

Now, $\mathrm{E}[Y \big| U] \geq_{\mathrm{cx}} \mathrm{E}[\tilde{Y} \big| \tilde{U}]$ if, and only if, $p_{11} \leq \tilde{p}_{11}$. Thus, we see that in Case 1, $(X, Y) \leq_{\mathrm{GDO_3\text{-}cx}} (\tilde{X}, \tilde{Y})$ if, and only if, $(\tilde{X}, \tilde{Y})$ is, informally, "more positively dependent" than (X, Y).

Case 2: $\boxed{p_{00}p_{11} \leq p_{01}p_{10} \text{ and } \tilde{p}_{00}\tilde{p}_{11} \leq \tilde{p}_{01}\tilde{p}_{10}.}$ In this case X and Y are *negatively dependent* (for instance, $\mathrm{Cov}(X,Y) \leq 0$) and also $\tilde{X}$ and $\tilde{Y}$ are negatively dependent. Note that from the inequalities $p_{00}p_{11} \leq p_{01}p_{10}$ and $\tilde{p}_{00}\tilde{p}_{11} \leq \tilde{p}_{01}\tilde{p}_{10}$ it follows that $\frac{p_{11}}{p_{1\cdot}} \leq \frac{p_{01}}{p_{0\cdot}}$ and $\frac{\tilde{p}_{11}}{p_{1\cdot}} \leq \frac{\tilde{p}_{01}}{p_{0\cdot}}$. A computation, similar to the one in Case 1, yields

$$\mathrm{E}[Y|U] = \begin{cases} 0, & \text{with probability } \frac{p_{11}}{p_{1\cdot}}; \\ p_{1\cdot}, & \text{with probability } \frac{p_{01}}{p_{0\cdot}} - \frac{p_{11}}{p_{1\cdot}}; \\ 1, & \text{with probability } 1 - \frac{p_{01}}{p_{0\cdot}}; \end{cases}$$

and

$$\mathrm{E}[\tilde{Y}|\tilde{U}] = \begin{cases} 0, & \text{with probability } \frac{\tilde{p}_{11}}{p_{1\cdot}}; \\ p_{1\cdot}, & \text{with probability } \frac{\tilde{p}_{01}}{p_{0\cdot}} - \frac{\tilde{p}_{11}}{p_{1\cdot}}; \\ 1, & \text{with probability } 1 - \frac{\tilde{p}_{01}}{p_{0\cdot}}. \end{cases}$$

Now, $\mathrm{E}[Y|U] \geq_{\mathrm{cx}} \mathrm{E}[\tilde{Y}|\tilde{U}]$ if, and only if, $p_{11} \geq \tilde{p}_{11}$. Thus we see that in Case 2, $(X,Y) \leq_{\mathrm{GDO_3\text{-}cx}} (\tilde{X},\tilde{Y})$ if, and only if, $(\tilde{X},\tilde{Y})$ is, informally, "more negatively dependent" than (X,Y).

Note that in the above cases we have $(X,Y) \leq_{\mathrm{GDO_3\text{-}cx}} (\tilde{X},\tilde{Y}) \Longleftrightarrow (Y,X) \leq_{\mathrm{GDO_3\text{-}cx}} (\tilde{Y},\tilde{X})$.

For the cases in which one of (X,Y) and $(\tilde{X},\tilde{Y})$ is positively dependent and the other is negatively dependent, it is possible to make a similar analysis; we do not present the details here. ■

1.4.2 The Order $\leq_{\mathbf{GDO_4\text{-}st}}$

The univariate order $\leq_{\mathrm{st}}$ is reflexive and transitive. Thus we have the following result.

Proposition 1.4.17. *The order* $\leq_{\mathrm{GDO_4\text{-}st}}$ *satisfies property* (O1).

We do not know whether $\leq_{\mathrm{GDO_4\text{-}st}}$ satisfies property (O2). But, following the comments in Remark 1.2.1 we conjecture that it does not. However, we have the following results.

Proposition 1.4.18. *The order* $\leq_{\mathrm{GDO_4\text{-}st}}$ *satisfies property* (O3).

Proof: Let (X,Y) be a random vector. Furthermore, let ϕ be a one-to-one measurable function, and consider $(\phi(X),Y)$. Let $G(x,u)$ be defined as in Eq. (1.3.2), and define G_ϕ by

$$G_\phi(x,u) = G(\phi^{-1}(x),u).$$

Let U be a uniform random variable on $[0,1]$, which is independent of X. As in the proof of Proposition 1.4.2, Y can be expressed as $Y \stackrel{\text{st}}{=} G(X,U)$ and also as $Y \stackrel{\text{st}}{=} G_\phi(\phi(X),U)$. Therefore the function $d(u)$ that is defined in Eq. (1.3.10) can be expressed both as

$$d(u) = \text{Var}[Y \big| U = u] = \text{Var}[G(X,u)]$$

and as

$$d(u) = \text{Var}[Y \big| U = u] = \text{Var}[G_\phi(\phi(X),u)].$$

In other words, both vectors (X,Y) and $(\phi(X),Y)$ share the same function $d(u)$. Similarly, if $(\tilde{X},\tilde{Y})$ is another random vector then $(\tilde{X},\tilde{Y})$ and $(\phi(\tilde{X}),\tilde{Y})$ share the same function $\tilde{d}(u)$ defined in Eq. (1.3.10). Thus, if $(X,Y) \leq_{\text{GDO}_4\text{-st}} (\tilde{X},\tilde{Y})$ then $(\phi(X),Y) \leq_{\text{GDO}_4\text{-st}} (\phi(\tilde{X}),\tilde{Y})$.

Next, it is easy to see that if $(X,Y) \leq_{\text{GDO}_4\text{-st}} (\tilde{X},\tilde{Y})$ then $(X,l(Y)) \leq_{\text{GDO}_4\text{-st}} (\tilde{X},l(\tilde{Y}))$ for every linear function l.

Property (O3) now follows from the above two observations. ■

Proposition 1.4.19. *The order* $\leq_{\text{GDO}_4\text{-st}}$ *satisfies property* (O4).

Proof: Let (X,Y) be a random vector. Recall the notation $X^\perp$ and $Y^\perp$ from Eq. (1.2.10). We want to prove that

$$(X^\perp, Y^\perp) \leq_{\text{GDO}_4\text{-st}} (X,Y)).$$

Let U and $U^\perp$ be uniform (0,1) random variables such that U and X are independent and $U^\perp$ and $X^\perp$ are independent. Let G and $G^\perp$ be functions such that

$$Y = G(X,U) \quad \text{and} \quad Y^\perp = G^\perp(X^\perp, U^\perp).$$

With this notation we want to prove that

$$\text{Var}[Y^\perp \big| U^\perp] \leq_{\text{a.s}} \text{Var}[Y \big| U]. \tag{1.4.32}$$

Since $Y^\perp$ is independent of $X^\perp$ we see that $[Y^\perp \big| U^\perp = u] \stackrel{\text{st}}{=} h(u)$ for some function h. Hence $\text{Var}[Y^\perp \big| U^\perp = u] = 0$ for all $u \in (0,1)$. Therefore $\text{Var}[Y^\perp \big| U^\perp] = 0$, whereas $\text{Var}[Y \big| U] \geq 0$, and Eq. (1.4.32) follows. ■

In the following example it is shown that the order $\leq_{\text{GDO}_4\text{-st}}$ does not have property (O5).

Example 1.4.20. Let X be a nonnegative continuous random variable with a finite variance, and let U be uniform (0,1) random variable that is independent of X. Define

$$Y = X \cdot F_Z^{-1}(U)$$

where Z is a standard normal random variable—this is a special case of heteroscedastic regression model in Eq. (1.4.21). Using the notation in Eq. (1.3.8) we have

$$G(X, U) = X \cdot F_Z^{-1}(U).$$

Denote the marginal distributions of X and Y by F_X and F_Y, respectively. A straightforward computation shows that

$$\mathrm{Var}(Y) = \mathrm{E}(X^2).$$

Now, recall the notation $X^\top$ and $Y^\top$ from Eq. (1.2.11). Explicitly, let $X^\top$ be a random variable such that $X^\top \stackrel{\text{st}}{=} X$, and define $Y^\top$ by $Y^\top = F_Y^{-1}(F_X(X^\top))$. Formally, as in Eq. (1.3.4), let $U^\top$ be a uniform (0,1) random variable that is independent of $X^\top$, and let $G^\top$ be a function such that $Y^\top = G^\top(X^\top, U^\top)$. The function $g^\top(x, u) = F_Y^{-1}(F_X(x))$ is actually independent of u. Therefore $\mathrm{Var}[Y^\top \big| U^\top] \stackrel{\text{a.s.}}{=} \mathrm{Var}[Y^\top] = \mathrm{Var}[Y] = \mathrm{E}[X^2]$. On the other hand, $\mathrm{Var}[Y \big| U = u] = \mathrm{Var}(X)(F_Z^{-1}(u))^2$ for all $u \in (0, 1)$. For u large enough near 1, we have that $\mathrm{Var}[Y \big| U = u] \geq \mathrm{E}[X^2] = \mathrm{Var}[Y]$. It follows that $\mathrm{Var}[Y \big| U] \not\leq_{\text{st}} \mathrm{Var}[Y] \stackrel{a.s.}{=} \mathrm{Var}[Y^\top \big| U^\top]$. That is, $(X, Y) \not\leq_{\text{GDO}_4\text{-st}} (X^\top, Y^\top)$. ■

Although, as Example 1.4.20 shows, it is not always true that $(X, Y) \leq_{\text{GDO}_4\text{-st}} (X^\top, Y^\top)$, a weaker stochastic inequality still holds, as is described in Proposition 1.4.21 below. We recall that for two random variables Z and $\tilde{Z}$ we denote $Z \leq_{\text{icv}} \tilde{Z}$ if $\mathrm{E}\phi(Z) \leq \mathrm{E}\phi(\tilde{Z})$ for every increasing concave function ϕ for which the above expectations are well defined. For a detailed study of the univariate order $\leq_{\text{icv}}$ see, for example, Müller and Stoyan [335] or Shaked and Shanthikumar [426]. The order $\leq_{\text{GDO}_4\text{-icv}}$, that is mentioned in Proposition 1.4.21 below, is the one obtained from Eq. (1.3.13) with $\leq_{\text{magnitude}}$ being $\leq_{\text{icv}}$.

Proposition 1.4.21. *With the notation in Eq.* (1.2.11)*, for every random vector* (X, Y)*, we have*

$$(X, Y) \leq_{\text{GDO}_4\text{-icv}} (X^\top, Y^\top). \tag{1.4.33}$$

Proof: Let U be as in Eq. (1.3.1), and let $U^{\top}$ be similarly defined in relation to $(X^{\top}, Y^{\top})$. Note that actually $U^{\top}$ is independent of $Y^{\top}$. In order to prove Eq. (1.4.33) we need to show that

$$\mathrm{Var}[Y \big| U] \le_{\mathrm{icv}} \mathrm{Var}[Y^{\perp} \big| U^{\perp}]. \tag{1.4.34}$$

Let ϕ be an increasing concave function. Then

$$\mathrm{E}[\phi(\mathrm{Var}[Y \big| U])] \le \phi(\mathrm{E}[\mathrm{Var}[Y \big| U]]) \le \phi(\mathrm{Var}[U]) = \phi(\mathrm{Var}[Y^{\perp} \big| U^{\perp}]),$$

where the first inequality follows from Jensen's Inequality, the second inequality follows from the monotonicity of ϕ and the fact that $\mathrm{E}[\mathrm{Var}[Y \big| U]] \le \mathrm{Var}[U]$ (this follows from Eq. (1.2.8) or from Eq. (2.3) in Shaked et al. [428]), and the equality follows from Eq. (1.2.11) and the independence of $Y^{\perp}$ and $U^{\perp}$. This establishes Eq. (1.4.34). ■

A combination of Propositions 1.4.19 and 1.4.21 gives the following analog of Proposition 5.4 in Shaked et al. [428].

Corollary 1.4.22. *With the notation in Eqs.* (1.2.10) *and* (1.2.11)*, for every random vector* (X, Y)*, we have*

$$(X^{\perp}, Y^{\perp}) \le_{\mathrm{GDO_4\text{-}st}} (X, Y) \le_{\mathrm{GDO_4\text{-}icv}} (X^{\top}, Y^{\top}).$$

In relation to the measures $\delta_2(X, Y)$ and $\tau_2(X, Y)$ [see Eqs. (1.4.16) and (1.4.17)], the order $\le_{\mathrm{GDO_4\text{-}st}}$ has the following sensible property.

Theorem 1.4.23. *If* $(X, Y) \le_{\mathrm{GDO_4\text{-}st}} (\tilde{X}, \tilde{Y})$ *then*

$$\tau_2(X, Y) \le \tau_2(\tilde{X}, \tilde{Y}) \quad \textit{and} \quad \delta_2(X, Y) \le \delta_2(\tilde{X}, \tilde{Y}).$$

Proof: As in the proof of Theorem 1.4.12, it is sufficient to prove the result for τ_2. The assumption $(X, Y) \le_{\mathrm{GDO_4\text{-}st}} (\tilde{X}, \tilde{Y})$ means

$$\mathrm{Var}[Y \big| U] \le_{\mathrm{st}} \mathrm{Var}\big[\tilde{Y} \big| \tilde{U}\big],$$

which implies

$$\mathrm{E}[\mathrm{Var}[Y \big| U]] \le \mathrm{E}\big[\mathrm{Var}\big[\tilde{Y} \big| \tilde{U}\big]\big].$$

The inequality $\tau_2(X, Y) \le \tau_2(\tilde{X}, \tilde{Y})$ now follows from Eq. (1.4.19). ■

Example 1.4.24. Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be bivariate normal random vectors as in Example 1.4.14. Then Y and $\tilde{Y}$ can be expressed as in Eqs. (1.4.28) and (1.4.29). We compute

$$\mathrm{Var}[Y|U] = \rho^2 \quad \text{and} \quad \mathrm{Var}[\tilde{Y}|\tilde{U}] = \tilde{\rho}^2,$$

that is, here both $\mathrm{Var}[Y|U]$ and $\mathrm{Var}[\tilde{Y}|\tilde{U}]$ are degenerate random variables. It is easy to see that

$$\mathrm{Var}[Y|U] \leq_{\mathrm{st}} \mathrm{Var}[\tilde{Y}|\tilde{U}] \Longleftrightarrow |\rho| \leq |\tilde{\rho}|,$$

that is,

$$(X,Y) \leq_{\text{GDO}_4\text{-st}} (\tilde{X},\tilde{Y}) \Longleftrightarrow |\rho| \leq |\tilde{\rho}|.$$

It is of interest to note that in comparing two bivariate normal random vectors, the above equivalence is similar to the equivalence

$$(X,Y) \leq_{\text{GDO}_2\text{-st}} (\tilde{X},\tilde{Y}) \Longleftrightarrow |\rho| \leq |\tilde{\rho}|$$

that was obtained in Shaked et al. [428]. ■

Example 1.4.25. Let (X,Y) and $(\tilde{X},\tilde{Y})$ be random vectors with discrete joint probability mass functions as in Example 1.4.16. Below we find necessary and sufficient conditions on the parameters p_{ij}'s and $\tilde{p}_{ij}$'s that imply $(X,Y) \leq_{\text{GDO}_4\text{-st}} (\tilde{X},\tilde{Y})$.

Let U be a uniform (0,1) random variable, independent of X, and consider the representation (1.4.30) for Y. Similarly, let $\tilde{U}$ be a uniform (0,1) random variable, independent of $\tilde{X}$, and consider the representation (1.4.31) for $\tilde{Y}$. As in Example 1.4.16, consider the following two cases:

Case 1: $\boxed{p_{00}p_{11} \geq p_{01}p_{10} \text{ and } \tilde{p}_{00}\tilde{p}_{11} \geq \tilde{p}_{01}\tilde{p}_{10}.}$ Note that from the inequality $p_{00}p_{11} \geq p_{01}p_{10}$ it follows that $\frac{p_{01}}{p_{0\cdot}} \leq \frac{p_{11}}{p_{1\cdot}}$. Then from Eq. (1.4.30) we see that

$$[Y|U=u] = \begin{cases} 0, & \text{if } u \in [0, \frac{p_{01}}{p_{0\cdot}}]; \\ 1-X, & \text{if } u \in [\frac{p_{01}}{p_{0\cdot}}, \frac{p_{11}}{p_{1\cdot}}]; \\ 1, & \text{if } u \in [\frac{p_{11}}{p_{1\cdot}}, 1]. \end{cases}$$

Hence

$$\mathrm{Var}[Y|U] = \begin{cases} p_{0\cdot}p_{1\cdot}, & \text{with probability } \frac{p_{11}}{p_{1\cdot}} - \frac{p_{01}}{p_{0\cdot}}; \\ 0, & \text{with probability } 1 - \left(\frac{p_{11}}{p_{1\cdot}} - \frac{p_{01}}{p_{0\cdot}}\right). \end{cases}$$

Similarly,

$$\mathrm{Var}[\tilde{Y}\big|\tilde{U}] = \begin{cases} p_{0\cdot}p_{1\cdot}, & \text{with probability } \frac{\tilde{p}_{11}}{p_{1\cdot}} - \frac{\tilde{p}_{01}}{p_{0\cdot}}; \\ 0, & \text{with probability } 1 - \left(\frac{\tilde{p}_{11}}{p_{1\cdot}} - \frac{\tilde{p}_{01}}{p_{0\cdot}}\right). \end{cases}$$

Thus, $\mathrm{Var}[Y\big|U] \leq_{\mathrm{st}} \mathrm{Var}[\tilde{Y}\big|\tilde{U}]$ if, and only if, $p_{11} \leq \tilde{p}_{11}$ (which is equivalent to $p_{01} \geq \tilde{p}_{01}$). Thus we see that in Case 1, $(X, Y) \leq_{\mathrm{GDO_4\text{-}st}} (\tilde{X}, \tilde{Y})$ if, and only if, $(\tilde{X}, \tilde{Y})$ is, informally, "more positively dependent" than (X, Y).

Case 2: $\boxed{p_{00}p_{11} \leq p_{01}p_{10} \text{ and } \tilde{p}_{00}\tilde{p}_{11} \leq \tilde{p}_{01}\tilde{p}_{10}.}$ Note that from the inequality $p_{00}p_{11} \leq p_{01}p_{10}$ it follows that $\frac{p_{11}}{p_{1\cdot}} \leq \frac{p_{01}}{p_{0\cdot}}$. A computation, similar to the one in Case 1, yields

$$\mathrm{Var}[Y\big|U] = \begin{cases} p_{0\cdot}p_{1\cdot}, & \text{with probability } \frac{p_{01}}{p_{0\cdot}} - \frac{p_{11}}{p_{1\cdot}}; \\ 0, & \text{with probability } 1 - \left(\frac{p_{01}}{p_{0\cdot}} - \frac{p_{11}}{p_{1\cdot}}\right); \end{cases}$$

and

$$\mathrm{Var}[\tilde{Y}\big|\tilde{U}] = \begin{cases} p_{0\cdot}p_{1\cdot}, & \text{with probability } \frac{\tilde{p}_{01}}{p_{0\cdot}} - \frac{\tilde{p}_{11}}{p_{1\cdot}}; \\ 0, & \text{with probability } 1 - \left(\frac{\tilde{p}_{01}}{p_{0\cdot}} - \frac{\tilde{p}_{11}}{p_{1\cdot}}\right). \end{cases}$$

Now, $\mathrm{Var}[Y\big|U] \leq_{\mathrm{st}} \mathrm{Var}[\tilde{Y}\big|\tilde{U}]$ if, and only if, $p_{11} \geq \tilde{p}_{11}$. Thus, we see that in Case 2, $(X, Y) \leq_{\mathrm{GDO_4\text{-}st}} (\tilde{X}, \tilde{Y})$ if, and only if, $(\tilde{X}, \tilde{Y})$ is, informally, "more negatively dependent" than (X, Y).

Note that in this example we have that $(X, Y) \leq_{\mathrm{GDO_4\text{-}st}} (\tilde{X}, \tilde{Y}) \Longleftrightarrow (Y, X) \leq_{\mathrm{GDO_4\text{-}st}} (\tilde{Y}, \tilde{X})$.

For the cases in which one of (X, Y) and $(\tilde{X}, \tilde{Y})$ is positively dependent and the other is negatively dependent, it is possible to make a similar analysis; we do not give the details here either. ∎

1.5 An Application in Reliability Theory

In the previous section we identified conditions which lead to various GDOs, and we derived some properties of these GDOs. Now we describe a practical situation in which random vectors, that are ordered with respect to $\leq_{\mathrm{GDO_3\text{-}cx}}$, yield useful inequalities.

First we derive a result involving the behavior of order statistics that are associated with random vectors that are ordered with respect to the order $\leq_{\mathrm{GDO_3\text{-}cx}}$. We then apply it to problems that arise in reliability theory.

We recall the definition of the univariate increasing convex order (for a detailed study of the this order, see, e.g., Müller and Stoyan [335] or Shaked and Shanthikumar [426]). Let Z and $\tilde{Z}$ be two univariate random variables. It is said that Z is smaller than $\tilde{Z}$ with respect to the univariate increasing convex order (denoted as $Z \leq_{\text{icx}} \tilde{Z}$) if $\mathrm{E}[\phi(Z)] \leq \mathrm{E}[\phi(\tilde{Z})]$ for every increasing convex function ϕ for which the above expectations are well defined.

Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be two random vectors with the same marginal distributions, and let U and $\tilde{U}$ be as in Eq. (1.3.8). Denote, as in Eq. (1.3.9), the corresponding regression functions by k and $\tilde{k}$. Let $(U_1, Y_1), (U_2, Y_2), \ldots, (U_n, Y_n)$ be n independent copies of (U, Y), and let $(\tilde{U}_1, \tilde{Y}_1), (\tilde{U}_2, \tilde{Y}_2), \ldots, (\tilde{U}_n, \tilde{Y}_n)$ be n independent copies of $(\tilde{U}, \tilde{Y})$. Denote the order statistics that correspond to $U_1, U_2, \ldots, U_n$ by $U_{(1)} \leq U_{(2)} \leq \cdots \leq U_{(n)}$ and the order statistics that correspond to $\tilde{U}_1, \tilde{U}_2, \ldots, \tilde{U}_n$ by $\tilde{U}_{(1)} \leq \tilde{U}_{(2)} \leq \cdots \leq \tilde{U}_{(n)}$. In the following result we assume that k and $\tilde{k}$ are strictly increasing, but from Proposition 1.4.4, it is seen that this is not a very restrictive assumption.

Theorem 1.5.1. *Let (X, Y) and $(\tilde{X}, \tilde{Y})$ be two random vectors with the same marginal distributions, and let U and $\tilde{U}$ be as in Eq. (1.3.8). If the corresponding functions k and $\tilde{k}$ are strictly increasing, and if $(X, Y) \leq_{\text{GDO}_3\text{-cx}} (\tilde{X}, \tilde{Y})$, then*

$$\tilde{k}(\tilde{U}_{(n)}) \leq_{\text{icx}} k(U_{(n)}). \tag{1.5.1}$$

Proof: The proof is similar to the proof of Proposition 6.1 in Shaked et al. [428]. Let $\mathrm{E}[Y \big| U]_{(1)} \leq \mathrm{E}[Y \big| U]_{(2)} \leq \cdots \leq \mathrm{E}[Y \big| U]_{(n)}$ be the order statistics that correspond to a sample of n independent copies of $\mathrm{E}[Y \big| U]$. Similarly, let $\mathrm{E}[\tilde{Y} \big| \tilde{U}]_{(1)} \leq \mathrm{E}[\tilde{Y} \big| \tilde{U}]_{(2)} \leq \cdots \leq \mathrm{E}[\tilde{Y} \big| \tilde{U}]_{(n)}$ be the order statistics that correspond to a sample of n independent copies of $\mathrm{E}[\tilde{Y} \big| \tilde{U}]$. Since k is strictly increasing, it is apparent that $\mathrm{E}[Y \big| U]_{(i)} \stackrel{\text{st}}{=} k(U_{(i)})$, $i = 1, 2, \ldots, n$. From the hypothesis assumption we have $\mathrm{E}[\tilde{Y} \big| \tilde{U}] \leq_{\text{cx}} \mathrm{E}[Y \big| U]$, and this implies $\mathrm{E}[\tilde{Y} \big| \tilde{U}] \leq_{\text{icx}} \mathrm{E}[Y \big| U]$. The stochastic inequality (1.5.1) now follows from Corollary 4.A.16 of Shaked and Shanthikumar [426]. ■

Consider n reliability items that are going to be put in a parallel system. Each of the items is tested before it is put into the system, and as a result of the test, it may be reshaped or adjusted, and the adjustment may affect the item lifetime. Let $X_1, X_2, \ldots, X_n$

be the (random) results of the tests, and let $U_1, U_2, \ldots, U_n$ be the corresponding (random) "noises" as interpreted after Eq. (1.3.3) (the units of which are irrelevant for our conclusions below as can be seen from the discussion in Remark 1.4.6). Suppose that the X_is are not observable but that the realizations of the "noises" U_is can be observed (though not controlled). Let $Y_1, Y_2, \ldots, Y_n$ be the (random) adjusted lifetimes given the above tests results. We assume that the pairs $(U_1, Y_1), (U_2, Y_2), \ldots, (U_n, Y_n)$ are independent and identically distributed. We want to compare the above situation to another situation in which a different type of test is performed on the items before they are put into the system. This time the tests results are denoted by $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_n$, and $\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_n$ are the corresponding adjusted lifetimes of the items, given the above other tests results. We denote the corresponding observable "noises" by $\tilde{U}_1, \tilde{U}_2, \ldots, \tilde{U}_n$. Here too we assume that the pairs $(\tilde{U}_1, \tilde{Y}_1), (\tilde{U}_2, \tilde{Y}_2), \ldots, (\tilde{U}_n, \tilde{Y}_n)$ are independent and identically distributed.

If we assume that k and $\tilde{k}$ are strictly increasing, then the (random) lifetime of the system in the former case (using a notation from the proof of Theorem 1.5.1) is $\mathrm{E}[Y\big|U]_{(n)} \stackrel{\text{st}}{=} k(U_{(n)})$, whereas the corresponding lifetime in the latter case is $\mathrm{E}[\tilde{Y}\big|\tilde{U}]_{(n)} \stackrel{\text{st}}{=} \tilde{k}(\tilde{U}_{(n)})$. Supposing that $(\tilde{X}_1, \tilde{Y}_1)$ is more globally dependent than (X_1, Y_1) with respect to the order $\leq_{\text{GDO}_3\text{-cx}}$, we see from Theorem 1.5.1 that a "less accurate" test procedure (i.e., one in which a lifetime is more influenced by the "noise") yields a longer system lifetime, in the sense of $\leq_{\text{icx}}$. This fact can be useful to reliability engineers when they construct the testing procedure.

Of course, the costs of the test procedures, and the unusual cost of observing the "noise," also need to be taken into account by the engineers, but we do not discuss this issue here.

Acknowledgments

Moshe Shaked is supported by NSA grant H98230-12-1-0222. Miguel A. Sordo is supported by Ministerio de Ciencia e Innovación (grant MTM2009-08326) and Consejería de Economía Innovación y Ciencia (grant P09-SEJ-4739). Alfonso Suárez-Llorens is supported by Ministerio de Ciencia e Innovación (grant MTM2009-08326) and Consejería de Economía Innovación y Ciencia (grant P09-SEJ-4739).

Chapter 2
Duality Theory and Transfers for Stochastic Order Relations

Alfred Müller

Abstract: In this paper it will be demonstrated how functional analytic tools from duality theory can be used to give interesting characterizations of stochastic order relations for discrete distributions in terms of mass transfer principles. A general result for a large class of integral stochastic orders will be derived, and it will be shown that this applies to many important examples like usual stochastic order, convex order, supermodular order, directional convex order, and orthant orders.

2.1 Introduction

Many stochastic order relations have a nice interpretation in terms of a sequence of *mass transfers*, if the involved distributions have a finite support. We explain the idea by a simple example concerning

A. Müller (✉)
Department Mathematik, Universität Siegen, D–57068 Siegen, Germany
e-mail: mueller@mathematik.uni-siegen.de

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_2,

the usual stochastic order $\leq_{\rm st}$ for real-valued random variables. Let X, Y be real-valued random variables with

$$\mathrm{P}\{X = 0\} = \mathrm{P}\{X = 2\} = \frac{1}{4}, \quad \mathrm{P}\{X = 3\} = \frac{1}{2},$$

and

$$\mathrm{P}\{Y = 1\} = \mathrm{P}\{Y = 2\} = \frac{1}{4}, \quad \mathrm{P}\{Y = 4\} = \frac{1}{3}, \quad \mathrm{P}\{Y = 5\} = \frac{1}{6}.$$

Then $X \leq_{\rm st} Y$ holds. One way to see this is by using the fact that there are random variables $\hat{X}, \hat{Y}$ with the same marginal distributions as X and Y having the property that $\hat{X} \leq \hat{Y}$ almost surely. This can be obtained, e.g., by choosing the following joint distribution:

$$\begin{aligned}
&\mathrm{P}\{\hat{X} = 0, \hat{Y} = 1\} = \mathrm{P}\{\hat{X} = 2, \hat{Y} = 2\} = \frac{1}{4}, \\
&\mathrm{P}\{\hat{X} = 3, \hat{Y} = 4\} = \frac{1}{3} \quad \text{ and } \mathrm{P}\{\hat{X} = 3, \hat{Y} = 5\} = \frac{1}{6}.
\end{aligned}$$

This statement can be reinterpreted as saying that the distribution of Y can be obtained from the distribution of X by a sequence of three *mass transfers* moving probability mass upwards, namely moving probability mass 1/4 upwards from 0 to 1, moving probability mass 1/3 upwards from 3 to 4, and moving probability mass 1/6 upwards from 3 to 5 (the probability mass of 1/4 in the point 2 is not moved).

Using a little bit more formal notation, we can write this as follows. Let δ_x denote the point mass in $x \in \mathbb{R}$, then we can write the distributions of X and Y as

$$P_X = \frac{1}{4}\delta_0 + \frac{1}{4}\delta_2 + \frac{1}{2}\delta_3, \quad P_Y = \frac{1}{4}\delta_1 + \frac{1}{4}\delta_2 + \frac{1}{3}\delta_4 + \frac{1}{6}\delta_5.$$

The fact that P_Y can be obtained from P_X by three mass transfers moving probability mass upwards can be written as

$$P_Y - P_X = \frac{1}{4}(\delta_1 - \delta_0) + \frac{1}{3}(\delta_4 - \delta_3) + \frac{1}{6}(\delta_5 - \delta_3).$$

Recall that the difference of two measures is called a *signed measure.* Thus the statement that P_Y is obtained from P_X by a sequence of mass transfers moving probability mass upwards can be stated in mathematical terms as $P_Y - P_X$ is a positive linear combination of signed measures of the form $\delta_y - \delta_x$ with $x < y$.

The aim of this paper is to demonstrate that such a principle of transfers holds for many well-known stochastic order relations and that there is a unified principle on how to prove this in general by using functional analytic results from duality theory. The use of methods from duality theory for the analysis of stochastic order relations is not new, see, e.g., [76, 331].

The study of mass transfer principles as described above has recently found increasing interest in the economics literature in the context of comparing multivariate risks, see, e.g., [109, 318, 334]. Indeed, the basic principle that is used in this paper has already been used in [318, 334] for the special cases of supermodular ordering and inframodular ordering.

2.2 Transfers and Integral Stochastic Orders

In this section the general principle of a transfer is introduced. To do so, the basic facts about signed measures are needed. Let $S \subseteq \mathbb{R}^d$ be a Borel subset of some Euclidean space and let $\mathcal{S}$ be the Borel-σ-algebra on S. If μ^+ and μ^- are two finite measures on $(S, \mathcal{S})$, the difference $\mu = \mu^+ - \mu^-$ is a *signed measure*. Such a signed measure is a mapping $\mu : \mathcal{S} \to \mathbb{R}$ that is σ-additive with $\mu(\emptyset) = 0$. By the well-known Hahn-Jordan decomposition theorem we have for any signed measure μ on $(S, \mathcal{S})$ a unique decomposition $\mu = \mu^+ - \mu^-$ as a difference between two measures with the property that there is a measurable subset $E \subset S$ such that $\mu^+(E) = 0$ and $\mu^-(S \backslash E) = 0$. These measures μ^+ and μ^- are then called the positive and negative parts, respectively. The measure $|\mu| := \mu^+ + \mu^-$ is called the total variation and $\|\mu\| := \mu^+(S) + \mu^-(S)$ is the total variation norm. Denote by $\mathbb{M}$ the set of all signed measures on S with finite total variation norm $\|\mu\| < \infty$ and with the property that $\mu^+(S) = \mu^-(S)$. Notice that for any two probability measures P, Q, the difference $Q - P \in \mathbb{M}$ and that in fact any $\mu \in \mathbb{M}$ is a multiple of such a difference of two probability measures, i.e., $\mathbb{M}$ is the linear space spanned by the differences of probability measures.

A degenerate probability measure on $\boldsymbol{x}$ is denoted $\delta_{\boldsymbol{x}}$. Given two probability measures P, Q supported on a finite subset of $\mathbb{R}^d$, call the signed measure $Q - P$ a *transfer* from P to Q. If

$$(Q - P)^- = \sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{x}_i} \quad \text{and} \quad (Q - P)^+ = \sum_{i=1}^{m} \beta_i \delta_{\boldsymbol{y}_i},$$

then the transfer $Q - P$ removes probability mass α_i from point $\boldsymbol{x}_i$, $i = 1, \ldots, n$ and adds probability mass β_i to $\boldsymbol{y}_i$, $i = 1, \ldots, m$. To indicate this transfer we write

$$\sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{x}_i} \to \sum_{i=1}^{m} \beta_i \delta_{\boldsymbol{y}_i}.$$

In the examples that we consider later on we will also illustrate this graphically by green and red points. The meaning will be that the red points will depict the points $\boldsymbol{x}_i$ where we remove mass, and the green points will depict the points $\boldsymbol{y}_i$ where the mass is moved to. Thus an increasing transfer will look as follows (Fig. 2.1):

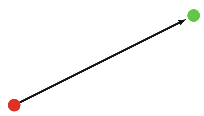


Figure 2.1: Increasing transfer

Next we define classes of functions that are generated by sets of transfers. In the following we denote by $\mathcal{C}$ the set of all continuous functions $f : S \to \mathbb{R}$.

Definition 2.2.1. Consider a set $M \subset \mathbb{M}$ of transfers and the class $\mathcal{F} \subset \mathcal{C}$ of continuous functions f such that

$$\sum_{i=1}^{m} \beta_i f(\boldsymbol{y}_i) \geq \sum_{i=1}^{n} \alpha_i f(\boldsymbol{x}_i)$$

whenever $\mu \in M$, where $\mu := \sum_{i=1}^{m} \beta_i \delta_{\boldsymbol{y}_i} - \sum_{i=1}^{n} \alpha_i \delta_{\boldsymbol{x}_i}$. The class $\mathcal{F}$ is said to be *induced* by M.

Any class of functions $\mathcal{F}$ defines a so-called *integral stochastic order*. Stochastic orders of this type have been considered in detail in [306, 331]. In the following we will denote the mentioned stochastic orders with the symbols used in the monographs [335, 426].

Definition 2.2.2. A probability measure P is dominated by a probability measure Q with respect to the integral order $\leq_{\mathcal{F}}$ (denoted $P \leq_{\mathcal{F}} Q$) if

$$\int u \mathrm{d}P \leq \int u \mathrm{d}Q \quad \text{for all} \quad u \in \mathcal{F}.$$

The class $\mathcal{F}$ of functions is called the *generator* of the order $\leq_{\mathcal{F}}$.

We will now demonstrate that many important examples of stochastic orders have a generator $\mathcal{F}$ that is induced by a set of transfers $M \subset \mathbb{M}$. In all cases that we consider it doesn't matter whether or not we assume continuity of the function in $\mathcal{F}$. This follows from the results in [120]. In some cases there are different possibilities to define the transfers. We will speak of *simple transfers* if they move mass only from a very small number of points to a small number of other points, typically from at most two points to at most two other points.

Throughout this paper the Euclidean space $\mathbb{R}^d$ will be endowed with the natural componentwise order, where for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ we write $\boldsymbol{x} \leq \boldsymbol{y}$ if $x_i \leq y_i$ for $i = 1, \ldots, d$.

In the following descriptions of transfers $\eta \in [0,1]$ will always describe the total mass that is moved by the transfer.

Simple Increasing Transfer

Given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$ with $\boldsymbol{x} \leq \boldsymbol{y}$ and $\eta \in [0,1]$ a simple transfer $\eta\delta_{\boldsymbol{x}} \to \eta\delta_{\boldsymbol{y}}$ is called *increasing*. The reverse transfer is called *decreasing*.

If M is the set of increasing transfers then it is obvious that this induces the class $\mathcal{F}$ of continuous increasing functions. Thus in this case $\leq_{\mathcal{F}}$ is the usual stochastic order $\leq_{\mathrm{st}}$ (Fig. 2.2).

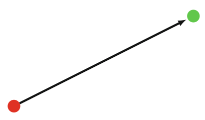


Figure 2.2: Increasing transfer

Simple Convex Transfer

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ and $\alpha, \beta, \gamma, \varepsilon \in [0,1]$ such that

$$\begin{aligned} &\boldsymbol{z} = \alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y}, \quad \boldsymbol{w} = \beta \boldsymbol{y} + (1-\beta)\boldsymbol{x}, \\ &\gamma \boldsymbol{x} + (1-\gamma)\boldsymbol{y} = \varepsilon \boldsymbol{z} + (1-\varepsilon)\boldsymbol{w}, \end{aligned}$$

a simple transfer $\eta\,(\varepsilon\delta_{\boldsymbol{z}} + (1-\varepsilon)\delta_{\boldsymbol{w}}) \to \eta\,(\gamma\delta_{\boldsymbol{x}} + (1-\gamma)\delta_{\boldsymbol{y}})$ is called *convex*. The reverse transfer is called *concave*. When $\alpha = \beta$, hence $\gamma = \varepsilon = 1/2$, the transfer is called symmetric. Notice that if $\alpha = 1-\beta$, then $\boldsymbol{w} = \boldsymbol{z}$ (Fig. 2.3).

General Convex Transfer

For a discrete measure $P = \sum_{i=1}^m \alpha_i \delta_{\boldsymbol{x}_i}$ the *barycenter* is defined as

$$\mathrm{bar}(P) = \frac{1}{\sum_{i=1}^m \alpha_i} \sum_{i=1}^m \alpha_i \boldsymbol{x}_i$$

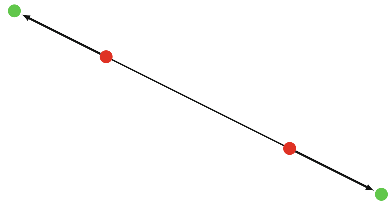


Figure 2.3: Simple convex transfer

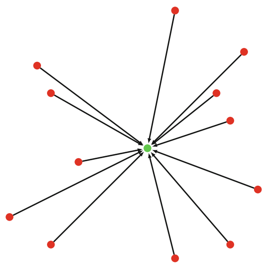


Figure 2.4: General concave transfer (fusion)

A *general convex transfer* is a transfer $\eta\delta_{\mathrm{bar}(P)} \to \eta P = \eta \sum_{i=1}^{m} \alpha_i \delta_{\boldsymbol{x}_i}$ for some $\eta \in [0,1]$. The reverse transfer is called *concave* (Fig. 2.4).

The general (non-simple) convex transfers are obtained by iterating simple convex transfers. In dimension $d = 1$ a convex transfer is nothing else than a mean-preserving spread, as studied by [394–396]. In dimension d concave transfers are related to fusions [148].

The class $\mathcal{F}$ generated by the set of all convex transfers is just the class of all convex functions. Thus $\leq_{\mathcal{F}}$ is the convex order $\leq_{\mathrm{cx}}$.

Simple Supermodular Transfer

The following notation is used here:

$$\begin{aligned} \boldsymbol{x} \vee \boldsymbol{y} &:= (\max\{x_1, y_1\}, \ldots, \max\{x_d, y_d\}), \\ \boldsymbol{x} \wedge \boldsymbol{y} &:= (\min\{x_1, y_1\}, \ldots, \min\{x_d, y_d\}). \end{aligned}$$

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ such that

$$\boldsymbol{x} = \boldsymbol{z} \wedge \boldsymbol{w}, \quad \boldsymbol{y} = \boldsymbol{z} \vee \boldsymbol{w},$$

a simple transfer $\eta\left(\frac{1}{2}\delta_{\boldsymbol{z}} + \frac{1}{2}\delta_{\boldsymbol{w}}\right) \to \eta\left(\frac{1}{2}\delta_{\boldsymbol{x}} + \frac{1}{2}\delta_{\boldsymbol{y}}\right)$ is called *supermodular*. The reverse transfer is called *submodular*.

The class $\mathcal{F}$ generated by the set of all supermodular transfers is just the class of all supermodular functions. Thus $\leq_{\mathcal{F}}$ is the supermodular order $\leq_{\mathrm{sm}}$ (Fig. 2.5).

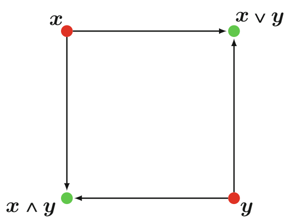


Figure 2.5: Supermodular transfer

Simple Directionally Convex Transfer

Given $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{z} \in \mathbb{R}^d$ and $\gamma, \varepsilon \in [0,1]$ such that $\boldsymbol{x} \le \boldsymbol{w} \le \boldsymbol{y}$, $\boldsymbol{x} \le \boldsymbol{z} \le \boldsymbol{y}$ and

$$\gamma \boldsymbol{x} + (1-\gamma)\boldsymbol{y} = \varepsilon \boldsymbol{z} + (1-\varepsilon)\boldsymbol{w}, \tag{2.2.1}$$

a simple transfer $\eta\,(\varepsilon\delta_{\boldsymbol{z}} + (1-\varepsilon)\delta_{\boldsymbol{w}}) \to \eta\,(\gamma\delta_{\boldsymbol{x}} + (1-\gamma)\delta_{\boldsymbol{y}})$ is called *directionally convex*. The reverse transfer is called *directionally concave*. When $\gamma = \varepsilon = 1/2$, the transfer is called symmetric.

The class $\mathcal{F}$ generated by the set of all directionally convex transfers is just the class of all directionally convex functions. Thus $\le_{\mathcal{F}}$ is the convex order $\le_{\text{dcx}}$. Notice that Eq. (2.2.1) ensures that the points $\boldsymbol{x}, \boldsymbol{w}, \boldsymbol{y}, \boldsymbol{z}$ are the vertices of a parallelogram (Fig. 2.6).

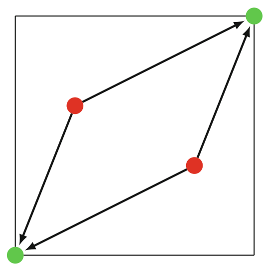


Figure 2.6: Directionally convex transfer

Δ-Monotone Transfer

Let $\boldsymbol{x} \le \boldsymbol{y}$ with strict inequality $x_i < y_i$ in k variables $i_1, \dots, i_k$ for some $k \in \{1, \dots, d\}$. Then $[\boldsymbol{x}, \boldsymbol{y}] := \{\boldsymbol{z} : \boldsymbol{x} \le \boldsymbol{z} \le \boldsymbol{y}\}$ is a k-dimensional hyperbox with 2^k vertices. Let V_o be the subset of vertices $\boldsymbol{z}$ with the property that the number of components with $z_i = x_i,\ i \in \{i_1, \dots, i_k\}$ is odd. Similarly, V_e shall be the subset of vertices where this number is even.

A transfer removing mass η from each of the vertices in V_o and moving mass η to each of the vertices in V_e is called a Δ-*monotone transfer*.

The class $\mathcal{F}$ generated by the set of all Δ-monotone transfers is the class of all Δ-monotone functions. To see this, recall the definition of a Δ-monotone function $f : \mathbb{R}^d \to \mathbb{R}$. For a function $f : \mathbb{R}^d \to \mathbb{R}$ define the *difference operators*

$$\Delta_i^\epsilon f(\boldsymbol{x}) \;=\; f(\boldsymbol{x}+\epsilon \boldsymbol{e}_i) - f(\boldsymbol{x}),$$

where $\boldsymbol{e}_i$ is the i-th unit vector and $\epsilon > 0$. The function f is said to be Δ-*monotone*, if for every subset $J = \{i_1, \dots, i_k\} \subset \{1, \dots, d\}$ and every $\epsilon_1, \dots, \epsilon_k > 0$

$$\Delta_{i_1}^{\epsilon_1} \dots \Delta_{i_k}^{\epsilon_k} f(\boldsymbol{x}) \;\geq\; 0 \quad \text{for all } \boldsymbol{x}.$$

Defining $\boldsymbol{y}$ by $y_i := x_i + \epsilon_i$ if $i \in \{i_1, \dots, i_k\}$ and $y_i = x_i$ otherwise, it is easy to see that

$$\Delta_{i_1}^{\epsilon_1} \dots \Delta_{i_k}^{\epsilon_k} f(\boldsymbol{x}) = \sum_{\boldsymbol{z} \in V_e} f(\boldsymbol{z}) - \sum_{\boldsymbol{z} \in V_o} f(\boldsymbol{z}).$$

Thus the class $\mathcal{F}$ generated by the set of all Δ-monotone transfers is the class of all Δ-monotone functions, and therefore in this case $\leq_{\mathcal{F}}$ is the upper orthant order $\leq_{\mathrm{uo}}$, as this is generated by the class of Δ-monotone functions, see [335], Theorem 3.3.15.

In the same theorem one can find the related result that the lower orthant order $\leq_{lo}$ is generated by the class of functions with the property that $\boldsymbol{x} \mapsto -f(-\boldsymbol{x})$ is Δ-monotone. We call these functions Δ-*antitone transfer*. Therefore a similar concept of transfer can be defined which leads to the lower orthant order. One just has to replace in the definition of the Δ-monotone transfer the sets V_o and V_e by the sets $\bar{V}_o$ and $\bar{V}_e$, where $\bar{V}_o$ is the subset of vertices $\boldsymbol{z}$ with the property that $(-1)^k$ times the number of components with $z_i = x_i,\ i \in \{i_1, \dots, i_k\}$ is odd. Similarly, $\bar{V}_e$ shall be the subset of vertices where $(-1)^k$ times the number of components with $z_i = x_i,\ i \in \{i_1, \dots, i_k\}$ is even. We will denote this as a Δ-*antitone transfer* (Fig. 2.7).

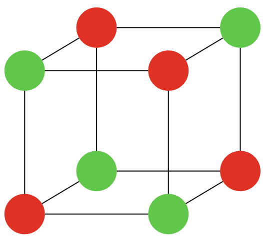


Figure 2.7: Δ-monotone transfer in dimension $d = 3$

2.3 Duality Theory

The proof of our main results in the next section requires some results from a part of functional analysis that is known as *duality theory*. For a detailed description of that general theory we refer to [89]. Here we will describe the most important facts adapted to our setting.

For $S \subset \mathbb{R}^d$ compact, denote by $\mathcal{C}$ the set of continuous functions on S. By the compactness assumption on S these functions are all bounded and therefore integrable with respect to any $\mu \in \mathbb{M}$.

In functional analysis it is convenient to describe integrals as a bilinear form $\langle f, \mu \rangle = \int f \mathrm{d}\mu = \int f \mathrm{d}\mu^+ - \int f \mathrm{d}\mu^-$.

A pair (E, F) of vector spaces is said to be in *duality*, if there is a bilinear mapping $\langle \cdot, \cdot \rangle : E \times F \to \mathbb{R}$. The duality is said to be *strict*, if for each $0 \neq x \in E$ there is a $y \in F$ with $\langle x, y \rangle \neq 0$ and if for each $0 \neq y \in F$ there is an $x \in E$ with $\langle x, y \rangle \neq 0$.

Unfortunately the duality $(\mathbb{M}, \mathcal{C})$ is not strict, as $\langle f, \mu \rangle = 0$ for all $\mu \in \mathbb{M}$ only implies f to be constant. But strict duality can be obtained by identifying functions which differ only by a constant. Formally, define an equivalence relation $f \sim g$ if $f - g$ is constant. Equivalently, fix some $s_0 \in S$ and require $f(s_0) = 0$. Denote the corresponding quotient space by $\mathcal{C}_\sim$.

Lemma 2.3.1. *$\mathbb{M}$ and $\mathcal{C}_\sim$ are in strict duality under the bilinear mapping*

$$\begin{aligned} \langle \cdot, \cdot \rangle &: \mathbb{M} \times \mathcal{C}_\sim \to \mathbb{R}, \\ \langle \mu, f \rangle &= \int f \mathrm{d}\mu. \end{aligned}$$

A crucial role in our further investigations is played by the bipolar theorem for convex cones. The notion of polar is introduced following the notation of [89].

The *polar* M° of a set $M \subset E$ (in the duality (E, F)) is defined as

$$M^\circ = \{ y \in F : \ \langle x, y \rangle \geq -1 \text{ for all } x \in M \}. \tag{2.3.1}$$

The polar of a set $N \subset F$ is defined analogously.

Given a vector space V, a subset $K \subset V$ is called a *cone* if $x \in K$ implies $\alpha x \in K$ for all $\alpha \geq 0$. Given any subset $M \subset V$, the convex cone $\mathrm{co}(M)$ generated by M is the smallest convex cone that contains M.

Define the *dual cone* of an arbitrary set $M \subset E$ by

$$M^* = \{y \in F : \langle x, y\rangle \geq 0 \text{ for all } x \in M\}.$$

It is easy to see that M^* is a convex cone. Moreover, notice that for a convex cone K the polar and dual cones coincide: $K^\circ = K^*$.

For any duality (E, F) define the *weak topology* $\sigma(E, F)$ on E as the weakest topology on E such that the mappings $x \mapsto \langle x, y\rangle$ are continuous for all $y \in F$. Now the bipolar theorem for convex cones can be stated as follows [89, Corollary 22.10].

Theorem 2.3.2. *Suppose E and F are in strict duality and $X \subset E$ is an arbitrary set. Then X^{**} is the weak closure of the convex cone generated by X.*

2.4 Main Results

Theorem 2.3.2 will be the key to prove many results of the following type:

Let M be a set of transfers, and let $\mathcal{F}$ be the class of functions induced by M. Then for probability measures P and Q with finite support $P \leq_{\mathcal{F}} Q$ holds if and only if Q can be obtained from P by a finite number of transfers from M.

Indeed we can show the following general result.

Theorem 2.4.1. *Assume that the convex cone $co(M)$ generated by M is weakly closed in the duality $(\mathbb{M}, \mathcal{C}_\sim)$, and let $\mathcal{F}$ be the class of functions induced by M. Then for probability measures P and Q with finite support $P \leq_{\mathcal{F}} Q$ holds if and only if Q can be obtained from P by a finite number of transfers from M.*

Proof: $P \leq_{\mathcal{F}} Q$ holds if and only if $\int f \mathrm{d}P \leq \int f \mathrm{d}Q$ or equivalently $\int f \mathrm{d}(Q - P) \geq 0$ for all $f \in \mathcal{F}$, where $Q - P$ is a signed measure in $\mathbb{M}$. Using the terminology of duality theory from the last section, this can be rewritten as $Q - P \in \mathcal{F}^*$. The fact that $\mathcal{F}$ is the class of functions induced by M can be rewritten as $\mathcal{F} = M^*$: thus $Q - P \in M^{**}$. Therefore it follows from Theorem 2.3.2 that $Q - P$ is in the weak closure of the convex cone generated by M. If we assume that the convex cone $co(M)$ is weakly closed this implies that $Q - P \in co(M)$; thus $Q - P = \sum_{i=1}^n \eta_i \mu_i$ with $\eta_i > 0$ and $\mu_i \in M$. As P and Q are probability measures, it is possible to choose $\eta_i \le 1$. But this means that Q can be obtained from P by a finite number of transfers $\eta_i \mu_i$, $i = 1, \ldots, n$. ■

The crucial mathematical assumption that has to be shown in examples thus is the property that $co(M)$ is weakly closed. This has typically to be done by showing that we can choose an appropriate finite S such that $supp(P) \cup supp(Q) \subset S$ having the following property: if $M(S)$ is the restriction of M to transfers moving only mass within S then $M(S)$ is compact and induces a class of functions $f : S \to \mathbb{R}$ such that all these functions $f : S \to \mathbb{R}$ coincide with the restrictions of the functions $f : \mathbb{R}^d \to \mathbb{R}$ in $\mathcal{F}$. This follows from the following simple corollary of Theorem 2.4.1.

Corollary 2.4.2. *Assume that S is finite and that M is compact. Then $co(M)$ is weakly closed.*

Proof: As S is finite, the set of signed measures can be identified with a finite-dimensional Euclidean space $\mathbb{R}^{|S|}$ where the elements of $\mathbb{R}^{|S|}$ are the counting densities of the signed measures. The weak topology in this case is just the usual topology of pointwise convergence. It is well known that in a finite-dimensional Euclidean space the convex hull of a compact set is closed, and thus also $co(M)$ is closed. ■

2.5 Examples

It will be shown now case by case that for all important examples mentioned in Sect. 2.2 we can find an appropriate finite S.

Usual Stochastic Order

For the usual stochastic order $\le_{\mathrm{st}}$ on $\mathbb{R}^d$ we can choose the finite set $S = supp(P) \cup supp(Q)$. The set of all increasing transfers on S induces the class of functions $f : S \to \mathbb{R}$ that are increasing as a function on S. It is clear that the restriction of any increasing function $f : \mathbb{R}^d \to \mathbb{R}$ to S is an increasing function on S. On the other hand, any increasing function $f : S \to \mathbb{R}$ can be extended to an increasing function $f : \mathbb{R}^d \to \mathbb{R}$ by defining

$$f(\boldsymbol{x}) := \sup\{f(\boldsymbol{y}) : \boldsymbol{y} \in S, \boldsymbol{y} \le \boldsymbol{x}\}, \quad \boldsymbol{x} \in \mathbb{R}^d,$$

with the convention that $\sup \emptyset := \inf\{f(\boldsymbol{y}) : \boldsymbol{y} \in S\}$.

This gives a new proof of the following known result.

Theorem 2.5.1. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \leq_{\mathrm{st}} \boldsymbol{Y}$.

(b) $P_{\boldsymbol{Y}}$ *can be obtained from* $P_{\boldsymbol{X}}$ *by a finite number of simple increasing transfers.*

(c) There are random vectors $\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}}$ *on a common probability space, having the same distributions as* $\boldsymbol{X}, \boldsymbol{Y}$ *with the property that* $\hat{\boldsymbol{X}} \leq \hat{\boldsymbol{Y}}$ *almost surely.*

Proof: The equivalence of (a) and (b) follows immediately from Theorem 2.4.1 and the discussion above. For the implication of (b) to (c) it is sufficient to consider the case that $P_{\boldsymbol{Y}}$ can be obtained from $P_{\boldsymbol{X}}$ by a simple increasing transfer $\eta\delta_{\boldsymbol{x}} \to \eta\delta_{\boldsymbol{y}}$. In this case it is necessary that $P_{\boldsymbol{X}}(\{\boldsymbol{x}\}) \geq \eta$. Choose any nonatomic probability space, on which we can define a random vector $\hat{\boldsymbol{X}}$ with distribution $P_{\boldsymbol{X}}$. Then there is a set A with $P_{\boldsymbol{X}}(A) = \eta$ and $\hat{\boldsymbol{X}}(\omega) = \boldsymbol{x}$ if $\omega \in A$. Define $\hat{\boldsymbol{Y}}(\omega) = \boldsymbol{y}$ if $\omega \in A$, and $\hat{\boldsymbol{Y}}(\omega) = \hat{\boldsymbol{X}}(\omega)$ otherwise. Then obviously $\hat{\boldsymbol{X}}, \hat{\boldsymbol{Y}}$ have the desired properties. The implication of (c) to (a) is obvious. ■

Convex Orders

For the order $\leq_{\mathrm{cx}}$ on $\mathbb{R}^d$, we can again choose the finite set $S = supp(P) \cup supp(Q)$. However, in this case we cannot work with the set of simple convex transfers, as these require that we can move the mass along some line, whereas it may happen that S does not contain any three points on a line. Consider as an example on $\mathbb{R}^2$

$$P = \delta_{(1,1)} \quad \text{and} \quad Q = \frac{1}{6}\delta_{(0,3)} + \frac{1}{6}\delta_{(2,3)} + \frac{1}{3}\delta_{(0,0)} + \frac{1}{3}\delta_{(2,0)}.$$

Then clearly $P \leq_{\mathrm{cx}} Q$, but we cannot obtain Q from P by simple convex transfers within $S = supp(P) \cup supp(Q)$, as there do not exist any simple convex transfers on S. Therefore the class of functions $\mathcal{F}$ on S induced by the set of simple convex transfers would be the set of all functions on S which does not coincide with the restriction of convex functions on $\mathbb{R}^2$ to S. It would be possible to obtain Q from P by simple convex transfers, if we add the points $(1,0)$ and $(1,2)$ to S. However, it is not easy to see in general which points one has to add to $supp(P) \cup supp(Q)$ to get an appropriate set S to work with.

A much better approach is to work with general convex transfers and with $S = supp(P) \cup supp(Q)$. Using that approach we can show the following result.

Theorem 2.5.2. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \leq_{\mathrm{cx}} \boldsymbol{Y}$.

(b) $P_{\boldsymbol{Y}}$ *can be obtained from* $P_{\boldsymbol{X}}$ *by a finite number of general convex transfers.*

(c) $P_{\boldsymbol{Y}}$ *can be obtained from* $P_{\boldsymbol{X}}$ *by a finite number of simple convex transfers.*

Proof: To show the equivalence of (a) and (b), let M be the set of all general convex transfers on an arbitrary finite set S. Then M induces the class $\mathcal{F}$ of discretely convex functions on S, which is defined as the set of convex functions on $\mathbb{R}^d$ restricted to S. Thus any restriction of a convex function to S is obviously in $\mathcal{F}$. Vice versa any function $f \in \mathcal{F}$ can be extended constructively to a convex function on $\mathbb{R}^d$ by taking the supremum over all affine functions on $\mathbb{R}^d$ which are smaller or equal to f on the finite set S. Moreover, M is compact and thus the equivalence of (a) and (b) follows from Theorem 2.4.1.

The equivalence of (b) and (c) follows from the fact that a general convex transfer can be obtained by a finite number of simple convex transfers. This is most easily seen by looking at the reverse general concave transfer, which is also known as a *fusion* of the probability mass on n points to one point, its barycenter. It is clear that this fusion can be done in at most n steps always taking the fusion of the mass in two points to their barycenter, which in each step yields a simple concave transfer, diminishing the number of points in the support by one. ■

Results very similar to Theorem 2.5.2 can be found in [148, 149]. In the univariate case a convex transfer can also be considered as a discrete version of a mean-preserving spread as considered in [394].

We can show a similar result for the increasing convex order $\leq_{\mathrm{icx}}$.

Theorem 2.5.3. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \leq_{\mathrm{icx}} \boldsymbol{Y}$.

(b) $P_{\boldsymbol{Y}}$ *can be obtained from* $P_{\boldsymbol{X}}$ *by a finite number of transfers, which are either convex transfers or increasing transfers.*

Proof: The proof is very similar to the proof of Theorem 2.5.2. The only change is that we have to choose M as the union of the sets of all convex transfers and all increasing transfers on an arbitrary finite set S. Then M induces the class $\mathcal{F}$ of discretely convex and increasing functions on S. It is obvious that any restriction of an increasing convex function to S is in $\mathcal{F}$. Vice versa any function $f \in \mathcal{F}$ can be extended constructively to an increasing convex function on $\mathbb{R}^d$ by taking the supremum over all affine functions on $\mathbb{R}^d$ which are smaller or equal to f on the finite set S. Moreover, M is compact and thus the equivalence of (a) and (b) follows from Theorem 2.4.1. ∎

Supermodular and Directionally Convex Order

For the supermodular order $\le_{\mathrm{sm}}$ on $\mathbb{R}^d$ we choose as the finite set S the smallest product set containing $supp(P) \cup supp(Q)$, i.e., if $supp(P) \cup supp(Q) = \{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ with $\boldsymbol{x}_i = (x_{i1}, \dots, x_{id})$ then we choose

$$S = \{x_{11}, \dots, x_{n1}\} \times \cdots \times \{x_{1d}, \dots, x_{nd}\}. \tag{2.5.1}$$

This is obviously a lattice so that it holds for any $\boldsymbol{x}, \boldsymbol{y} \in S$ that $\boldsymbol{x} \vee \boldsymbol{y} \in S$ and $\boldsymbol{x} \wedge \boldsymbol{y} \in S$. Using this setting we can show the following representation of supermodular ordering.

Theorem 2.5.4. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \le_{\mathrm{sm}} \boldsymbol{Y}$.

(b) *$P_{\boldsymbol{Y}}$ can be obtained from $P_{\boldsymbol{X}}$ by a finite number of supermodular transfers.*

Proof: Let M be the set of supermodular transfers on the set S described in Eq. (2.5.1). Then M induces the class $\mathcal{F}$ of supermodular functions on S. The restriction of any supermodular function on $\mathbb{R}^d$ to the sublattice S obviously is supermodular on S. On the other hand, any supermodular function $f : S \to \mathbb{R}$ can be extended to a supermodular function $f : \mathbb{R}^d \to \mathbb{R}$ as follows. For a finite subset $S_i \subset \mathbb{R}$ we define the projection mapping $h_i : \mathbb{R} \to S_i$ by $h_i(x) := \sup\{y \in S_i : y \le x\}$ with the convention $\sup \emptyset := \inf S_i$. For a finite product set $S = S_1 \times \cdots \times S_d \subset \mathbb{R}^d$ we then define the projection mapping $h : \mathbb{R}^d \to S$ by $h(\boldsymbol{x}) := (h_1(x_1), \dots, h_d(x_d))$. It is easy to see that $h(\boldsymbol{x}) \vee h(\boldsymbol{y}) = h(\boldsymbol{x} \vee \boldsymbol{y})$ and $h(\boldsymbol{x}) \wedge h(\boldsymbol{y}) = h(\boldsymbol{x} \wedge \boldsymbol{y})$.

Therefore a supermodular function $f : S \to \mathbb{R}$ can be extended to a supermodular function $f : \mathbb{R}^d \to \mathbb{R}$ by defining for any $\boldsymbol{x} \in \mathbb{R}^d$

$$f(\boldsymbol{x}) := f(h(\boldsymbol{x})). \tag{2.5.2}$$

Moreover, as S is finite, the set M is compact and thus the equivalence of (a) and (b) follows from Theorem 2.4.1. ∎

In the bivariate case a result similar to Theorem 2.5.4 can be found in [448]. For the case of higher dimensions this seems to be new.

For directionally convex order $\le_{\text{dcx}}$ we can also use S as defined in Eq. (2.5.1). For the proof of the corresponding representation result we can refer to [334], where the equivalent case of inframodular order has been considered in detail. In particular, it is shown in Sect. 4.3 in that paper how any discretely directionally convex function on S can be extended to a directional convex function on $\mathbb{R}^d$ using a componentwise linear extension which is similar to the extension of a subcopula to a copula described in [415]. We get the following result for directional convex order.

Theorem 2.5.5. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \le_{\text{dcx}} \boldsymbol{Y}$.

(b) $P_{\boldsymbol{Y}}$ *can be obtained from* $P_{\boldsymbol{X}}$ *by a finite number of directionally convex transfers.*

Orthant Orders

As a last example we will consider the lower orthant order $\le_{\text{lo}}$ and the upper orthant order $\le_{\text{uo}}$. Here we will generalize recent results of [109], where a representation result by transfers has been proved for these orders under the restrictions that $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same marginals. Decancq [109] used a completely different approach with a tedious constructive proof. We will use duality theory to show the same result without the restriction of having the same marginals.

Theorem 2.5.6. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \le_{\text{uo}} \boldsymbol{Y}$.

(b) $P_{\boldsymbol{Y}}$ *can be obtained from* $P_{\boldsymbol{X}}$ *by a finite number of* Δ*-monotone transfers.*

Proof: Again we choose S as in Eq. (2.5.1). Let M be the class of Δ-monotone transfers on S. Then M induces the class $\mathcal{F}$ of Δ-monotone functions on S. The restriction of any Δ-monotone function on $\mathbb{R}^d$ to the sublattice S obviously is Δ-monotone on S. On the other hand, any Δ-monotone function $f : S \to \mathbb{R}$ can be extended to a Δ-monotone function $f : \mathbb{R}^d \to \mathbb{R}$ by using the same construction as in Eq. (2.5.2). Moreover, as S is finite, M is compact and thus the equivalence of (a) and (b) follows from Theorem 2.4.1. ∎

Notice that in case $d = 1$ the order $\leq_{\text{uo}}$ is just the usual stochastic order $\leq_{\text{st}}$ and in that case a Δ-monotone transfer is nothing else than an increasing transfer. In case $d = 2$ the order $\leq_{\text{uo}}$ is equivalent to the increasing supermodular order $\leq_{ism}$ and indeed in that case a Δ-monotone transfer is either an increasing transfer (if the hyperbox $[\boldsymbol{x}, \boldsymbol{y}]$ is degenerated to a line segment) or a supermodular transfer (if the hyperbox in nondegenerated).

With exactly the same idea of proof we get the following corresponding result for the lower orthant order $\leq_{\text{lo}}$.

Theorem 2.5.7. *For random vectors $\boldsymbol{X}, \boldsymbol{Y}$ with finite support the following statements are equivalent:*

(a) $\boldsymbol{X} \leq_{\text{lo}} \boldsymbol{Y}$.

(b) $P_{\boldsymbol{Y}}$ can be obtained from $P_{\boldsymbol{X}}$ by a finite number of Δ-antitone transfers.

Chapter 3
Reversing Conditional Orderings

Rachele Foschi and Fabio Spizzichino

Abstract: We analyze some specific aspects concerning conditional orderings and relations among them. To this purpose we define a suitable concept of *reversed conditional ordering* and prove some related results. In particular we aim to compare the univariate stochastic orderings $\leq_{\mathrm{st}}$, $\leq_{\mathrm{hr}}$, and $\leq_{\mathrm{lr}}$ in terms of differences among different notions of conditional orderings. Some applications of our result to the analysis of positive dependence will be detailed. We concentrate attention to the case of a pair of scalar random variables X, Y. Suitable extensions to multivariate cases are possible.

R. Foschi (✉)
Laboratory of Innovation Management and Economics, IMT Advanced Studies Lucca, Piazza S. Ponziano 6, 55100 Lucca, Italy
e-mail: rachele.foschi@imtlucca.it

F. Spizzichino
Dipartimento di Matematica G. Castelnuovo, Università degli Studi di Roma La Sapienza, Piazzale Aldo Moro 5, 00185 Roma, Italy
e-mail: fabio.spizzichino@uniroma1.it

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_3,

3.1 Introduction

Stochastic orderings between random variables (or random vectors) constitute primary tools for the description and the characterization of concepts of stochastic dependence. On the one hand the relevant literature in this direction is very well established (see in particular [426] and the references contained therein). On the other hand it still continues to offer various suggestions for interesting work. Here we consider stochastic orderings between (one-dimensional) conditional distributions, also called *conditional orderings*. We will analyze some specific aspects concerning such orderings and relations among them.

Especially in a statistical setting, the following problem is of interest: what can be said about dependence of X w.r.t. Y (where X and Y are random variables or random vectors) when we assume that Y is stochastically increasing w.r.t. X in some specified sense? Attention to this topic has been given several times in the literature under different standpoints or different languages. One can see in particular the basic paper [155], [440, Chapter 3] and the recent papers [105, 106]. Related to this theme one can also see [410].

Some detailed aspects of this theme will be considered here for the special case of two scalar random variables. As a motivating purpose, we aim to compare the univariate stochastic orderings $\leq_{\mathrm{st}}$, $\leq_{\mathrm{hr}}$, and $\leq_{\mathrm{lr}}$ in terms of differences among notions of conditional orderings.

For scalar random variables X and Y, we consider different conditional orderings of the form

$$\mathcal{L}\left(Y|X \in I\right) \leq_{*} \mathcal{L}\left(Y|X \in I'\right), \tag{3.1.1}$$

where I, I' are intervals of different types and $\leq_{*}$ stands for $\leq_{\mathrm{st}}$, $\leq_{\mathrm{hr}}$, or $\leq_{\mathrm{lr}}$. In a few words, we can summarize our work by saying that we analyze implications or equivalences concerning such relations. Along this direction we will show some simple results that, at the best of our knowledge, have not been pointed out so far.

A concept of *reversed conditional ordering* will in particular emerge as natural from our discussion and our results will point out some symmetries existing between the mentioned univariate stochastic orderings $\leq_{\mathrm{st}}$, $\leq_{\mathrm{hr}}$, $\leq_{\mathrm{lr}}$ and different types of conditional orderings [where a "type" of ordering can be defined in terms of the possible choices for the intervals I, I' appearing in Eq. (3.1.1)]. A main result in this direction is Theorem 3.2.6.

It is clear that conditional orderings define special notions of positive dependence (see in particular [91] and references cited therein). In this paper, we will analyze positive dependence properties corresponding to the considered conditional orderings and we will see how results concerning implications and equivalences between conditional orderings can be translated in terms of dependence notions.

Then we will point out some direct applications of our results to dependence notions related with conditions of *default contagion* and to the case of *conditional independence* between X and Y.

More in details, this paper is organized as follows. In Sect. 3.2 we first introduce some formal concepts needed to give a general definition of *reversion* of a conditional ordering. Then (Theorem 3.2.6) we point out a specific property, related with conditioning, of the $\leq_{\text{st}}$ order. Definitions and results given in Sect. 3.2 will be directly applied in Sect. 3.3, where we detail the specific conditional orderings of our interest and present Theorem 3.3.2. In a few words we show how each conditional ordering of the form (3.1.1) is equivalent to one of the form

$$\mathcal{L}\left(X|Y \in J\right) \leq_{\widetilde{*}} \mathcal{L}\left(X|Y \in J'\right), \tag{3.1.2}$$

for suitable choice of the stochastic order $\leq_{\widetilde{*}}$ and of the intervals J, J'.

The equivalence of a conditional ordering of the form (3.1.1) or of the form (3.1.2) with a corresponding concept of dependence will be treated in Sect. 3.4. Arguments presented therein will directly suggest the definition of a new positive dependence concept that is related with the notion of stochastic increasing and that we denote by SIRL. Section 3.5 presents two different types of applications of Theorem 3.3.2: concepts of default contagion and cases of conditional independence between X and Y. The latter application adds some potentially useful insight about positive dependence of conditionally independent random variables that are stochastically increasing w.r.t. a conditioning variable Z. Finally, we present a short discussion with some concluding remarks in Sect. 3.6. In Appendix, we recall some notation and basic facts about stochastic dependence and copulas (see also [211, 355]).

The choice of restricting our analysis to pairs of scalar random variables, besides allowing us to simplify notation and definitions, is also motivated by specially relevant symmetries related with reversing conditional orderings. Our arguments, however, admit suitable generalizations to the multivariate case.

3.2 The Role of Usual Stochastic Ordering in Conditioning

Let X, Y be two real-valued random variables. Let furthermore E and E' be random events; $\mathcal{L}(X)$, $\mathcal{L}(X|E)$, $\mathcal{L}(Y)$, $\mathcal{L}(Y|E')$ will denote, respectively, the probability laws of X, X conditional on E, Y, Y conditional on E'. As a first issue in this section, we give a suitably general definition of *stochastic monotonicity* of Y w.r.t. X. To this purpose, it is convenient to define an order on the class $\mathcal{I}$ of all the intervals of $\mathbb{R}_+$.

Definition 3.2.1. For two intervals I, I' belonging to $\mathcal{I}$, we set $I \Subset I'$ if

$$\inf I < \inf I' \text{ or } \inf I = \inf I',\ \sup I' < \sup I.$$

The relation $\Subset$ is symmetric, anti-reflexive, and transitive and it defines a total order on $\mathcal{I}$.

Remark 3.2.2. From an intuitive point of view, the relation $I \Subset I'$ has the following meaning: for a lifetime X, the condition of belonging to I' is (in a special sense adapt for our context) more restrictive than the one of belonging to I. In fact, when $\inf I < \inf I'$, requiring that X reaches $\inf I'$ is a stronger condition than requiring that it reaches $\inf I$. If $\inf I = \inf I'$, we take into account the length of the interval: in this case it is more restrictive requiring that X falls into the shorter interval. Thus, we have, e.g., for any $x' > x > 0$, $(0, +\infty) \Subset (x, +\infty) \Subset (x, x') \Subset \{x\}$.

Let $\leq_*$ denote a univariate stochastic order and let $\mathcal{A}, \mathcal{B}$ be two subclasses of $\mathcal{I}$. On the basis of Definition 3.2.1, we can now give the following definition.

Definition 3.2.3. Y is $(\leq_*,\ \mathcal{A},\ \mathcal{B})$ stochastically increasing in X or stochastically increasing w.r.t. X in the sense $(\leq_*,\ \mathcal{A},\ \mathcal{B})$, if and only if, for any $I \in \mathcal{A}$, $I' \in \mathcal{B}$, $I \Subset I'$,

$$\mathcal{L}(Y|X \in I) \leq_* \mathcal{L}(Y|X \in I').$$

More shortly, we will also write $Y \uparrow_{(\leq_*,\ \mathcal{A},\ \mathcal{B})} X$.

Remark 3.2.4. By Definition 3.2.3, we obtain the classical notion of Y stochastically increasing in X (SI($Y|X$), see, e.g., Appendix section) by setting

$$* = \mathrm{st},\ \mathcal{A} = \mathcal{B} = \{(x-\varepsilon, x+\varepsilon) | \varepsilon > 0,\ x > \varepsilon\}.$$

Definition 3.2.5. Let $\mathcal{A}$, $\mathcal{B}$, $\widetilde{\mathcal{A}}$, $\widetilde{\mathcal{B}}$ be classes of intervals and $*, \tilde{*}$ be stochastic orderings. The relation $Y \uparrow_{(\le_*,\ \mathcal{A},\ \mathcal{B})} X$ is *reverted* by $X \uparrow_{(\le_{\tilde{*}},\ \widetilde{\mathcal{A}},\ \widetilde{\mathcal{B}})} Y$ if and only if, for any $I \in \mathcal{A}$, $I' \in \mathcal{B}$, two sets $J \in \widetilde{\mathcal{A}}$, $J' \in \widetilde{\mathcal{B}}$ exist, such that

$$\mathcal{L}(Y|X \in I) \le_* \mathcal{L}(Y|X \in I') \iff \mathcal{L}(X|Y \in J) \le_{\tilde{*}} \mathcal{L}(X|Y \in J').$$

In this paper, we consider the special cases of Definition 3.2.3 obtained by combining the following choices:

(1) $* = \mathrm{st}$

(2) $* = \mathrm{hr}$

(3) $* = \mathrm{lr}$

(A) $\mathcal{L}(Y) \le_* \mathcal{L}(Y|X > x) \quad \forall x > 0$

(B) $\mathcal{L}(Y|X > x) \le_* \mathcal{L}(Y|X > x') \quad \forall x < x'$

(C) $\mathcal{L}(Y|X = x) \le_* \mathcal{L}(Y|X = x') \quad \forall x < x'$

By using the notation of Definition 3.2.3, we respectively have

(A) $\mathcal{A} = \{\mathbb{R}_+\}$ and $\mathcal{B} = \{(x, +\infty) | x > 0\}$

(B) $\mathcal{A} = \mathcal{B} = \{(x, +\infty) | x > 0\}$

(C) $\mathcal{A} = \mathcal{B} = \{(x-\varepsilon, x+\varepsilon) | \varepsilon > 0,\ x > \varepsilon\}$

The following result points out a property of the usual stochastic order that is relevant in our setting. This result will allow us to find triples $(\le_*,\ \mathcal{A},\ \mathcal{B})$, $(\le_{\tilde{*}},\ \widetilde{\mathcal{A}},\ \widetilde{\mathcal{B}})$ satisfying Definition 3.2.5.

Theorem 3.2.6. *Let A, B, B' be intervals, with $B \Subset B'$. Then intervals A', D, D' exist, with $D \Subset D'$ and such that*

$$\mathcal{L}(Y|Y \in A, X \in B) \le_{\mathrm{st}} \mathcal{L}(Y|Y \in A, X \in B') \tag{3.2.1}$$

is equivalent to

$$\mathcal{L}(X|X \in A', Y \in D) \le_{\mathrm{st}} \mathcal{L}(X|X \in A', Y \in D'). \tag{3.2.2}$$

Proof: By definition of usual stochastic order, inequalities (3.2.1) and (3.2.2) also read as

$$\mathrm{P}\{Y > y | Y \in A, X \in B\} \leq \mathrm{P}\{Y > y | Y \in A, X \in B'\} \quad \forall\, y \geq 0, \tag{3.2.3}$$

$$\mathrm{P}\{X > x | X \in A', Y \in D\} \leq \mathrm{P}\{X > x | X \in A', Y \in D'\} \quad \forall\, x \geq 0. \tag{3.2.4}$$

By Bayes' theorem, Eq. (3.2.4) can be rewritten as

$$\frac{\mathrm{P}\{X > x, X \in A', Y \in D\}}{\mathrm{P}\{X \in A', Y \in D\}} \leq \frac{\mathrm{P}\{X > x, X \in A', Y \in D'\}}{\mathrm{P}\{X \in A', Y \in D'\}}$$

and subsequently

$$\frac{\mathrm{P}\{Y \in D | X > x, X \in A'\}}{\mathrm{P}\{Y \in D | X \in A'\}} \leq \frac{\mathrm{P}\{Y \in D' | X > x, X \in A'\}}{\mathrm{P}\{Y \in D' | X \in A'\}}.$$

On its turn, Eq. (3.2.3) becomes

$$\frac{\mathrm{P}\{Y > y, Y \in A | X \in B\}}{\mathrm{P}\{Y \in A | X \in B\}} \leq \frac{\mathrm{P}\{Y > y, Y \in A | X \in B'\}}{\mathrm{P}\{Y \in A | X \in B'\}}.$$

Thus, Eqs. (3.2.3) and (3.2.4) are equivalent if and only if, for some $x, y > 0$,

$$\begin{cases} D = A \\ D' = A \cap (y, +\infty) \\ A' = B \\ B' = A' \cap (x, +\infty). \end{cases} \tag{3.2.5}$$

Since A is given, $D \Subset D'$. In fact, $\inf D' = \max(y, \inf A) \geq \inf A$. We notice that, since intervals are connected sets, $\inf I = \inf I'$ implies $I \subset I'$ or $I' \subset I$. Hence the condition $\inf I = \inf I'$ implies $I \cup I' \Subset I \cap I'$; therefore

- If $y > \inf A$, $A \Subset A \cap (y, +\infty)$
- If $y \leq \inf A$, $A \cup (A \cap (y, +\infty)) \Subset A \cap (y, +\infty)$, that is, again, $A \Subset A \cap (y, +\infty)$

∎

Remark 3.2.7. Theorem 3.2.6 implicitly provides a condition for the existence of the intervals A', D, D' and therefore for a conditional order being revertible. We notice that Eq. (3.2.5) is a system of four equations in the three unknowns A', D, D'. The last equation allows us to check the existence of solutions for the system. Nonexistence of solutions means that the considered conditional order is not revertible. When the solution exists, but at least one of the intervals A', D, D' is empty, then the considered conditional order is not revertible either.

Let A, A', B, D be intervals as in Theorem 3.2.6.

Proposition 3.2.8. *$A \Subset A'$ if and only if $D \Subset B$.*

Proof: The claim straightly follows by the definition, given in the proof of Theorem 3.2.6, of the sets A' and D in terms of A and B. ∎

Theorem 3.2.6 will have a basic role for our purposes in the next section. In this respect, it is useful to recall that the stochastic orderings $\leq_{\mathrm{hr}}$, $\leq_{\mathrm{lr}}$ can be characterized in terms of $\leq_{\mathrm{st}}$ (see [426]). More precisely, one has

Proposition 3.2.9. *$\mathcal{L}(S) \leq_{\mathrm{hr}} \mathcal{L}(T)$ if and only if, for any $t \geq 0$,*

$$\mathcal{L}(S - t|S > t) \leq_{\mathrm{st}} \mathcal{L}(T - t|T > t).$$

Proposition 3.2.10. *$\mathcal{L}(S) \leq_{\mathrm{lr}} \mathcal{L}(T)$ if and only if $\mathcal{L}(S \mid S \in A) \leq_{\mathrm{st}} \mathcal{L}(T \mid T \in A)$ for any measurable set A.*

In Definition 3.2.5, the interchange in the role of the variables X, Y corresponds to a change in the stochastic order and in the choice of the conditioning events. We notice that there is a *balance* between the strength of the stochastic order and the strength of the conditioning events. Actually, in view of Propositions 3.2.9 and 3.2.10, Proposition 3.2.8 guarantees that if $\tilde{*}$ is stronger than $*$, then $D \Subset B$ and $D' \Subset B'$.

3.3 Remarkable Properties of Conditional Orderings and Related Inversions

In this section we analyze in details the conditional orderings defined by the positions (1)–(3), (A)–(C), mentioned in Sect. 3.2. In particular by applying Theorem 3.2.6, we will find, for each entry, the corresponding *reversed* ordering.

By combining (A)–(C) with cases (1)–(3), we obtain the following matrices $M(Y|X)$, $M(X|Y)$:

	A	B	C
1	A1$(Y\|X)$	B1$(Y\|X)$	C1$(Y\|X)$
2	A2$(Y\|X)$	B2$(Y\|X)$	C2$(Y\|X)$
3	A3$(Y\|X)$	B3$(Y\|X)$	C3$(Y\|X)$

(3.3.1)

	A	B	C
1	A1$(X\|Y)$	B1$(X\|Y)$	C1$(X\|Y)$
2	A2$(X\|Y)$	B2$(X\|Y)$	C2$(X\|Y)$
3	A3$(X\|Y)$	B3$(X\|Y)$	C3$(X\|Y)$

(3.3.2)

Each entry of the two matrices $M(Y|X), M(X|Y)$ is a property of conditional order. For example, A1$(Y|X)$ means $\mathcal{L}(Y) \leq_{st} \mathcal{L}(Y|X > x) \quad \forall x > 0$.

Remark 3.3.1. Heuristically speaking, the conditional orderings appearing in the matrices become stronger and stronger when reading their entries "from above to below" or "from left to right." In view of the chain of implications

$$\leq_{\mathrm{lr}} \Rightarrow \leq_{\mathrm{hr}} \Rightarrow \leq_{\mathrm{st}}, \tag{3.3.3}$$

we immediately obtain

$$A3 \Longrightarrow A2 \Longrightarrow A1, \quad B3 \Longrightarrow B2 \Longrightarrow B1, \quad C3 \Longrightarrow C2 \Longrightarrow C1.$$

On the other hand, the implications

$$C1 \Longrightarrow B1 \Longrightarrow A1, \quad C2 \Longrightarrow B2 \Longrightarrow A2, \quad C3 \Longrightarrow B3 \Longrightarrow A3$$

follow from the relation $\mathbb{R}_+ \Supset (x, +\infty) \Supset \{x\}$ in view of Theorem 3.2.6 and Proposition 3.2.8.

In principle the two matrices $M(Y|X), M(X|Y)$ present $18 = 9+9$ different properties for the joint law of (X, Y). Based on the existing literature (see, e.g., [155, 426, 440]), we can guess however the existence of some equivalences among them. Actually a complete catalogue of equivalences can be established between pairs of them in terms of Definition 3.2.5 and Theorem 3.2.6. More precisely, we have the following result:

Theorem 3.3.2. *The matrix $M(X|Y)$ is the transpose of $M(Y|X)$.*

Proof: We apply Theorem 3.2.6 to any property corresponding to the entries of $M(Y|X)$. We provide the detailed proof for the elements in the first column of $M(Y|X)$ (thus obtaining the first row of $M(X|Y)$). The other equivalences can be proven by analogous arguments.

A1$(Y|X)$ satisfies condition (3.2.1) with $A = B = \mathbb{R}_+$, $B' = (x', +\infty)$, $\forall x' > 0$. Equation (3.2.5) in the proof of Theorem 3.2.6 allows us to write the equivalent condition (3.2.2) with $A' = D = \mathbb{R}_+$, $D' = (y, +\infty)$, $\forall y > 0$. Finally, we see that $B' = (x', +\infty)$ satisfies the last condition of the system (3.2.5), $B' = A' \cap (x, +\infty)$. Hence, by Theorem 3.2.6, A1$(Y|X)$ is equivalent to the inequality (3.2.2) that turns out to be equivalent to A1$(X|Y)$.

In view of Proposition 3.2.9, A2$(Y|X)$ is equivalent to condition (3.2.1) with $A = (t, +\infty)$, $\forall t \geq 0$; $B = \mathbb{R}_+$, $B' = (x', +\infty)$, $\forall x' > 0$. By Theorem 3.2.6, A2$(Y|X)$ is equivalent to Eq. (3.2.2) with $A' = \mathbb{R}_+$, $D = (t, +\infty)$, $\forall t \geq 0$; $D' = (\max(t, y), +\infty)$, $\forall y > 0$. Since $\max(t, y) \geq y$, Eq. (3.2.2) turns out to be equivalent to B1$(X|Y)$.

In view of Proposition 3.2.10, A3$(Y|X)$ is equivalent to condition (3.2.1) with $A = (t - \varepsilon, t + \varepsilon)$, $\forall \varepsilon > 0$, $t > \varepsilon$; $B = \mathbb{R}_+$, $B' = (x', +\infty)$, $\forall x' > 0$. By Theorem 3.2.6, A3$(Y|X)$ is equivalent to Eq. (3.2.2) with $A' = \mathbb{R}_+$, $D = (t - \varepsilon, t + \varepsilon)$, $\forall \varepsilon > 0$, $t > \varepsilon$; $D' = (t - \varepsilon, t + \varepsilon) \cap (y, +\infty)$, $\forall y > 0$. $D' = (\max(y, t - \varepsilon), t + \varepsilon) \subset D$ and therefore $D \Subset D'$ also holds. In other words, in the limit for ε going to 0, D' collapses in a point on the right of $\{t\}$. Therefore Eq. (3.2.2) turns out to be equivalent to C1$(X|Y)$. ∎

In view of Theorem 3.3.2, the table in Eq. (3.3.1) also reads

A1$(Y\|X)$	B1$(Y\|X)$	C1$(Y\|X)$
B1$(X\|Y)$	B2$(Y\|X)$	C2$(Y\|X)$
C1$(X\|Y)$	C2$(X\|Y)$	C3$(Y\|X)$

(3.3.4)

Remark 3.3.3. 1. The hazard rate order $\leq_{\text{hr}}$ can be characterized as follows (see [426]): $Y \geq_{\text{hr}} X$ if and only if

$$\frac{\overline{G}_X(t)}{\overline{G}_Y(t)} \textit{ is decreasing in } t. \tag{3.3.5}$$

Equation (3.3.5) allows us to define $\leq_{\text{hr}}$ even if one or both the distributions to be compared are not absolutely continuous.

2. For what concerns the likelihood ratio order, a characterization not involving densities can be given as follows: $X \leq_{\mathrm{lr}} Y$ if and only if
$$\mathrm{P}\{X \in A\}\mathrm{P}\{Y \in B\} \geq \mathrm{P}\{X \in B\}\mathrm{P}\{Y \in A\} \tag{3.3.6}$$
for all measurable sets A and B such that $A \leq B$, where $A \leq B$ means that $x \in A$ and $y \in B$ imply that $x \leq y$.

All our results and the conditional orderings considered so far do not require absolute continuity.

In the absolutely continuous case, the proof of some of the equivalences in Theorem 3.3.2 could also have been directly obtained by applying Bayes' formula. In particular, we can, for instance, argue as follows:

-
$$\begin{aligned}
\mathrm{C2}(Y|X) &\iff \mathcal{L}(Y|X=x) \leq_{\mathrm{hr}} \mathcal{L}(Y|X=x') \quad \forall x < x' \\
&\iff \frac{\overline{G}_Y(y'|X=x)}{\overline{G}_Y(y|X=x)} \leq \frac{\overline{G}_Y(y'|X=x')}{\overline{G}_Y(y|X=x')} \quad \forall x < x',\ y < y' \\
&\iff \frac{g_X(x'|Y>y)}{g_X(x|Y>y)} \leq \frac{g_X(x'|Y>y')}{g_X(x|Y>y')} \quad \forall x < x',\ y < y' \\
&\iff \mathcal{L}(X|Y>y) \leq_{\mathrm{lr}} \mathcal{L}(X|Y>y') \quad \forall y < y' \\
&\iff \mathrm{B3}(X|Y),
\end{aligned}$$

-
$$\begin{aligned}
\mathrm{A3}(Y|X) &\iff \mathcal{L}(Y) \leq_{\mathrm{lr}} \mathcal{L}(Y|X>x) \\
&\iff \overline{G}_X(x|Y=y) \leq \overline{G}_X(x|Y=y') \quad \forall y < y' \\
&\iff \mathcal{L}(X|Y=y) \leq_{\mathrm{st}} \mathcal{L}(X|Y=y') \\
&\iff \mathrm{C1}(X|Y),
\end{aligned}$$

-
$$\begin{aligned}
&\mathrm{B2}(Y|X) \\
\iff\ & \mathcal{L}(Y|X>x) \leq_{\mathrm{hr}} \mathcal{L}(Y|X>x') \quad \forall x < x' \\
\iff\ & \overline{G}_Y(y'|X>x')\overline{G}_Y(y|X>x) \geq \overline{G}_Y(y|X>x')\overline{G}_Y(y'|X>x) \\
& \quad \forall x < x',\ y < y' \\
\iff\ & \overline{G}_X(x'|Y>y')\overline{G}_X(x|Y>y) \geq \overline{G}_X(x|Y>y')\overline{G}_X(x'|Y>y) \\
& \quad \forall x < x',\ y < y' \\
\iff\ & \mathcal{L}(X|Y>y) \leq_{\mathrm{hr}} \mathcal{L}(X|Y>y') \quad \forall y < y' \\
\iff\ & \mathrm{B2}(X|Y).
\end{aligned}$$

Remark 3.3.4. The properties A1, B2, C3 lie on the main diagonal of the matrix $M(Y|X)$. Therefore, in view of Theorem 3.3.2, they must be symmetric w.r.t. X, Y, i.e., $A1(Y|X) = A1(X|Y)$, $B2(Y|X) = B2(X|Y)$, $C3(Y|X) = C3(X|Y)$. In these cases, we notice that the interchange between X, Y does not require a *balancing* between a change in the stochastic order and in the conditioning events.

Remark 3.3.5. When X, Y are exchangeable, $M(Y|X) = M(X|Y)$. Therefore, since $M(Y|X)$ coincides with its own transpose, it is symmetric.

3.4 Conditional Orderings and Dependence

All the entries of the matrices $M(Y|X), M(X|Y)$ can be seen as dependence properties for the pair (X, Y). In some cases the correspondences are well known or immediately follow by definitions. This is the case for A1, B1, C1, C3 (see Appendix and the table in Eq. (3.4.1) below). This section will be devoted to analyzing the remaining ones.

The following result is a simple consequence of Bayes' formula.

Proposition 3.4.1. *$B2(Y|X) \iff (X, Y)$ is* RCSI.

Proof: By Eq. (3.3.5), $B2(Y|X)$, i.e., $\mathcal{L}(Y|X > x) \leq_{\text{hr}} \mathcal{L}(Y|X > x')$ for any $x < x'$, $y < y'$, can be rewritten as

$$\overline{G}_Y(y'|X > x')\overline{G}_Y(y|X > x) \geq \overline{G}_Y(y|X > x')\overline{G}_Y(y'|X > x),$$

for any $x < x'$, $y < y'$. By applying Bayes' formula, we obtain the inequality

$$\overline{F}(x', y')\overline{F}(x, y) \geq \overline{F}(x, y')\overline{F}(x', y)$$

that is the definition of $\overline{F}$ TP_2, i.e., (X, Y) is right corner set increasing (RCSI) (see Appendix and [355]). ■

In view of the above arguments, we can rewrite $M(Y|X)$ in the following form:

PQD	RTI($Y|X$)	SI($Y|X$)	
RTI($X|Y$)	RCSI	C2($Y|X$)	(3.4.1)
SI($X|Y$)	C2($X|Y$)	PLRD	

On the main diagonals of $M(Y|X)$ and $M(X|Y)$, we find the symmetric dependence properties positive quadrant dependent (PQD), RCSI, positive likelihood ratio dependent (PLRD).

We aim now at completing the correspondences between dependence properties and entries of $M(Y|X)$.

Also C2$(Y|X)$, C2$(X|Y)$ can be interpreted as conditions of stochastic dependence. However they do not correspond, as far as we know, to any definition introduced so far in the literature. In this respect we state the following:

Proposition 3.4.2. *C2$(Y|X)$ holds if and only if $Y - t|Y > t$ is* SI *in X for any $t > 0$.*

Proof: By Eq. (3.3.5), the condition C2$(Y|X)$, i.e.,

$$\mathcal{L}(Y|X = x) \leq_{\mathrm{hr}} \mathcal{L}(Y|X = x'),$$

is equivalent to

$$\overline{G}_Y(y'|X = x')\overline{G}_Y(y|X = x) \geq \overline{G}_Y(y|X = x')\overline{G}_Y(y'|X = x)$$

for any $x < x'$, $y < y'$, that is, $\dfrac{\overline{G}_Y(y'|X = x)}{\overline{G}_Y(y|X = x)}$ is increasing in x. Since $y < y'$, we can write $y' = y + t$, for $t > 0$, and thus

$$\frac{\overline{G}_Y(y + t|X = x)}{\overline{G}_Y(y|X = x)} = \overline{G}_Y(y + t|X = x, Y > y) \uparrow x, \tag{3.4.2}$$

for all $t > 0$. ■

In view of the above result, it is natural to give the following.

Definition 3.4.3. Y is stochastically increasing in X in the residual lifetime (in short SIRL$(Y|X)$) if SI$(Y - t|Y > t, X)$ for any $t > 0$.

The matrix $M(Y|X)$ can be finally rewritten as

PQD	RTI$(Y\|X)$	SI$(Y\|X)$
RTI$(X\|Y)$	RCSI	SIRL$(Y\|X)$
SI$(X\|Y)$	SIRL$(X\|Y)$	PLRD

(3.4.3)

Condition SIRL$(Y|X)$ is a stronger dependence property than SI$(Y|X)$. In particular, the chain of equivalences (3.3.3) implies PLRD $\Longrightarrow$ SIRL $\Longrightarrow$ RCSI. Concerning the identification of SIRL with a property of stochastic dependence, it is of interest to note that the following result relates such a property with the behavior of the survival copula $\hat{C}(u, v) = \overline{F}(\overline{G}_X^{-1}(u), \overline{G}_Y^{-1}(v))$ (see also Appendix).

Proposition 3.4.4. *Y is* SIRL *in X if and only if $\frac{\partial \hat{C}}{\partial u}(u,v)$ is* TP_2*.*

Proof: The term $\overline{G}_Y(y+t|X=x, Y>t)$ can also be rewritten as

$$\frac{\mathrm{P}\{X=x, Y>y+t\}}{\mathrm{P}\{X=x, Y>t\}} = \frac{\frac{\partial \hat{C}}{\partial u}(\overline{G}_X(x), \overline{G}_Y(y+t))}{\frac{\partial \hat{C}}{\partial u}(\overline{G}_X(x), \overline{G}_Y(t))}.$$

Since this function has to be increasing in x, by adopting the change of variables

$$u = \overline{G}_X(x),\ u' = \overline{G}_X(x'),\ v = \overline{G}_Y(y+t),\ z = \overline{G}_Y(t),$$

we obtain

$$\frac{\partial \hat{C}(u,v)}{\partial u}\frac{\partial \hat{C}(u',z)}{\partial u} \leq \frac{\partial \hat{C}(u',v)}{\partial u}\frac{\partial \hat{C}(u,z)}{\partial u}. \tag{3.4.4}$$

Since $x < x' < t < y+t$ and, therefore, $u > u',\ z > v$, Eq. (3.4.4) corresponds to $\frac{\partial \hat{C}}{\partial u}(u,v)$ being TP_2. ■

In view of Proposition 3.4.4 and Definition 3.7.3 in Appendix, the conditional orderings treated here can be written as properties of a copula. This circumstance is not surprising, since conditional orderings are not affected by strictly increasing (deterministic) transformations, as it happens for copulas.

If the copula is symmetric, in particular in the exchangeable case, $M(Y|X) = M(X|Y)$, then we have in particular (for both the matrices)

B1 $\Longleftrightarrow$ A2

B3 $\Longleftrightarrow$ C2

C1 $\Longleftrightarrow$ A3

In the case when the copula is Archimedean, we have the further equivalences

B1 $\Longleftrightarrow$ B2 (see, e.g., [44])

C1 $\Longleftrightarrow$ C2

as stated by the following proposition.

Proposition 3.4.5. *Let $\hat{C}$ be an Archimedean copula. Then* SI $\Longrightarrow$ SIRL*.*

Proof: Let be $\hat{C}(u,v) = \phi(\phi^{-1}(u) + \phi^{-1}(v))$. $\hat{C}$ is SI if and only if, for any $a, x \geq 0$, $\dfrac{\phi'(x+a)}{\phi'(x)}$ is increasing w.r.t. x, i.e., for any $x < x'$,

$$\frac{\phi'(x+a)}{\phi'(x)} \leq \frac{\phi'(x'+a)}{\phi'(x')}. \tag{3.4.5}$$

In view of Proposition 3.4.4, the thesis consists in $\dfrac{\partial \hat{C}}{\partial u}(u,v)$ being TP_2, i.e., for any $u, u', v, v' \in [0,1],\ u < u', v < v'$,

$$\phi'(\alpha' + \beta)\phi'(\alpha + \beta') \leq \phi'(\alpha + \beta)\phi'(\alpha' + \beta'), \tag{3.4.6}$$

where $\alpha = \phi^{-1}(u),\ \alpha' = \phi^{-1}(u'),\ \beta = \phi^{-1}(v),\ \beta' = \phi^{-1}(v')$ and therefore $\alpha' < \alpha,\ \beta' < \beta$. By choosing $x = \alpha' + \beta',\ x' = \alpha + \beta',\ a = \beta - \beta'$, we obtain that Eq. (3.4.5) implies Eq. (3.4.6). ■

3.5 Applications of Theorem 3.3.2

The circumstance that each entry of the matrix $M(Y|X)$ is equivalent to an entry of $M(X|Y)$ turns out to be of interest in several different contexts. Some of them will be sketched in what follows.

3.5.1 Default Contagion and Dynamic Dependence Properties

In several applied models, the following type of conditional ordering can be of interest:

$$\mathcal{L}(Y|X = x) \leq_* \mathcal{L}(Y|X > x). \tag{3.5.1}$$

It is natural to wonder whether property (3.5.1) is different from those appearing in $M(Y|X)$. Actually, such stochastic inequalities are equivalent to the ones appearing in column B of $M(Y|X)$. More precisely we can state

Proposition 3.5.1. *For* $* = \text{st}, \text{hr}, \text{lr}$*, condition* (3.5.1) *is equivalent to*

$$\mathcal{L}(Y|X > x) \leq_* \mathcal{L}(Y|X > x'),\ \forall x < x'.$$

Proof:

- $$\begin{aligned} & \mathcal{L}(Y|X=x) \leq_{\text{st}} \mathcal{L}(Y|X>x) \\ \iff & \frac{\overline{G}_X(x)}{g_X(x)} \leq \frac{\overline{G}_X(x|Y>y)}{g_X(x|Y>y)} \\ \iff & \mathcal{L}(X) \leq_{\text{hr}} \mathcal{L}(X|Y>y), \end{aligned}$$

  and this, in view of Theorem 3.3.2, is equivalent to B1$(Y|X)$.

- $$\begin{aligned} & \mathcal{L}(Y|X=x) \leq_{\text{hr}} \mathcal{L}(Y|X>x) \\ \iff & \frac{\overline{G}_Y(y'|X=x)}{\overline{G}_Y(y|X=x)} \leq \frac{\overline{G}_Y(y'|X>x)}{\overline{G}_Y(y|X>x))}, \ \forall y<y' \\ \iff & \frac{\overline{G}_X(x|Y>y)}{g_X(x|Y>y)} \leq \frac{\overline{G}_X(x|Y>y')}{g_X(x|Y>y')}, \ \forall y<y' \\ \iff & \mathcal{L}(X|Y>y) \leq_{\text{hr}} \mathcal{L}(X|Y>y'), \end{aligned}$$

  that is B2$(Y|X)$.

- $$\begin{aligned} & \mathcal{L}(Y|X=x) \leq_{\text{lr}} \mathcal{L}(Y|X>x) \\ \iff & \frac{\overline{G}_X(x|Y=y)}{g_X(x|Y=y)} \leq \frac{\overline{G}_X(x|Y=y')}{g_X(x|Y=y')} \\ \iff & \mathcal{L}(X|Y=y) \leq_{\text{hr}} \mathcal{L}(X|Y=y'). \end{aligned}$$

  That is B3$(Y|X)$. ■

We notice that all the three parts of the above proof are based on Bayes' formula and on Theorem 3.3.2.

In different frameworks, especially in the case when X, Y are non-negative random variables with the meaning of failure times, default times, or times to events, it is important to establish whether the following condition holds:

$$\mathcal{L}(Y|X=t, Y>t) \leq_* \mathcal{L}(Y|X>t, Y>t). \qquad (3.5.2)$$

The stochastic inequality (3.5.2) is a special case of *dependence properties of dynamic type* studied in the field of reliability; see in particular

[13, 14, 361, 420]. The same inequality is also closely related to the concept of *default contagion* introduced in the literature on financial risk; see, e.g., [314] and references therein. Some relations between the inequalities in Eq. (3.5.1) and in Eq. (3.5.2) are made precise by the following result.

Proposition 3.5.2. *For* $*$ = hr, lr, *condition* (3.5.1) *implies Eq.* (3.5.2)*.*

Proof:

- For $* = \mathrm{lr}$, the implication straightly follows by [426, Theorem 1.C.6].

- For $* = \mathrm{hr}$, it is a consequence of the implication:

$$\mathcal{L}(S) \leq_{\mathrm{hr}} \mathcal{L}(T) \implies \mathcal{L}(S|S > t) \leq_{\mathrm{hr}} \mathcal{L}(T|T > t), \ \forall t \geq 0.$$

■

Propositions 3.5.1 and 3.5.2, applied to the results of the previous section, lead us to link dependence properties to default contagion properties.

Corollary 3.5.3.

(i) $\mathrm{SIRL}(X|Y) \implies \mathcal{L}(Y|X = t, Y > t) \leq_{\mathrm{lr}} \mathcal{L}(Y|X > t, Y > t)$

(ii) $\mathrm{SIRL}(Y|X) \implies \mathcal{L}(X|X > t, Y = t) \leq_{\mathrm{lr}} \mathcal{L}(X|X > t, Y > t)$

(iii) $\mathrm{RCSI}(X, Y) \implies \begin{cases} \mathcal{L}(Y|X = t, Y > t) \leq_{\mathrm{hr}} \mathcal{L}(Y|X > t, Y > t) \\ \mathcal{L}(X|X > t, Y = t) \leq_{\mathrm{hr}} \mathcal{L}(X|X > t, Y > t) \end{cases}$

Remark 3.5.4. $\mathrm{RCSI}(X, Y)$ also implies default contagion in the usual stochastic order sense, i.e.,

$$\mathcal{L}(Y|X = t, Y > t) \leq_{\mathrm{st}} \mathcal{L}(Y|X > t, Y > t)$$

and

$$\mathcal{L}(X|X > t, Y = t) \leq_{\mathrm{st}} \mathcal{L}(X|X > t, Y > t).$$

3.5.2 Conditional Independence and Reversed Markov Processes

Let X, Y be conditionally independent w.r.t. a random variable Z. It is well known that, when X and Y are both stochastically increasing w.r.t. Z, in some suitable sense, then the pair (X, Y) manifests some corresponding property of positive dependence. A rich literature has been devoted to this issue, for the general case of $n \geq 2$ conditionally independent variables. See, in particular, [217, 230, 277, 429, 446].

In this section, we show some specific aspects of this topic for the case of the bivariate dependence notions appearing in the matrices $M(Y|X)$, $M(Y|X)$.

The following definition of transitivity of a dependence property is relevant in the present setting.

Definition 3.5.5. Let $\mathcal{D}$ be a dependence property. We say that $\mathcal{D}$ is *transitive under conditional independence*, in short *c.i.-transitive*, if the conditions:

(i) (X, Z) satisfies $\mathcal{D}$

(ii) (Z, Y) satisfies $\mathcal{D}$

(iii) X, Y conditionally independent given Z

imply that (X, Y) satisfies $\mathcal{D}$.

For example, the following implication is clear (under the assumption that X, Y are conditionally independent given Z): if

$$\mathcal{L}\left(Z|X = x'\right) \leq_{\mathrm{st}} \mathcal{L}\left(Z|X = x''\right) \text{ for any } x' < x''$$

and

$$\mathcal{L}\left(Y|Z = z'\right) \leq_{\mathrm{st}} \mathcal{L}\left(Y|Z = z''\right) \text{ for any } z' < z'',$$

then

$$\mathcal{L}\left(Y|X = x'\right) \leq_{\mathrm{st}} \mathcal{L}\left(Y|X = x''\right) \text{ for any } x' < x'';$$

in other words $C1(Z|X)$ and $C1(Y|X)$ imply $C1(Y|Z)$, i.e., SI is c.i.-transitive.

The notions of PQD, right tail increasing (RTI), and PLRD are also c.i.-transitive. It is also obvious, by definition of c.i.-transitivity and by Theorem 3.3.2, that a property $\mathcal{D}$ is c.i.-transitive if and only if its reversed property $\mathcal{D}^*$ is such.

Proposition 3.5.6. *Let X, Y be conditionally independent given Z. Then the following implications hold:*

(a) If (X, Z) and (Y, Z) are PQD*, then (X, Y) is* PQD*.*

(b) If (X, Z) and (Y, Z) are PLRD*, then (X, Y) is* PLRD*.*

(c) If $\mathrm{A3}(X|Z)$ *and* $\mathrm{SI}(Y|Z)$ *hold, then* $\mathrm{SI}(Y|X)$*.*

(d) If $\mathrm{B3}(X|Z)$ *and* $\mathrm{SIRL}(Y|Z)$ *hold, then* $\mathrm{SIRL}(Y|X)$*.*

Proof: The proof of (a) and (b) is immediate, by taking into account Definition 3.5.5 and the fact that PQD and PLRD are symmetric. The proof of (c) is also almost obvious. In fact, by Theorem 3.3.2, $\mathrm{A3}(X|Z)$ is equivalent to $\mathrm{C1}(Z|X)$, i.e., $\mathrm{SI}(X|Z)$. Then $\mathrm{C1}(Y|X)$ follows by c.i.-transitivity of $C1$. This is similar for item (d). ■

Items (c) and (d) of Proposition 3.5.6 are slightly different from other results given in the literature cited above. In fact, in such results, one considers random variables $T_1, \ldots, T_n$ that are conditionally independent given Z and it is typically assumed that one and the same conditional ordering holds for all the (possibly different) conditional distributions of T_j given Z. Compare in particular item (c) with [217] or [446, p. 138].

Interest of c.i.-transitive dependence properties also arises in a natural way in the analysis of real-valued Markov processes. Consider a Markov process in discrete time, $X_0, X_1, \ldots$, with transition kernel $p(x|x')$, and, for $n = 2, 3, \ldots$, let $p^{(n)}(x|x')$ be the transition kernel in n steps.

Lemma 3.5.7. *Let $p(x|x')$ satisfy a dependence property $\mathcal{D}$. If $\mathcal{D}$ is transitive, then also $p^{(n)}(x|x')$ satisfies $\mathcal{D}$, for $n = 2, 3, \ldots$.*

Proof: $p^{(2)}(x|x')$ satisfies $\mathcal{D}$ just by Definition 3.5.5. For $n = 3, 4, \ldots$, the claim follows by induction. In fact (X_0, X_2) satisfies D, (X_2, X_3) satisfies $\mathcal{D}$, and X_0, X_3 are conditionally independent w.r.t. X_2, and so on. ■

We obtain the following easy result:

Proposition 3.5.8. *Let $n = 1, 2, \ldots$.*

(a) If (X_0, X_1) is PQD*, then also (X_n, X_0) is* PQD*.*

(b) If (X_0, X_1) is PLRD*, then also (X_n, X_0) is* PLRD*.*

(c) If X_1 *is* SI *in* X_0*, then* (X_n, X_0) *satisfies* $\mathrm{A3}(X_0|X_n)$*.*

(d) If (X_0, X_1) *is* $\mathrm{SIRL}(X_1|X_0)$*, then* (X_n, X_0) *is* $\mathrm{B3}(X_0|X_n)$*.*

3.6 Discussion and Conclusions

We conclude this paper with some comments and final remarks.

Our main results are Theorems 3.2.6 and 3.3.2. These results present advantages of both theoretical and technical type. More precisely, Theorem 3.3.2 provides a unified framework for the proofs of the identity between the matrix $M(Y|X)$ and the transpose of the matrix $M(X|Y)$. Theorem 3.2.6 is a general result about the usual stochastic order that can be applied to Theorem 3.3.2, by using Propositions 3.2.9 and 3.2.10. We notice that Theorem 3.2.6 does not require any regularity condition on the probability distributions of the involved variables. In particular it does not rely on absolute continuity, thus making Theorem 3.3.2 independent of this assumption as well.

The use of Theorem 3.2.6 in the proof of Theorem 3.3.2 is made possible by Propositions 3.2.9 and 3.2.10, which allow us to express $\leq_{\mathrm{lr}}$, $\leq_{\mathrm{hr}}$ in terms of $\leq_{\mathrm{st}}$.

Theorem 3.2.6 could be similarly applied to conditional orderings involving other stochastic orders that can still be expressed in terms of $\leq_{\mathrm{st}}$. As a first instance, we can refer to the reversed hazard rate order that can be characterized in terms of $\leq_{\mathrm{st}}$ similarly to the hazard rate order. However also other stochastic orderings can be linked to $\leq_{\mathrm{st}}$ by means of different characterizations: this is the case, for example, of the mean residual life order, the harmonic mean residual life order, the convex order, and the dispersive order (see [426], in particular Theorems 2.A.4, 2.B.2, 3.A.4, and 3.B.13).

Initially our work has been inspired by the purpose of comparing the univariate stochastic orderings $\leq_{\mathrm{st}}$, $\leq_{\mathrm{hr}}$, $\leq_{\mathrm{lr}}$ in terms of differences among notions of conditional orderings. In this respect, we actually pointed out some symmetries that exist between the notions of stochastic orders $\leq_{\mathrm{st}}$, $\leq_{\mathrm{hr}}$, $\leq_{\mathrm{lr}}$ on one side and conditional orderings on the other side. Our results also provide some insight about the concepts of positive dependence. In particular, by means of Theorem 3.3.2, we can explain the superpositions between different pairs of concepts of positive dependence.

Theorems 3.2.6 and 3.3.2 are also useful to understanding the link between notions of default contagion and other notions of positive dependence properties (see Corollary 3.5.3).

All our results can be easily extended to a particular case of multivariate conditioning, i.e., in the case when conditioning events differ each other by the specification of only one variable, as it happens, for example, between

$$E_0 = \{T_1 = t_1, \dots, T_k = t_k, T_{k+1} > 0, T_{k+2} > 0, T_{k+3} > s_3, \dots, T_n > s_{n-k}\} \text{ and}$$
$$E = \{T_1 = t_1, \dots, T_k = t_k, T_{k+1} > s_1, T_{k+2} > 0, T_{k+3} > s_3, \dots, T_n > s_{n-k}\}.$$

In view of transitivity of conditioning, we can read a comparison of the kind $\mathcal{L}(T_{k+2}|E_0) \leq_* \mathcal{L}(T_{k+2}|E)$ as

$$\mathcal{L}(T_{k+2}|E_0) \leq_* \mathcal{L}((T_{k+2}|E_0)|T_{k+1} > s_1)$$

and trace it back to the bivariate case treated so far, with

$$X = T_{k+1}|E_0, \ Y = T_{k+2}|E_0.$$

Our results can be suitably extended from the bivariate to the multivariate case. Provided appropriate "weak" notions of multivariate stochastic orders are considered, such an extension can be developed along the same lines of this present paper. This will be the subject of some future work.

3.7 Appendix

This section is devoted to recalling basic definitions and theorems concerning stochastic dependence and copulas. For further details, see, e.g., [355].

We will consider the following positive dependence properties:

Definition 3.7.1.

(i) (X, Y) is PQD if

$$\overline{F}(x, y) \geq \overline{G}_X(x)\overline{G}_Y(y)$$

(ii) Y is right tail increasing in X (RTI($Y|X$)) if $\dfrac{\overline{F}(x, y)}{\overline{G}_X(x)} \uparrow x$

(iii) (X, Y) is RCSI if for $\forall\ x \leq x',\ y \leq y'$,

$$\overline{F}(x, y)\overline{F}(x', y') \geq \overline{F}(x', y)\overline{F}(x, y')$$

i.e., $\overline{F}$ is Totally Positive of order 2 (TP_2).

We will consider here only the "joint" absolutely continuous case, with f joint density. We consider then also

Definition 3.7.2.

(i) Y is stochastically increasing in X (SI($Y|X$)) if

$$\overline{G}_Y(y|X = x) \uparrow x.$$

(ii) (X, Y) is PLRD if f is TP_2.

As known, the dependence properties of a joint distribution are actually properties of its copula only, i.e., they do not involve the margins.

We recall that a copula can be seen as the distribution function of two r.v.s U, V, uniformly distributed on $[0, 1]$. In the specific, we considered here the survival copula $\hat{C}(u, v) = \overline{F}\left(\overline{G}_X^{-1}(u), \overline{G}_Y^{-1}(v)\right)$.

The dependence properties in Definitions 3.7.1 and 3.7.2 can be restated by the following:

Definition 3.7.3.

(i) (X, Y) is PQD if and only if $\hat{C}$ is PQD, i.e., if and only if $\hat{C}(u, v) \geq uv$.

(ii) RTI($Y|X$) if and only if $\hat{C}$ is LTD($V|U$), i.e., if and only if

$$\frac{\hat{C}(u, v)}{u} \downarrow u \ \ \forall v.$$

(iii) (X, Y) is RCSI if and only if $\hat{C}$ is TP_2.

(iv) SI($Y|X$) if and only if $\hat{C}$ is SI($V|U$), i.e., if and only if

$$\frac{\partial \hat{C}(u, v)}{\partial u} \downarrow u \ \ \forall v.$$

(v) (X, Y) is PLRD if and only if $\dfrac{\partial^2 \hat{C}(u, v)}{\partial u \partial v}$ is TP_2.

Acknowledgements

We thank the editors for inviting us to contribute to this volume in honor of Moshe Shaked.

Part II

Stochastic Comparison of Order Statistics

Chapter 4
Multivariate Comparisons of Ordered Data

Félix Belzunce

Abstract: In this paper we present a review for some of the main results about multivariate comparisons of ordered data.

4.1 Introduction

One of the most prolific fields where stochastic orders find their main application is the field of ordered data. The stochastic comparison of ordered data has received a great attention along the last two decades and the number of papers devoted to this topic has increased significantly. The main models of ordered data that have been considered are order statistics, record values, and generalized order statistics. Reviews of such results for order statistics can be found in Boland et al. [73, 74]. More recently the case of dependent observations has received an increasing attention and results in this direction can be found in Navarro and Shaked [352], Navarro et al. [347], Navarro [338], and

F. Belzunce (✉)
Departmento Estadística e Investigación Operativa, Universidad de Murcia,
30100 Espinardo, Murcia, Spain
e-mail: belzunce@um.es

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_4,

more recently Belzunce et al. [50]. For a review on the topic of comparisons of record values the reader may refer to Belzunce et al. [52] and results for the comparison of generalized order statistics can be found in Franco et al. [168], Belzunce et al. [54], Khaledi [227], Hu and Zhuang [204, 205], Khaledi and Kochar [234], Qiu and Wu [379], and Xie and Hu [479].

From previous papers and the references therein one can guess the great variety of results that we can find. In fact we can find results about comparisons of ordered data from one population and from two populations and comparisons for spacings of ordered data and results in the univariate and multivariate setting for these topics. This paper is intended to be a review for some of the main results about multivariate comparisons of ordered data from two populations. Of course it is not the intention of this paper to be an exhaustive review on this topic but to provide a first account of these results for the reader interested on the topic. Anyway additional references have been included for those interested on some of the other topics mentioned above. In any case I wish to apologize for those references that have not been included that for sure should be mentioned in this paper.

The organization of this paper is the following. In Sect. 4.2, I will recall the multivariate stochastic orders that we are going to consider in this paper. In Sect. 4.3 I will make a review of some models of ordered data and its relationships. In Sect. 4.4 I will recall some of the main results on the topic and to finish, in Sect. 4.5 I will provide some additional references.

In this paper "increasing" and "decreasing" mean "nondecreasing" and "nonincreasing," respectively. Also given a random variable X with continuous distribution function F and density f, $\overline{F} \equiv 1 - F$ denotes the survival function, $f/\overline{F}$ is the hazard rate, and F^{-1} denotes the inverse of the distribution function which is understood to be the left continuous one. By $\stackrel{\text{st}}{=}$ we denote equality in law.

4.2 Multivariate Orders

In this section we provide a review of the stochastic orders that we are going to consider along this paper. For a general reference on definitions and properties of stochastic orders we need to mention the two books by Shaked and Shanthikumar [422, 426] and the book by Müller and Stoyan [335].

We start by considering the usual multivariate stochastic order. Recall that given two n-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, then $\boldsymbol{X}$ is said to be less than $\boldsymbol{Y}$ in the multivariate stochastic order (denoted by $\boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y}$) if $\mathrm{E}[\phi(\boldsymbol{X})] \leq \mathrm{E}[\phi(\boldsymbol{Y})]$, for all increasing function $\phi : \mathbb{R}^n \mapsto \mathbb{R}$, provided the previous expectations exist. In the univariate case given two random variables X and Y with distribution functions F and G, respectively, $X \leq_{\text{st}} Y$ if, and only if, $\overline{F} \leq \overline{G}$. This partial order, in the univariate case, can be extended to a weaker criteria replacing increasing functions by increasing convex functions. Given two random variables X and Y, we say that X is less than Y in the increasing convex order, denoted by $X \leq_{\text{icx}} Y$, if

$$\mathrm{E}[\phi(X)] \leq \mathrm{E}[\phi(Y)],$$

for all convex increasing convex functions ϕ, for which the involved expectations exist. In the multivariate setting, there are several possible ways to extend this concept, depending on the kind of convexity that we consider.

Given two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, we say that $\boldsymbol{X}$ is less than $\boldsymbol{Y}$ in the multivariate increasing convex order, denoted by $\boldsymbol{X} \leq_{\text{icx}} \boldsymbol{Y}$, if

$$\mathrm{E}[\phi(\boldsymbol{X})] \leq \mathrm{E}[\phi(\boldsymbol{Y})], \tag{4.2.1}$$

for all increasing convex functions $\phi : \mathbb{R}^n \mapsto \mathbb{R}$, for which the involved expectations exist. It is clear from the definition that $\boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y} \Longrightarrow \boldsymbol{X} \leq_{\text{icx}} \boldsymbol{Y}$.

Some other suitable classes of functions defined on $\mathbb{R}^n$ can also be considered to extend convex orders to the multivariate case by means of a difference operator. To be specific, let Δ_i^ε be the ith difference operator defined for a function $\phi : \mathbb{R}^n \to \mathbb{R}$ as

$$\Delta_i^\varepsilon \phi(\boldsymbol{x}) = \phi(\boldsymbol{x} + \varepsilon \boldsymbol{e}_i) - \phi(\boldsymbol{x}),$$

where $\boldsymbol{e}_i = (0, \ldots, 0, \overbrace{1}^{i}, 0, \ldots, 0)$. A function ϕ is said to be directionally convex if $\Delta_i^\varepsilon \Delta_j^\delta \phi(\boldsymbol{x}) \geq 0$ for all $1 \leq i \leq j \leq n$ and $\varepsilon, \delta \geq 0$. We observe that directionally convex functions are also known as ultramodular functions; see, for example, Marinacci and Montrucchio [305]. A function ϕ is said to be supermodular if $\Delta_i^\varepsilon \Delta_j^\delta \phi(\boldsymbol{x}) \geq 0$ for all $1 \le i < j \le n$ and $\varepsilon, \delta \ge 0$. If ϕ is twice differentiable, then it is directionally convex if $\partial^2\phi/\partial x_i \partial x_j \ge 0$ for every $1 \le i \le j \le n$, and it is supermodular if $\partial^2\phi/\partial x_i \partial x_j \ge 0$ for every $1 \le i < j \le n$. Clearly a function ϕ is directionally convex if it is supermodular and it is componentwise convex.

When we consider increasing directionally convex functions in Eq. (4.2.1), then we say that $\boldsymbol{X}$ is less than $\boldsymbol{Y}$ in the increasing directionally convex order, denoted by $\boldsymbol{X} \le_{\text{idir-cx}} \boldsymbol{Y}$. The increasing directionally convex order is weaker than the multivariate increasing convex order, i.e., $\boldsymbol{X} \le_{\text{icx}} \boldsymbol{Y} \Longrightarrow \boldsymbol{X} \le_{\text{idir-cx}} \boldsymbol{Y}$. Note that the increasing directionally convex order compares not only the dependence structures of two random vectors but also the variability of the marginals.

Let us consider now one order stronger than the stochastic order, which is of interest mainly in reliability theory. First we give the definition in the univariate case. Given two nonnegative random variables X and Y with hazard rates r and s, respectively, then X is said to be less than Y in the hazard rate order (denoted by $X \le_{\text{hr}} Y$) if $r(t) \ge s(t)$ for all $t \ge 0$. This order can be characterized in terms of the survival functions as follows: $X \le_{\text{hr}} Y$ if, and only if, $\overline{F}/\overline{G}$ is decreasing, where $\overline{F}$ and $\overline{G}$ are the survival functions of X and Y, respectively. This characterization allows the definition for any pair of random variables. To extend the definition to the multivariate case, we need first to consider the notion of multivariate conditional hazard rate functions. Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a nonnegative n-dimensional random vector. For $t \ge 0$ let h_t denote the event

$$h_t = \{\boldsymbol{X}_I = \boldsymbol{t}_I, \boldsymbol{X}_{\bar{I}} > t\boldsymbol{e}\},$$

where $\boldsymbol{e} = (1, \ldots, 1)$, $I = \{i_1, \ldots, i_k\}$ is a subset of $\{1, \ldots, n\}$, $\bar{I}$ is its complement with respect to $\{1, \ldots, n\}$, $\boldsymbol{X}_I$ and $\boldsymbol{t}_I$ denote the vectors formed by the components of $\boldsymbol{X}$ and $\boldsymbol{t}$ with indices in I, and $0 < t_{i_j} \le t$ for all $j = 1, \ldots, k$. The event h_t is called a *history at the point* t of the random vector $\boldsymbol{X}$. Given a history h_t as above, and an $i \in \bar{I}$, we define its *multivariate conditional hazard rate*, at the point t, as follows:

$$\lambda_{i|I}(t|\boldsymbol{t}) = \lim_{\Delta t \to 0+} \frac{1}{\Delta t} \mathrm{P}\{t < X_i \le t + \Delta t | h_t\}.$$

Now let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two n-dimensional random vectors with multivariate conditional hazard rate functions $\eta_{\cdot|\cdot}(\cdot|\cdot)$ and $\lambda_{\cdot|\cdot}(\cdot|\cdot)$, respectively. We say that $\boldsymbol{X}$ is smaller than $\boldsymbol{Y}$ in the multivariate hazard rate order, denoted by $\boldsymbol{X} \le_{\text{hr}} \boldsymbol{Y}$, if

$$\eta_{i|I\cup J}(x|\boldsymbol{s}_{I\cup J}) \geq \lambda_{i|I}(x|\boldsymbol{t})$$

whenever $I \cap J = \varnothing$, $\mathbf{0} \leq \boldsymbol{s}_I \leq \boldsymbol{t} \leq x\boldsymbol{e}$, and $\mathbf{0} \leq \boldsymbol{s}_J \leq x\boldsymbol{e}$, where $i \in \overline{I \cup J}$.

In some situations it is not possible to provide an explicit expression for the distribution function and therefore is not possible to check some of the previous orders. An alternative is to use the density functions (or the probability mass function in the case of discrete random variables) to compare two random variables. Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two n-dimensional random vectors with joint density functions f and g, respectively. We say that $\boldsymbol{X}$ is less than $\boldsymbol{Y}$ in the multivariate likelihood ratio order, denoted by $\boldsymbol{X} \leq_{\text{lr}} \boldsymbol{Y}$, if

$$f(\boldsymbol{x})g(\boldsymbol{y}) \leq f(\boldsymbol{x} \wedge \boldsymbol{y})g(\boldsymbol{x} \vee \boldsymbol{y})$$

for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in $\mathbb{R}^n$, where $\wedge$ and $\vee$ denote the componentwise minimum and maximum operations, respectively.

In the univariate case, given two random variables X and Y with density functions f and g, respectively, we say that X is less than Y in the likelihood ratio order, denoted by $X \leq_{\text{lr}} Y$, if $f(t)g(s) \leq f(s)g(t)$ for all $s < t \in \mathbb{R}$.

Among the likelihood ratio, hazard rate, and stochastic orders we have the following implications:

$$\boldsymbol{X} \leq_{\text{lr}} \boldsymbol{Y} \Longrightarrow \boldsymbol{X} \leq_{\text{hr}} \boldsymbol{Y} \Longrightarrow \boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y}.$$

Now we finish considering a comparison in variability through the dispersive order. Given two random variables X and Y with distribution functions F and G, we say that X is smaller than Y in the *dispersive order*, denoted by $X \leq_{\text{disp}} Y$, if $F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$ whenever $0 \leq \alpha \leq \beta \leq 1$. That is, the dispersive order requires that the difference between any two quantiles of X be smaller than the difference between corresponding quantiles of Y.

There does not exist, in the literature, a unique extension of the definition of the univariate dispersive order to the multivariate case. Rather, one can find in the literature different multivariate extensions that generalize different useful characterizations of the univariate dispersive order to the multivariate case. Here we describe two such extensions. First, in the multivariate case, given an n-dimensional random vector $\boldsymbol{X}$, and a $\boldsymbol{u} = (u_1, \ldots, u_n)$ in $[0,1]^n$, one way to define the multivariate $\boldsymbol{u}$-quantile for $\boldsymbol{X}$, denoted as $\hat{x}(\boldsymbol{u}) = (\hat{x}_1(u_1), \hat{x}_2(u_1, u_2), \ldots, \hat{x}_n(u_1, \ldots, u_n))$, is as follows:

$$\hat{x}_1(u_1) \equiv F_1^{-1}(u_1)$$

and

$$\hat{x}_i(u_1, \ldots, u_i) \equiv F_{i|1,\ldots,i-1}^{-1}(u_i) \quad \text{for } i = 2, \ldots, n,$$

where $F_{i|1,\ldots,i-1}^{-1}$ is the quantile function of the random variable $[X_i | X_1 = \hat{x}_1(u_1), \ldots, X_{i-1} = \hat{x}_{i-1}(u_1, \ldots, u_{i-1})]$.

This known construction is widely used in simulation theory and is named the *standard construction*. It satisfies most of the important properties of the quantile function in the univariate case.

Given two n-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, with standard constructions $\hat{x}(\boldsymbol{u})$ and $\hat{y}(\boldsymbol{u})$, respectively, Shaked and Shanthikumar [425] studied the following condition:

$$\hat{y}(\boldsymbol{u}) - \hat{x}(\boldsymbol{u}) \quad \text{is increasing in } \boldsymbol{u} \in (0,1)^n \tag{4.2.2}$$

as a multivariate generalization of the dispersive order. We will say that $\boldsymbol{X}$ is smaller than $\boldsymbol{Y}$ in the variability order, denoted by $\boldsymbol{X} \leq_{\text{var}} \boldsymbol{Y}$, if Eq. (4.2.2) holds.

Another multivariate generalization based on the standard construction was given by Fernández-Ponce and Suárez-Llorens [161]. We will say that $\boldsymbol{X}$ is smaller than $\boldsymbol{Y}$ in the dispersive order, denoted by $\boldsymbol{X} \leq_{\text{disp}} \boldsymbol{Y}$, if

$$\|\hat{x}(\boldsymbol{v}) - \hat{x}(\boldsymbol{u})\| \leq \|\hat{y}(\boldsymbol{v}) - \hat{y}(\boldsymbol{u})\|$$

for all $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^n$, where $\| \cdot \|$ denotes the Euclidean norm.

Shaked and Shanthikumar [425] introduce condition (4.2.2) to identify pairs of multivariate functions of random vectors that are ordered in the st:icx order. Given two random variables X and Y, we say that X is smaller than Y in the st:icx order, denoted by $X <_{\text{st:icx}} Y$, if $\mathrm{E}[h(X)] \leq \mathrm{E}[h(Y)]$ for all increasing functions h for which the expectations exist (i.e., if $X \leq_{\text{st}} Y$) and if, for all increasing convex functions h,

$$\mathrm{Var}[h(X)] \leq \mathrm{Var}[h(Y)],$$

provided the variances exist.

The result was given for conditionally increasing in sequence (CIS) random vectors. A random vector $(X_1, \ldots, X_n)$ is said to be *CIS* if, for $i = 2, 3, \ldots, n$,

$$(X_i \mid X_1 = x_1, \ldots, X_{i-1} = x_{i-1}) \leq_{\text{st}} (X_i \mid X_1 = x_1', \ldots, X_{i-1} = x_{i-1}')$$

whenever $x_j \leq x'_j$, $j = 1, 2, \ldots, i-1$.

Additionally we will use along this paper the CI and conditionally increasing in quantile (CIQ) dependence notions. We say that the random vector $(X_1, \ldots, X_n)$ is *conditionally increasing* (CI) if and only if the random vector $(X_{\pi(1)}, \ldots, X_{\pi(n)})$ is CIS for all permutations π of $\{1, 2, \ldots, n\}$.

The random vector $(X_1, \ldots, X_n)$ is said to be *CIQ* , if the standard construction $\hat{x}(\boldsymbol{u})$ is increasing in $\boldsymbol{u} \in (0,1)^n$.

Shaked and Shanthikumar [425] prove that given two nonnegative n-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, with the CIS property, if $\mathbf{X} \leq_{\text{var}} \boldsymbol{Y}$ then

$$\phi(\boldsymbol{X}) \leq_{\text{st:icx}} \phi(\boldsymbol{Y}),$$

for all increasing directionally convex function ϕ.

Recently Belzunce et al. [56] have proved that, for random vectors with the same CIQ copula,

$$\boldsymbol{X} \leq_{\text{disp}} \boldsymbol{Y} \iff \boldsymbol{X} \leq_{\text{var}} \boldsymbol{Y} \iff X_i \leq_{\text{disp}} Y_i, \text{ for all } i = 1, \ldots, n.$$

4.3 Multivariate Models with Ordered Components

Perhaps the most well-known model of a random vector with ordered components is the random vector of order statistics. This model arises in natural way when we arrange in increasing order a set of observations from a random variable. Another example is the case of epoch times of a counting process, like the case of a nonhomogeneous Poisson process. Epoch times of nonhomogeneous Poisson processes can be introduced as record values of a proper sequence of random variables, which is another typical example of ordered data. Given the similarity of several results for order statistics and record values Kamps [221] introduces the model of generalized order statistics. This model provides a unified approach to study order statistics and record values and several other models of ordered data.

First we provide the definition of generalized order statistics following Kamps [221, 222]:

Definition 4.3.1. Let $n \in \mathbb{N}$, $k \geq 1$, $m_1, \ldots, m_{n-1} \in \mathbb{R}$, $M_r = \sum_{j=r}^{n-1} m_j$, $1 \leq r \leq n-1$ be parameters such that $\gamma_r = k+n-r+M_r \geq 1$

for all $r \in 1, \ldots, n-1$, and let $\tilde{m} = (m_1, \ldots, m_{n-1})$, if $n \geq 2$ ($\tilde{m} \in \mathbb{R}$ arbitrary, if $n = 1$). We call uniform generalized order statistics to the random vector $(U_{(1,n,\tilde{m},k)}, \ldots, U_{(n,n,\tilde{m},k)})$ with joint density function

$$h(u_1, \ldots, u_n) = k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} (1 - u_j)^{m_j} \right) (1 - u_n)^{k-1}$$

on the cone $0 \leq u_1 \leq \cdots \leq u_n \leq 1$. Now given a distribution function F we call generalized order statistics based on F to the random vector

$$(X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)}) \equiv \left(F^{-1}(U_{(1,n,\tilde{m},k)}), \ldots, F^{-1}(U_{(n,n,\tilde{m},k)}) \right).$$

If F is an absolutely continuous distribution with density f, the joint density function of $(X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)})$ is given by

$$f(x_1, \ldots, x_n) = k \left(\prod_{j=1}^{n-1} \gamma_j \right) \left(\prod_{j=1}^{n-1} \overline{F}(x_j)^{m_j} f(x_j) \right) \overline{F}(x_n)^{k-1} f(x_n)$$

on the cone $F^{-1}(0) \leq x_1 \leq \cdots \leq x_n \leq F^{-1}(1)$.

In the special case when $m_1 = \cdots = m_{n-1} = m$, the variables $(X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)})$ are called m-GOSs and are denoted by $(X_{(1,n,m,k)}, \ldots, X_{(n,n,m,k)})$.

Among the different distributional properties we recall a property about the copula of a random vector of GOSs. A copula C is a cumulative distribution function with uniform marginals on $[0, 1]$. Furthermore, it has been shown that if F is an n-dimensional distribution function with marginal distribution functions $F_1, \ldots, F_n$, then there exists a copula C such that for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$, we have $F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$. Moreover, if $F_1, \ldots, F_n$ are continuous, then C is unique. For details on various properties of copulas, interested readers may refer to Nelsen [355]. For GOSs we have that two random vectors of GOSs with the same set of parameters and possibly based on different distributions have the same copula (see Belzunce et al. [56]).

Let us see now several models that are included in this model. As we have mentioned previously, order statistics and record values are a particular case of this model.

Taking $m_i = 0$ for all $i = 1, \ldots, n-1$ and $k = 1$ we get the random vector of order statistics $(X_{1:n}, X_{2:n}, \ldots, X_{n:n})$ from a set of n independent and identically distributed (i.i.d) observations $X_1, X_2, \ldots, X_n$

with common absolutely continuous distribution F, in particular, we get that $X_{i:n} \stackrel{\text{st}}{=} X_{(i,n,0,1)}$. However, the model of GOSs does not include order statistics when we remove the assumption of independent and/or identically distributed observations. If we consider the random vector of order statistics $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ from n independent but not necessarily identically distributed observations $X_1, X_2, \dots, X_n$, where, for all $i = 1, 2, \dots, n$, X_i has absolutely continuous distribution with density function f_i and distribution function F_i, the joint density function of $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ is given by (see Vaughan and Venables [461])

$$\begin{aligned} f_{1,2,\cdots,n:n}(x_1, x_2, \dots, x_n) &= \sum_{\pi \in \mathcal{P}} \prod_{i=1}^{n} f_{\pi_i}(x_i), \\ &\qquad \text{for all } x_1 < x_2 < \cdots < x_n, \end{aligned}$$

where $\mathcal{P}$ denotes the set of all $n!$ permutations of $\{1, \dots, n\}$ and $\pi = (\pi_1, \dots, \pi_n)$ is one specific permutation.

If we consider order statistics $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ from a random vector of n possibly dependent observations $(X_1, X_2, \dots, X_n)$, with absolutely continuous distribution and joint density function f, the joint density of $(X_{1:n}, X_{2:n}, \dots, X_{n:n})$ is given by

$$\begin{aligned} f_{1,2,\cdots,n:n}(x_1, x_2, \dots, x_n) &= \sum_{\pi \in \mathcal{P}} f(x_{\pi_1}, x_{\pi_2}, \dots, x_{\pi_n}), \\ &\qquad \text{for all } x_1 < x_2 < \cdots < x_n. \end{aligned}$$

Let us consider the case of record values. Chandler [83] introduced the mathematical notion of record values to study, from a statistical point of view, sequences of record values that arise in practice. Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables, which can be considered as independent observations from the random variable X. Denote the cumulative distribution function of X by F, and assume that F is absolutely continuous. Also, denote by f the corresponding density function. Record values are defined by means of record times, so first let us recall the definition of record times. Given a sequence of i.i.d. random variables as above, the *record times* are the random variables

$$\begin{aligned} &L(1) = 1, \\ &L(n) = \min\{j > L(n-1) | X_j > X_{L(n-1)}\}, \quad n = 2, 3, \dots. \end{aligned}$$

The sequence of *record values* $X(n)$ is defined by

$$X(n) \equiv X_{L(n)}, \quad n = 1, 2, \ldots .$$

Taking $m_i = -1$ for all $i = 1, \ldots, n-1$ and $k = 1$ we get that $X(i) =_{st} X_{(i,n,-1,1)}$.

As mentioned above record values are related to the epoch times of a nonhomogeneous Poisson process. A counting process $\{N(t),\ t \geq 0\}$ is a *nonhomogeneous Poisson process* with intensity (or rate) function $r \geq 0$ if:

(a) $\{N(t),\ t \geq 0\}$ has the Markov property.

(b) $\mathrm{P}\{N(t + \Delta t) = n + 1 | N(t) = n\} = r(t)\Delta t + o(\Delta t)$, $n \geq 0$.

(c) $\mathrm{P}\{N(t + \Delta t) > n + 1 | N(t) = n\} = o(\Delta t)$, $n \geq 0$.

We assume that

$$\int_t^\infty r(u)\,\mathrm{d}u = \infty \quad \text{for all } t \geq 0; \tag{4.3.1}$$

this ensures that with probability 1 the process has a jump after any time point t.

A nonnegative function r which satisfies Eq. (4.3.1) can be interpreted as the hazard rate function of a random variable, and if $X_1, X_2, \ldots$ is a sequence of i.i.d. random variables with hazard rate function r then the epoch time of the nth jump, T_n, say, satisfies (see Gupta and Kirmani [185])

$$T_n \stackrel{\mathrm{st}}{=} X(n).$$

An extension of this counting process, where the epoch times are not included in the model of GOSs, is the nonhomogeneous pure birth process (NHPB). Such processes are called "relevation counting processes" in Pellerey et al. [372]. Epoch times of NHPB processes correspond also to Pfeifer's record values. Pfeifer [373] establishes a model of record values based on a double sequence of nonidentically distributed random variables.

A NHPB process is a Markovian point process in which the jump intensity at any time t depends not only on t but also on the number n of jumps before time t. Formally, we consider a point process $N = \{N(t),\, t \geq 0\}$ such that

(a) N has the Markov property

(b) $P\{N(t+\Delta t) = n+1 \big| N(t) = n\} = r_n(t)\Delta t + o(\Delta t)$. $n \geq 1$

(c) $P\{N(t+\Delta t) > n+1 \big| N(t) = n\} = o(\Delta t)$, $n \geq 1$.

where the r_n's are nonnegative functions that satisfy

$$\int_t^{\infty} r_n(x)\mathrm{d}x = \infty \text{ for all } t \geq 0. \tag{4.3.2}$$

Again previous condition ensures that r_n can be considered as a failure rate of a proper distribution. Now denote by $\overline{F}_n$ and f_n the survival and the density functions associated to r_n, then the joint density of $\boldsymbol{T} = (T_1, \ldots, T_n)$ is given by

$$h(x_1, x_2, \ldots, x_n) = \prod_{j=1}^{n-1} \frac{f_j(x_j)}{\overline{F}_{j+1}(x_j)} f_n(x_n) \quad \text{for } x_1 \leq x_2 \leq \cdots \leq x_n.$$

In a distributional theoretical sense, GOSs are contained in the model of epoch times of an NHPB process. Consider GOSs based on F with failure rate r and parameters k, n, and M_r, $r = 1, \ldots, n-1$, then

$$(X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)}) \stackrel{\text{st}}{=} (T_1, \ldots, T_n)$$

where T_i are the epoch times of a NHPB process with intensities $r_i = (k+n-i+M_i)r$, for $i : 1, \ldots, n$.

We finish with some additional particular cases of GOSs.

A generalization of the record values is the case in which $k \in \mathbb{N}$, obtaining what is called k-records.

Let $\{X_i\}_{i\in\mathbb{N}}$ be a sequence of i.i.d random variables with distribution function F and let $k \in \mathbb{N}$. Denoting by

$$(X_{j-k+1,j})_{j=k,k+1,\ldots},$$

the sequence of k-th largest order statistics, the random variables given by

$$\begin{aligned} &L^k(1) = 1 \\ &L^k(n+1) = \min\{j > L(n) : X_{j,j+k-1} > X_{L^k(n),L^k(n)+k-1}\} \end{aligned}$$

are called the record times, and the quantities

$$X_{L^k(n)} \equiv X_{L^k(n),L^k(n)+k-1}$$

are called k-th records or k-records, which are a special case of GOSs taking $m_i = -1$ for all $i = 1, \ldots, n-1$ and $k \in \mathbb{N}$.

A life-testing experiment of interest in reliability studies involves N independent and identically distributed random variables placed simultaneously on test and at the time of the m-th failure R_i surviving units are randomly censored from the test. The progressively type II censored order statistics arising from such a reliability experiment can be obtained from the model of GOSs by setting $n = m$, $m_i = R_i$, and $k = R_m + 1$.

An interesting model contained in the model of generalized order statistics is that of order statistics under multivariate imperfect repair; see Shaked and Shanthikumar [418]. Suppose n items start to function at the same time 0. Upon failure, an item undergoes a repair. If i items ($i = 0, 1, \ldots, n-1$) have already been scrapped, then, with probability p_{i+1}, the repair is unsuccessful and the item is scrapped, and with probability $1 - p_{i+1}$, the repair is successful and minimal.

Let us now consider n items with i.i.d. random lifetimes $X_1, \ldots, X_n$, with the same distribution F and density function f. Let $(X_{(1)}, \ldots, X_{(n)})$ be the ordered random lifetimes resulting from $X_1, \ldots, X_n$ under such a minimal repair policy. Then, the joint density function of $(X_{(1)}, \ldots, X_{(n)})$ is given by

$$f(t_1, \ldots, t_n) = n! \prod_{j=1}^{n} p_j f(t_j) \left[\overline{F}(t_j)\right]^{(n-j+1)p_j - (n-j)p_{j+1} - 1} \text{ for } 0 \le t_1 \le \cdots \le t_n.$$

It is evident that this is a particular case of the joint density function of generalized order statistics based on F for the choice of parameters $k = p_n$ and $m_j = (n-j+1)p_j - (n-j)p_{j+1} - 1$.

We do not want to finish this section without mentioning some additional applications of these models where the results provided in the next section can be applied. The main applications of order statistics are in statistics and reliability. In statistics, order statistics arise in natural way when we considered ordered data and reliability order statistics arise as a model for k-out-of-n systems. Given a system of n components we say that the system is a k-out-of-n system if the system works if and only if at least k of the components function, i.e., it works as long as at most $n-k$ components have failed. If we denote by $X_1, \ldots, X_n$ the random lifetimes of the components, then the random lifetime of the k-out-of-n system is given by the order statistics $X_{n-k+1:n}$. Other applications arise in auction theory and risk theory.

Record values appear in a natural way in real life, for example, records in sports and record values related to natural phenomena, like in meteorology and hydrology. In reliability theory, the times of repair of an item, which is being continuously minimally repaired, are the epoch times of a nonhomogeneous Poisson process. The nonhomogeneous Poisson process is also a common useful model in the study of software reliability growth. Applications of NHPB processes in insurance, reliability theory, epidemiology, and load-sharing models can be found in Pellerey et al. [372] and in Belzunce et al. [51].

4.4 Multivariate Comparisons of Ordered Data

In this section we make a review of several results about stochastic comparisons of random vectors of ordered data.

We start by considering results for the multivariate stochastic order. Next result can be obtained from a characterization of the multivariate stochastic order by construction on the same probability space (see Belzunce et al. [54]). Another easy proof is the following: It is known that the stochastic order of the marginals of two random vectors with the same copula implies the multivariate stochastic order for the two random vectors. If we consider two random vectors of GOSs with the same set of parameters, it is not difficult to show that if the underlying distributions of the two sets of GOSs are ordered in the stochastic order, then the marginals of the corresponding GOSs are ordered in the same sense; therefore, given that the two random vectors have the same copula, as mentioned in previous section, we can obtain the following result:

Theorem 4.4.1. *Let X and Y be random variables with distribution functions F and G, respectively, and let $\boldsymbol{X}=(X_{(1,n,\tilde{m},k)},\ldots,X_{(n,n,\tilde{m},k)})$ and $\boldsymbol{Y}=(Y_{(1,n,\tilde{m},k)},\ldots,Y_{(n,n,\tilde{m},k)})$ be random vectors of generalized order statistics based on F and G, respectively. If $X\leq_{\rm st}Y$ then $\boldsymbol{X}\leq_{\rm st}\boldsymbol{Y}$.*

Therefore this result extends previous results for order statistics in the i.i.d. case and record values (see Belzunce et al. [51]) and provides new results for the particular cases described in previous section. A result for the comparison of random vectors of order statistics, from possibly dependent and/or not identically distributed observations, follows easily given that the arrangement in increasing order of any set of observations is an increasing function. Therefore by the preservation of stochastic order under increasing transformations we have the following result.

Theorem 4.4.2. *Given two sets of random variables* $X_1, \ldots, X_n$ *and* $Y_1, \ldots, Y_n$*, if, for two permutations* π *and* π' *of* $\{1, \ldots, n\}$, $(X_{\pi_1}, \ldots, X_{\pi_n}) \leq_{st} (Y_{\pi'_1}, \ldots, Y_{\pi'_n})$*, then we have*

$$(X_{1:n}, \ldots, X_{n:n}) \leq_{st} (Y_{1:n}, \ldots, Y_{n:n}).$$

It is also possible to provide a result for the case of epoch times of NHPB processes (or Pfeiffer's record values). Let us consider two such processes, indexed by $i = 1, 2$, parameterized by the sets $\{r_{i,n},\ n \geq 1\}$ of hazard rate functions that satisfy Eq. (4.3.2). The corresponding epoch times will be denoted by $0 \equiv T_{i,0} \leq T_{i,1} \leq T_{i,2} \leq \cdots$. Let $\{X_{i,n},\ n \geq 1\}$, $i = 1, 2$, be two sets of independent absolutely continuous nonnegative random variables, where $X_{i,n}$ has the hazard rate function $r_{i,n}$.

Next we recall a result which gives conditions under which the epoch times of the two NHPB processes are ordered according to the usual stochastic order.

Theorem 4.4.3. *Let* $T_{i,n}$*'s be the epoch times of the two NHPB processes as described above. If* $X_{1,1} \leq_{\text{st}} X_{2,1}$ *and if* $X_{1,j} \leq_{\text{hr}} X_{2,j}$ *for* $j \geq 2$ *then*

$$(T_{1,1}, T_{1,2}, \ldots, T_{1,n}) \leq_{\text{st}} (T_{2,1}, T_{2,2}, \ldots, T_{2,n}), \qquad n \geq 1.$$

A proof of this result can be found in Belzunce et al. [51]. Using some general ideas from Shaked and Szekli [430] it is possible to construct an alternative, though lengthier, proof of this theorem.

The proof of the stochastic order, for the case of GOSs, based on the fact the two random vectors of GOSs share the same copula, is a very useful tool to provide results in this setting. In fact, based on this idea, Balakrishnan et al. [28] provide a result for the comparison in the idir-cx order of two random vectors of GOSs. The proof is based on the fact that, given two random vectors with a common CI copula, if the marginals are ordered in the increasing convex order, then the two random vectors are ordered in the increasing directionally convex order (see Balakrishnan et al. [28]). In the case of two random vectors of GOSs, with the same set of parameters, it is easy to see that the common copula is CI (in fact has the stronger MTP2 dependence property), so it only remains to provide conditions for the increasing convex order of the marginals. A result in this direction has been provided by Balakrishnan et al. [28] where they prove that a sufficient condition for the increasing convex order of GOSs, is the increasing convex order of the minimums. Therefore it is possible to provide the following result:

Theorem 4.4.4. *Let X and Y be two continuous random variables with distribution functions F and G, respectively. Let random vectors $\boldsymbol{X} = \left(X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)}\right)$ and $\boldsymbol{Y} = \left(Y_{(1,n,\tilde{m},k)}, \ldots, Y_{(n,n,\tilde{m},k)}\right)$ be generalized order statistics from F and G, respectively, with $m_i \geq -1$ for all i. If $X_{(1,n,\tilde{m},k)} \leq_{\text{icx}} Y_{(1,n,\tilde{m},k)}$, then $\boldsymbol{X} \leq_{\text{idir-cx}} \boldsymbol{Y}$.*

Again, as a particular case, we get results for the several models included in the model of GOSs. Of particular interest is the case of record values. In this case, the first component is equally distributed as the distribution from which the record values are arising from. Consequently, the increasing convex order of the distributions on which the two random vectors are based on is a sufficient condition for the comparison of the vector and in particular for the increasing convex order of record values from the two populations. It would be also interesting to see if this kind of arguments can be applied for the case of epoch times of NHPB processes. In this case, given that the vector of epoch times can be shown to be CI, we would need to check if two of these random vectors share the same copula and to provide conditions for the increasing convex order of the marginals. Belzunce and Shaked [57] provide a result for the increasing convex order of the marginals, and therefore, it only remains to be seen if the random vector of epoch times shares the same copula. To provide conditions under which two random vectors of epoch times of NHPB processes have the same copula is an open problem for future research.

Next we recall some results for the hazard rate order (see Belzunce et al. [54]).

Theorem 4.4.5. *Let X and Y be absolutely continuous random variables with distribution functions F and G, respectively, and let $\boldsymbol{X} = (X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)})$ and $\boldsymbol{Y} = (Y_{(1,n,\tilde{m},k)}, \ldots, Y_{(n,n,\tilde{m},k)})$ be random vectors of generalized order statistics based on F and G, respectively. If $X \leq_{\text{hr}} Y$, then $\boldsymbol{X} \leq_{\text{hr}} \boldsymbol{Y}$.*

There is no extension of this result to the case of order statistics from possibly dependent and/or not identically distributed observations. An extension to the case of epoch times of NHPB processes can be found in Belzunce et al. [51]; in fact, previous result can be derived from the next result, taking into account the relationship among GOSs and NHPB processes described in the previous section.

Theorem 4.4.6. *Let $T_{i,n}$ and $X_{i,n}$ be as in Theorem 4.4.3. If $X_{1,j} \leq_{\text{hr}} X_{2,j}$ for $j \geq 1$ then $(T_{1,1}, T_{1,2}, \ldots, T_{1,n}) \leq_{\text{hr}} (T_{2,1}, T_{2,2}, \ldots, T_{2,n})$ for all $n \geq 1$.*

Next we consider several results for the multivariate likelihood ratio order. The proof of the following result can be seen in Belzunce et al. [54]:

Theorem 4.4.7. *Let X and Y be absolutely continuous random variables with hazard rates r and s, respectively. Let two random vectors $\boldsymbol{X} = (X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)})$ and $\boldsymbol{Y} = (Y_{(1,n,\tilde{m},k)}, \ldots, Y_{(n,n,\tilde{m},k)})$ be generalized order statistics based on F and G, respectively. If either*

(a) $m_i \geq 0$ for all i, and $X \leq_{\text{lr}} Y$ or

(b) $m_i \geq -1$ for all i, $X \leq_{\text{hr}} Y$ and s/r is increasing

then $\boldsymbol{X} \leq_{\text{lr}} \boldsymbol{Y}$.

Given that the multivariate likelihood ratio order is preserved under marginalization, a consequence of this result is that of the likelihood ratio order of the marginals. When we compare marginals there are a lot of papers that have been devoted to the comparison of GOSs with different indexes and from different dimensions. Recently Balakrishnan et al. [27] provide a multivariate result in this direction which provides as a consequence results for the comparison of the marginals in the likelihood ratio order. In particular they provide a multivariate likelihood ratio ordering result for subsets of m-GOSs that we recall next.

Theorem 4.4.8. *Let X and Y be absolutely continuous variables with distributions F and G and with hazard rates r and s, respectively. Let $\boldsymbol{X} = (X_{(1,n,m,k)}, \ldots, X_{(n,n,m,k)})$ and $\boldsymbol{Y} = (Y_{(1,n',m,k)}, \ldots, Y_{(n',n',m,k)})$ be random vectors of m-GOSs based on distributions F and G, respectively. For $r_1 \leq r_2 \leq \cdots \leq r_i \leq n$, $r'_1 \leq r'_2 \leq \cdots \leq r'_i \leq n'$, $r'_1 - r_1 = r'_2 - r_2 = \cdots = r'_i - r_i \geq \max\{0, n' - n\}$, if either*

(a) $X \leq_{\text{lr}} Y$ *and* $m \geq 0$ *or*

(b) $X \leq_{\text{hr}} Y$, s/r *is increasing, and* $-1 \leq m < 0$

then $(X_{(r_1,n,m,k)}, X_{(r_2,n,m,k)}, \ldots, X_{(r_i,n,m,k)}) \leq_{\text{lr}} (Y_{(r'_1,n',m,k)}, Y_{(r'_2,n',m,k)}, \ldots, Y_{(r'_i,n',m,k)})$.

Now we consider a different problem. We consider stochastic comparisons of two random vectors of GOSs with different parameters but based on the same distribution F.

Theorem 4.4.9. *Let* $\boldsymbol{X} = (X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)})$ *and* $\boldsymbol{X}' = (X_{(1,n,\tilde{m'},k')}, \ldots, X_{(1,n,\tilde{m'},k')})$ *be random vectors of generalized order statistics based on the same distribution* F *and with parameters* k, m_i, $i : 1\ldots, n-1$ *and* k', m'_i, $i : 1\ldots, n-1$, *respectively. Then* $\boldsymbol{X} \leq_{\text{lr}} \boldsymbol{X}'$ *if, and only if,* $k \geq k'$ *and* $m_i \geq m'_i$, *for all* $i : 1, \ldots, n-1$.

Now a combination of Theorems 4.4.9 and 4.4.1, 4.4.4, 4.4.5, and 4.4.7, leads to the following result:

Theorem 4.4.10. *Let* X *and* Y *be absolutely continuous random variables with distribution functions* F *and* G, *respectively, and let* $\boldsymbol{X} = (X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)})$ *and* $\boldsymbol{Y} = (Y_{(1,n,\tilde{m}',k')}, \ldots, Y_{(n,n,\tilde{m}',k')})$ *be random vectors of generalized order statistics based on* F *and* G *and with parameters* k, m_i, $i = 1\ldots, n-1$ *and* k', m'_i, $i = 1\ldots, n-1$, *respectively. Let us suppose that* $k \geq k'$ *and* $m_i \geq m'_i$, *for all* $i = 1, \ldots, n-1$. *Then we have the following results:*

(i) *If* $X \leq_{\text{st}} Y$ *then* $\boldsymbol{X} \leq_{\text{st}} \boldsymbol{Y}$.

(ii) *If* $X_{(1,n,\tilde{m},k)} \leq_{\text{icx}} Y_{(1,n,\tilde{m},k)}$ *then* $\boldsymbol{X} \leq_{\text{idir-cx}} \boldsymbol{Y}$.

(iii) *If* $X \leq_{\text{hr}} Y$ *then* $\boldsymbol{X} \leq_{\text{hr}} \boldsymbol{Y}$.

(iv) *If*

(a) $m_i \geq 0$ *or* $m'_i \geq 0$ *for all* $i : 1, \ldots, n-1$, *and* $X \leq_{\text{lr}} Y$, *or*

(b) $m_i \geq -1$ *or* $m'_i \geq -1$ *for all* $i : 1, \ldots, n-1$, $X \leq_{\text{hr}} Y$ *and* s/r *is increasing,*

then $\boldsymbol{X} \leq_{\text{lr}} \boldsymbol{Y}$.

Now we consider some results for cases not covered by the GOS's model. First we consider a result for the comparison in the multivariate likelihood ratio order of two random vectors of order statistics in the case of independent and possibly not identically distributed observations. This result has been provided by Belzunce et al. [50].

Theorem 4.4.11. *Given two sets of n independent observations $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$, with absolutely continuous distributions, if $X_i \leq_{\rm lr} Y_j$, for all $i, j = 1, 2, \ldots, n$, then their corresponding random vectors of order statistics $(X_{1:n}, X_{2:n}, \ldots, X_{n:n})$ and $(Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n})$ satisfy*

$$(X_{1:n}, X_{2:n}, \ldots, X_{n:n}) \leq_{\rm lr} (Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}).$$

Belzunce et al. [50] provide also the following result for the comparison of subsets of order statistics $(X_{r_1:n}, X_{r_2:n}, \ldots, X_{r_k:n})$ and $(Y_{r'_1:n'}, Y_{r'_2:n'}, \ldots, Y_{r'_k:n'})$ in the multivariate likelihood ratio order:

Theorem 4.4.12. *Suppose two random vectors of order statistics $(X_{1:n}, X_{2:n}, \ldots, X_{n:n})$ and $(Y_{1:n'}, Y_{2:n'}, \ldots, Y_{n':n'})$ from two sets of independent observations $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_{n'}$, with absolutely continuous distributions, respectively. For $1 \leq r_1 < r_2 < \cdots < r_k \leq n$ and $1 \leq r'_1 < r'_2 < \cdots < r'_k \leq n'$, where $r'_1 - r_1 = r'_2 - r_2 = \cdots = r'_k - r_k \geq \max\{0, n' - n\}$, if $X_i \leq_{\rm lr} Y_j$, for all $i, j = 1, 2, \ldots, n$, then*

$$(X_{r_1:n}, X_{r_2:n}, \ldots, X_{r_k:n}) \leq_{\rm lr} (Y_{r'_1:n'}, Y_{r'_2:n'}, \ldots, Y_{r'_k:n'}).$$

To finish we recall some additional results for the case of dependent observations (see Belzunce et al. [50]).

Theorem 4.4.13. *Given two random vectors $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_n)$, with absolutely continuous distributions, if $(X_{\pi_1}, X_{\pi_2}, \ldots, X_{\pi_n}) \leq_{\rm lr} (Y_{\pi'_1}, Y_{\pi'_2}, \ldots, Y_{\pi'_n})$ for every pair of permutations $\pi, \pi' \in \mathcal{P}$, then the two corresponding order statistics satisfy*

$$(X_{1:n}, X_{2:n}, \ldots, X_{n:n}) \leq_{\rm lr} (Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}).$$

A particular interesting case of the previous results is the case of dependent and exchangeable observations. A random vector $(X_1, X_2, \ldots, X_n)$ is said to be exchangeable if for every permutation $\pi \in \mathcal{P}$, $(X_1, X_2, \ldots, X_n) \stackrel{\rm st}{=} (X_{\pi_1}, X_{\pi_2}, \ldots, X_{\pi_n})$.

Theorem 4.4.14. *Given two random vectors* $(X_1, X_2, \ldots, X_n)$ *and* $(Y_1, Y_2, \ldots, Y_n)$*, with exchangeable and absolutely continuous distributions, respectively, if* $(X_1, X_2, \ldots, X_n) \leq_{\mathrm{lr}} (Y_1, Y_2, \ldots, Y_n)$*, then the two corresponding order statistics satisfy*

$$(X_{1:n}, X_{2:n}, \ldots, X_{n:n}) \leq_{\mathrm{lr}} (Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}).$$

Apart of the application of these results to provide comparisons of the marginals in the univariate likelihood ratio order, these results can be also applied to provide comparisons of conditional ordered data. This topic has received a great attention along the last decade. Some previous results can be found in Langberg et al. [270] and Belzunce et al. [48]. More recently, results for conditional order statistics can be found in the recent works of Li and Zuo [298], Li and Chen [286], Asadi and Bairamov [18], Asadi [17], Li and Zhao [296], Khaledi and Shaked [235], Li and Zhao [297], and Sadegh [406]. In the case of independent and nonidentically distributed random variables, Zhao et al. [491] and Kochar and Xu [260] provide several results in this direction. Some results for the case of record values have been discussed by Khaledi and Shojaei [236] and Khaledi et al. [228]. Again it is possible to obtain additional results from general results in the setting of generalized order statistics; see, for example, Hu et al. [198], Xie and Hu [478], Zhao and Balakrishnan [489] and Zhuang et al. [495]. For a discussion of this topic in the case of the likelihood ratio order see Balakrishnan et al. [27] and Belzunce et al. [50].

To finish with the results for the likelihood ratio order we recall a result for the epoch times of NHPB processes.

Theorem 4.4.15. *Let* $T_{i,n}$ *and* $X_{i,n}$ *be as in Theorem 4.4.3. If* $X_{1,j} \leq_{\mathrm{hr}} X_{2,j}$*, and if* $r_{2,j}/r_{1,j}$ *is increasing, and if for* $j \geq 1$*,*

$$r_{2,j+1}(t) - r_{2,j}(t) \geq r_{1,j+1}(t) - r_{1,j}(t), \qquad t \geq 0,$$

then $(T_{1,1}, T_{1,2}, \ldots, T_{1,n}) \leq_{\mathrm{lr}} (T_{2,1}, T_{2,2}, \ldots, T_{2,n})$ *for all* $n \geq 1$*.*

To finish this section we consider a result for the multivariate dispersive orders. The proof is based on the remark made at the end of Sect. 4.2 and the fact that the dispersive order of the underlying distributions implies the dispersive order of the corresponding GOSs.

Theorem 4.4.16. *Let* X *and* Y *be random variables distribution functions* F *and* G*, respectively, and let* $\boldsymbol{X} = \left(X_{(1,n,\tilde{m},k)}, \ldots, X_{(n,n,\tilde{m},k)}\right)$

and $\boldsymbol{Y} = \left(Y_{(1,n,\tilde{m},k)}, \ldots, Y_{(n,n,\tilde{m},k)}\right)$ *be random vectors of generalized order statistics based on* F *and* G*, respectively. If* $X \leq_{\text{disp}} Y$ *then*

$$\boldsymbol{X} \leq_{\text{disp[var]}} \boldsymbol{Y}.$$

Let us observe that under this result we can provide comparisons in the st:icx order of increasing directionally convex transformations of the random vectors of GOSs as observed in Sect. 4.2. It is worthwhile to mention that Chen and Hu [86] and Xie and Hu [480] provide a whole bunch of results for the comparison in the var order of random vector of GOSs based on the same distribution F.

4.5 Some Additional Comments

This paper is not intended to be an exhaustive paper on multivariate comparisons of ordered data but a review on the main results of this area taking into account the most relevant stochastic orders like stochastic, increasing directionally convex, hazard rate, likelihood ratio, and dispersive orders. Anyway there are some other papers on the topic that we would like to mention. Avérous et al.[23], Khaledi and Kochar [234] and Kochar and Xu [257] provide several results to compare the degree of dependence between pairs of ordered data. This topic is an interesting topic and it is an open area for future research. Some other models of ordered data, not considered in this paper, are sequential order statistics. Zhuang and Hu [493] provide several results for this model. Another topic not covered in this paper is the multivariate comparison of random vectors of spacings. Bartosewicz [42] and Kochar[247] provide results for multivariate comparisons of spacings for order statistics in the multivariate stochastic and likelihood ratio orders. Belzunce et al. [51] give several results for spacings of epoch times of nonhomogeneous Poisson and pure birth processes. Results for the case of spacings of GOSs have been provided by Belzunce et al. [54], Fang et al. [156] and Zhuang and Hu [494].

Acknowledgment

Félix Belzunce is supported by Ministerio de Ciencia e Innovación-FEDER under grant MTM2009-08311.

Chapter 5
Sample Spacings with Applications in Multiple-Outlier Models

Nuria Torrado and Rosa E. Lillo

Abstract: The difference between successive order statistics is denoted as spacing in the literature. They appear in reliability theory, survival analysis, and systems engineering where the times between failures of components of a system correspond with the spacings associated with order statistics. The purpose of this work is to survey the literature about stochastic comparisons of spacings from one and two sequences of order statistics.

N. Torrado (✉)
Department of Statistics and Operations Research, Universidad Pública de Navarra, 31006 Pamplona, Spain
e-mail: nuria.torrado@gmail.com

R.E. Lillo
Department of Statistics, Universidad Carlos III de Madrid, Campus de Getafe, Madrid, Spain
e-mail: lillo@est-econ.uc3m.es

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_5,

5.1 Introduction

Suppose we want to compare two different distributions, the simplest way to do this is by their corresponding means and variances. However, such a comparison is not very informative and the means and the variances sometimes do not exist. Comparisons based on survival functions, hazard rate functions, and other suitable functions of probability distributions establish partial orders among random variables, which are called *stochastic orderings*. Since 1994 the theory of stochastic orderings has grown significantly; see, e.g., Shaked and Shanthikumar [426] for an excellent review of this field.

Models of ordered random variables, such as order statistics, record values, and generalized order statistics, have been studied extensively in the literature (cf. Arnold et al. [15, 16], Kamps [221, 222] and Balakrishnan and Rao [29, 30]). It is well known that these models of ordered random variables appear in the context of reliability theory as the lifetime of k-out-of-n systems and, also, as the epoch times of some stochastic counting processes such as nonhomogeneous Poisson processes and nonhomogeneous pure birth processes (see, e.g., Huang and Su [196] and Belzunce et al. [52]). Spacings, the differences between successive order statistics, have been also investigated in the last decades, see Hu et al. [200] and Wen et al. [474], among others. In the life-testing and reliability models they correspond to times elapsed between successive failures of k-out-of-n systems, and also, they are the inter-epoch times of some stochastic counting processes (cf. Belzunce et al. [52] and Khaledi and Kochar [232]). Here, we review the literature about the area of stochastic comparisons between spacings based on order statistics from one and two samples of random variables.

The organization of this work is as follows: In Sect. 5.2 we recall the definition of order statistics and spacings and introduce the probability density function of normalized and simple spacings. In Sect. 5.3, we investigate stochastic comparisons of spacings from one sample of random variables, and in Sect. 5.4, we do the same for two samples. Section 5.5 is devoted to stochastic comparisons of spacings from multiple-outlier models. Finally, we present some conclusions and open problems in Sect. 5.6.

5.2 Distributional Properties of Spacings

In this section we give the definition of spacings and some distributional properties of them that will be used throughout the work.

If the random variables $X_1, \ldots, X_n$ are arranged in ascending order of magnitude, then the i-th smallest of X_i's is denoted by $X_{i:n}$. The ordered quantities

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}\,, \tag{5.2.1}$$

are called *order statistics* (OS), and $X_{i:n}$ is the i-th order statistic. These random variables are of great interest in many areas of statistics, in particular, in the characterizations of distributions (cf. Deheuvels [115]) and testing problems (see, e.g., Berrendero and Cárcamo [60]), among others. Specifically, there is a very interesting application of OSs in reliability theory. The $(n-k+1)$-th OS in a sample of size n represents the life length of a k-out-of-n system which is an important technical structure. It consists of n components of the same kind with independent and identically distributed life lengths. All n components start working simultaneously, and the system works, if at least k components function; i.e., the system fails if $(n-k+1)$ or more components fail. Special cases of k-out-of-n systems are series and parallel systems.

Another interesting random variables are

$$D_{i:n} = X_{i:n} - X_{i-1:n} \quad \text{and} \quad D^*_{i:n} = (n-i+1)\left(X_{i:n} - X_{i-1:n}\right),$$

for $i = 1, \ldots, n$, with $X_{0:n} \equiv 0$, respectively, called *simple spacings* and *normalized spacings*. In the reliability context they correspond to times elapsed between successive failures.

For heterogeneous but independent exponential random variables, Kochar and Korwar [249] proved that for $i \in \{2, \ldots, n\}$, the distribution of the i-th normalized spacing, D^*_i, is a mixture of independent exponential random variables with p.d.f.

$$f^*_i(t) = \sum_{\mathbf{r}_n} \frac{\prod_{k=1}^{n} \lambda_k}{\prod_{k=1}^{n} \left(\sum_{j=k}^{n} \lambda_{r_j} \right)} \left(\frac{\sum_{j=i}^{n} \lambda_{r_j}}{n-i+1} \right) \exp\left(-t\, \frac{\sum_{j=i}^{n} \lambda_{r_j}}{n-i+1} \right), \tag{5.2.2}$$

where $\mathbf{r}_n = (r_1, \ldots, r_n)$ is a permutation of $(1, \ldots, n)$. Torrado et al. [453] rewrite Eq. (5.2.2) in a more useful way as

$$f_i^*(t) = \sum_{j=1}^{M_i} \Delta(\beta_{m_j}^{(i)}, n) \left(\frac{\beta_{m_j}^{(i)}}{n-i+1} \right) \exp\left(-t \, \frac{\beta_{m_j}^{(i)}}{n-i+1} \right), \quad (5.2.3)$$

where $M_i = \binom{n}{n-i+1}$,

$$\beta_{m_j}^{(i)} = \sum_{\ell=i}^{n} \lambda_{r_\ell}, \quad (5.2.4)$$

with m_j indicating a group of indices of size $n-i+1$, and

$$\Delta(\beta_{m_j}^{(i)}, n) = \sum_{\mathbf{r}_{i-1,m_j}} \left(\prod_{k \in H_{m_j}} \lambda_k \right) \left[\prod_{\ell=1}^{i-1} \left(\sum_{\substack{u=\ell \\ r_u \in H_{m_j}}}^{i-1} \lambda_{r_u} + \beta_{m_j}^{(i)} \right) \right]^{-1}, \quad (5.2.5)$$

where $H_{m_j} = \{1, \ldots, n\} - m_j$ and the outer summation are being taken over all permutations of the elements of H_{m_j}.

The distribution of D_i is also a mixture of independent exponential random variables, with p.d.f.

$$f_i(t) = \sum_{j=1}^{M_i} \Delta(\beta_{m_j}^{(i)}, n) \, \beta_{m_j}^{(i)} \, e^{-t\beta_{m_j}^{(i)}}, \quad (5.2.6)$$

where M_i, $\beta_{m_j}^{(i)}$ and $\Delta(\beta_{m_j}^{(i)}, n)$ are defined as before.

Before proceeding to the main results, let us first recall four lemmas, which will be used in the following sections.

Lemma 5.2.1 (Lemma 3.1. in Kochar and Korwar [249]). *Let $\Delta(\beta_{m_j}^{(i)}, n)$ be as defined in Eq. (5.2.5). Suppose that m_1 and m_2 are two subsets of $\{1, \ldots, n\}$ of size $n-i+1$ $(1 < i \leq n)$ and that they have all but one element in common. Denote the different elements in m_1 by a_1 and those in m_2 by a_2. Then*

$$\lambda_{a_1} \Delta(\beta_{m_1}^{(i)}, n) \geq \lambda_{a_2} \Delta(\beta_{m_2}^{(i)}, n), \qquad \text{if} \quad \lambda_{a_2} \geq \lambda_{a_1}.$$

Lemma 5.2.2 (Chebyshev's sum inequality, Theorem 1 in Mitrinovic [327]). *Let $a_1 \leq a_2 \leq \cdots \leq a_n$ and $b_1 \leq b_2 \leq \cdots \leq b_n$ be two increasing sequences of real numbers. Then*

$$n \sum_{i=1}^{n} a_i b_i \geq \sum_{i=1}^{n} a_i \cdot \sum_{i=1}^{n} b_i.$$

Lemma 5.2.3 (Lemma A.1. in Torrado et al. [453]). *Let* $\Delta(\beta^{(i)}_{m_j}, n)$ *be as in Eq. (5.2.5), and* $m_j = j$ *and* $m_k = (j, \ell)$*. Then*

$$\Delta(\beta^{(3)}_{(3,4)}, 4)\Delta(\beta^{(4)}_{u}, 4) \geq \Delta(\beta^{(3)}_{(u,4)}, 4)\Delta(\beta^{(4)}_{3}, 4) \geq \Delta(\beta^{(3)}_{(u,3)}, 4)\Delta(\beta^{(4)}_{4}, 4),$$

for $u = 1, 2$.

Lemma 5.2.4 (Lemma A.2. in Torrado et al. [453]). *Under the same assumptions as those in Lemma 5.2.3, then*

(a) $\Delta(\beta^{(3)}_{(2,u)}, 4)\Delta(\beta^{(4)}_{1}, 4) \geq \Delta(\beta^{(3)}_{(1,2)}, 4)\Delta(\beta^{(4)}_{u}, 4),$

(b) $\Delta(\beta^{(3)}_{(1,u)}, 4)\Delta(\beta^{(4)}_{2}, 4) \geq \Delta(\beta^{(3)}_{(1,2)}, 4)\Delta(\beta^{(4)}_{u}, 4),$

(c) if $\beta^{(3)}_{(1,u)} - \beta^{(4)}_{2} < 0$*, then*
$\Delta(\beta^{(3)}_{(2,u)}, 4)\Delta(\beta^{(4)}_{1}, 4) \geq \Delta(\beta^{(3)}_{(1,u)}, 4)\Delta(\beta^{(4)}_{2}, 4),$

for $u = 3, 4$.

Another tool used in the literature to obtain the distribution of spacings is the theory of permanent. Suppose $\boldsymbol{A} = (a_{i,j})$ is a square matrix of order n. Then the permanent of the matrix $\boldsymbol{A}$ is defined to be

$$\mathrm{Per}\boldsymbol{A} = \sum_{P} \prod_{j=1}^{n} a_{j,i_j}\,, \tag{5.2.7}$$

where $\sum_P$ denotes the sum over all $n!$ permutations $(i_1, i_2, \ldots, i_n)$ of $(1, 2, \ldots, n)$. The definition of the permanent in Eq. (5.2.7) is thus similar to that of the determinant except that it does not have the alternating sign (c.f. Balakrishnan [25]). If $\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n$ are vectors in $\mathbb{R}^n$, then we will denote by $[\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n]$ the permanent of the $n \times n$ matrix $(\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n)$. The permanent

$$\Bigg[\underbrace{\boldsymbol{d}_1}_{r_1}, \underbrace{\boldsymbol{d}_1}_{r_2}, \ldots\Bigg]$$

is obtained by taking r_1 copies of $\boldsymbol{d}_1$, r_2 copies of $\boldsymbol{d}_2$, and so on.

It is useful to represent the joint density functions of order statistics by using the theory of permanents when the underlying random variables are not identical (see [31, 32]) and also the density function of the spacing $D_{i:n}$ which is given by

$$f_{i:n}(t) = \frac{1}{(i-2)!(n-i)!} \int_0^\infty \left[\underbrace{\boldsymbol{F}(u)}_{i-2}, \boldsymbol{f}(u), \boldsymbol{f}(u+t), \underbrace{\bar{\boldsymbol{F}}(u)}_{n-i} \right] \mathrm{d}u \,,$$

for $t \in \mathbb{R}_+ \equiv [0, \infty)$ and $i = 2, \ldots, n$, where

$$\begin{aligned} \boldsymbol{f}(t) &= (f_1(t), \ldots, f_n(t))', \quad \boldsymbol{F}(t) = (F_1(t), \ldots, F_n(t))', \\ \bar{\boldsymbol{F}}(t) &= (\bar{F}_1(t), \ldots, \bar{F}_n(t))' \end{aligned}$$

with $\bar{F} = 1 - F$. The theory of permanent has been used in [200, 202, 474] in order to study stochastic orderings of spacings.

5.3 Stochastic Orderings of Spacings from One Sample

Here, we give briefly a review of stochastic orders related to the *location*, the *magnitude*, and the *dispersion* of random variables. These notions can be found in Shaked and Shanthikumar [426].

The first results that we discuss concern the usual stochastic ordering. Let X and Y be univariate random variables with cumulative distribution functions (cdf's) F and G, respectively. We say that X is smaller than Y in the *usual stochastic order* if $\bar{F}(t) \leq \bar{G}(t)$, for all t and in this case, we write $X \leq_{\mathrm{st}} Y$ or $F \leq_{\mathrm{st}} G$.

Many authors have studied the stochastic properties of spacings from restricted families of distributions, such as *decreasing hazard rate* and *increasing hazard rate* distributions. A random variable X (or its distribution) is said to be decreasing (increasing) hazard rate or DHR (IHR) if $\bar{F}$ is logconvex (logconcave). If $X_1, X_2, \ldots, X_n$ is a random sample from a DHR (IHR) distribution, then it has been proved by Sukhatme and Proschan [37] that the successive normalized spacings are stochastically increasing (decreasing), that is, for $i = 1, \ldots, n-1$,

$$D^*_{i:n} \leq_{\mathrm{st}} (\geq_{\mathrm{st}}) D^*_{i+1:n} \,. \tag{5.3.1}$$

They also showed that $D^*_{i:n+1} \leq_{\mathrm{st}} (\geq_{\mathrm{st}}) D^*_{i:n}$, when $n \geq i$ for fixed i.

The corresponding problem when the random variables have exponential distributions has also been well studied. Sukhatme [445] showed that the normalized spacings of a random sample from an exponential distribution are i.i.d. random variables having the same exponential distribution. For heterogeneous but independent exponential random variables, Pledger and Proschan [377] proved that if the scale parameters of the exponential distributions are not all equal then the i-th normalized spacing is stochastically smaller than the $(i+1)$-th normalized spacing.

Theorem 5.3.1 (Theorem 3.1 in Pledger and Proschan [377]). *If $X_1, X_2, \ldots, X_n$ are independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$, then*

$$D^*_{i:n} \leq_{\mathrm{st}} D^*_{i+1:n}, \qquad \textit{for } i = 1, \ldots, n-1.$$

Next we consider the hazard rate ordering. Recall that, given two random variables X and Y with hazard rate functions h_X and h_Y, respectively, X is said to be smaller than Y in the *hazard rate order*, denoted by $X \leq_{\mathrm{hr}} Y$ or $F \leq_{\mathrm{hr}} G$, if $h_X(t) \geq h_Y(t)$, for all t.

For one sample, Kochar and Kirmani [248] strengthened Eq. (5.3.1) from stochastic ordering to hazard rate ordering, i.e., for $i=1, \ldots, n-1$

$$D^*_{i:n} \leq_{\mathrm{hr}} D^*_{i+1:n}, \tag{5.3.2}$$

when the random variables are independent and identically distributed from a DHR distribution. Khaledi and Kochar [229] showed that $D^*_{i:n} \leq_{\mathrm{hr}} D^*_{j:m}$ whenever $j \geq i$ and $j - i \geq m - n$, if F is DHR. Hu and Wei [203] proved the following result for generalized spacings, $D_{j,i:n} = X_{j:n} - X_{i:n}$, for $1 \leq i < j \leq n$.

Theorem 5.3.2 (Theorem 4.1 in Hu and Wei [203]). *If $X_1, X_2, \ldots, X_n$ are independent random variables from a* DHR (IHR) *distribution then*

$$D_{j-1,i-1:n-1} \leq_{\mathrm{hr}} (\geq_{\mathrm{hr}}) D_{j,i:n}, \quad \textit{for all } 1 \leq i < j \leq n.$$

When $X_1, X_2, \ldots, X_n$ are independent exponential random variables with possibly unequal scale parameters, observing Eq. (5.2.3), note that $D^*_{i:n} \leq_{\mathrm{hr}} D^*_{i+1:n}$ if and only if

$$\frac{\sum_{j=1}^{M_i} \Delta(\beta^{(i)}_{m_j}, n)\, \frac{\beta^{(i)}_{m_j}}{n-i+1}\, e^{-t\beta^{(i)}_{m_j}/(n-i+1)}}{\sum_{j=1}^{M_i} \Delta(\beta^{(i)}_{m_j}, n)\, e^{-t\beta^{(i)}_{m_j}/(n-i+1)}} \geq \frac{\sum_{j=1}^{M_{i+1}} \Delta(\beta^{(i+1)}_{m_j}, n)\, \frac{\beta^{(i+1)}_{m_j}}{n-i}\, e^{-t\beta^{(i+1)}_{m_j}/(n-i)}}{\sum_{j=1}^{M_{i+1}} \Delta(\beta^{(i+1)}_{m_j}, n)\, e^{-t\beta^{(i+1)}_{m_j}/(n-i)}},$$

which can be rewritten as

$$\sum_{j=1}^{M_{i+1}} \sum_{k=1}^{M_i} \Delta(\beta^{(i)}_{m_k}, n) \Delta(\beta^{(i+1)}_{m_j}, n)\, e^{-t\left(\frac{\beta^{(i)}_{m_k}}{n-i+1} + \frac{\beta^{(i+1)}_{m_j}}{n-i}\right)} \left(\frac{\beta^{(i)}_{m_k}}{n-i+1} - \frac{\beta^{(i+1)}_{m_j}}{n-i}\right) \geq 0. \tag{5.3.3}$$

Throughout this article, we shall suppose, without loss of generality, that the λ_i's are in increasing order.

Next, we give a useful lemma which will be used later.

Lemma 5.3.3 (Lemma A.3. in Torrado et al. [453]). *Let $\beta_{m_k}^{(i)}$ be as defined in Eq.* (5.2.4), *then*

$$\sum_{k=1}^{M_i}\sum_{j=1}^{M_{i+1}}\left(\frac{\beta_{m_k}^{(i)}}{n-i+1}-\frac{\beta_{m_j}^{(i+1)}}{n-i}\right)=0.$$

Proof: Since

$$\binom{n}{n-i}\binom{n-1}{n-i}\frac{1}{n-i+1}=\binom{n}{n-i+1}\binom{n-1}{n-i-1}\frac{1}{n-i},$$

we have

$$\begin{aligned}
&\sum_{j=1}^{M_{i+1}}\sum_{k=1}^{M_i}\left(\frac{\beta_{m_k}^{(i)}}{n-i+1}-\frac{\beta_{m_j}^{(i+1)}}{n-i}\right)\\
&\quad=\sum_{k=1}^{M_i}M_{i+1}\,\frac{\beta_{m_k}^{(i)}}{n-i+1}-\sum_{j=1}^{M_{i+1}}M_i\,\frac{\beta_{m_j}^{(i+1)}}{n-i}\\
&\quad=\sum_{\ell=1}^{n}\binom{n}{n-i}\binom{n-1}{n-i}\frac{\lambda_\ell}{n-i+1}-\sum_{\ell=1}^{n}\binom{n}{n-i+1}\binom{n-1}{n-i-1}\frac{\lambda_\ell}{n-i}\\
&\quad=\left[\binom{n}{n-i}\binom{n-1}{n-i}\frac{1}{n-i+1}-\binom{n}{n-i+1}\binom{n-1}{n-i-1}\frac{1}{n-i}\right]\sum_{\ell=1}^{n}\lambda_\ell\\
&\quad=0.
\end{aligned}$$

■

Kochar and Korwar [249] established the likelihood ratio ordering between the first normalized spacing and the others. The likelihood ratio order is stronger than the hazard rate order, so in particular $D^*_{1:n}\le_{\mathrm{hr}}D^*_{2:n}$. They also conjectured that a result similar to Eq. (5.3.2) holds and proved this conjecture for $n=3$. The general case for any n is still an open problem in the literature. Torrado et al. [453] solved partially this problem. In particular, they proved that, for $n=4$, the successive normalized spacings from heterogeneous exponential random variables are increasing in hazard rate ordering, that is,

$$D^*_{1:4}\le_{\mathrm{hr}}D^*_{2:4}\le_{\mathrm{hr}}D^*_{3:4}\le_{\mathrm{hr}}D^*_{4:4}.$$

Therefore, they showed that, in general, the second normalized spacing is smaller than the third normalized spacing according to the hazard rate ordering.

Theorem 5.3.4 (Theorem 3.1. in Torrado et al. [453])**.** *Let $X_1, \ldots, X_n$ be independent exponential random variables such that X_i has a hazard rate λ_i , for $i = 1, \ldots, n$, then*

$$D^*_{2:n} \leq_{\text{hr}} D^*_{3:n}, \quad \textit{for all } n.$$

We turn to consider the simple spacings of the order statistics. Note that the probability $\Delta(\beta^{(i)}_{m_j}, n)$ in Eq. (5.2.6) is the same than in the p.d.f. of $D^*_{i:n}$. This condition is essential to prove whether $D_{i:n} \leq_{\text{hr}} D_{i+1:n}$ for $i = 1, \ldots, n-1$. We give below some advances on this. Specifically, we prove that the second simple spacing is smaller than the third simple spacing according to the hazard rate ordering.

Theorem 5.3.5 (Theorem 3.2. in Torrado et al. [453])**.** *Under the same assumptions as those in Theorem 5.3.4. Then,*

$$D_{2:n} \leq_{\text{hr}} D_{3:n}, \quad \textit{for any } n.$$

Next, we show that the successive simple spacings are increasing in hazard rate ordering for $n = 4$. Hu et al. [200] proved that $D_{1:n} \leq_{\text{lr}} D_{2:n}$, and by Theorem 5.3.5 we know that $D_{2:n} \leq_{\text{hr}} D_{3:n}$, so we have to show $D_{3:4} \leq_{\text{hr}} D_{4:4}$.

Theorem 5.3.6 (Theorem 4.2. in Torrado et al. [453])**.** *Under the same assumptions as those in Theorem 5.3.4, then*

$$D_{3:4} \leq_{\text{hr}} D_{4:4}.$$

Proof: We have to show

$$\sum_{j=1}^{M_4} \sum_{k=1}^{M_3} \Delta(\beta^{(3)}_{m_k}, 4) \Delta(\beta^{(4)}_{m_j}, 4)\, e^{-t\left(\beta^{(3)}_{m_k} + \beta^{(4)}_{m_j}\right)} \left(\beta^{(3)}_{m_k} - \beta^{(4)}_{m_j}\right) \geq 0.$$

Here, the matrix of $\beta^{(3)}_{m_k} - \beta^{(4)}_{m_j}$ is

$$\begin{pmatrix} \lambda_3+\lambda_4-\lambda_1 & \lambda_3+\lambda_4-\lambda_2 & \lambda_4 & \lambda_3 \\ \lambda_2+\lambda_4-\lambda_1 & \lambda_4 & \lambda_2+\lambda_4-\lambda_3 & \lambda_2 \\ \lambda_2+\lambda_3-\lambda_1 & \lambda_3 & \lambda_2 & \lambda_2+\lambda_3-\lambda_4 \\ \lambda_4 & \lambda_1+\lambda_4-\lambda_2 & \lambda_1+\lambda_4-\lambda_3 & \lambda_1 \\ \lambda_3 & \lambda_1+\lambda_3-\lambda_2 & \lambda_1 & \lambda_1+\lambda_3-\lambda_4 \\ \lambda_2 & \lambda_1 & \lambda_1+\lambda_2-\lambda_3 & \lambda_1+\lambda_2-\lambda_4 \end{pmatrix} \tag{5.3.4}$$

and we can use the same approach as in the proof of Theorem 5.3.5. It is easy to check that there are only four possible negative coefficients $a_{u,2} = \lambda_j + \lambda_k - \lambda_\ell$ for $j < k < \ell$ and $u \notin \{j, k, \ell\}$. Now, we consider the term $a_{u,1} = \lambda_j + \lambda_\ell - \lambda_k \geq 0$ for $u = 1, \ldots, 4$. Notice that $\exp\left\{-t\left(\beta^{(3)}_{(j,\ell)} + \beta^{(4)}_k\right)\right\} = \exp\left\{-t\left(\beta^{(3)}_{(j,k)} + \beta^{(4)}_\ell\right)\right\}$.

Now, if $u = 1$ or 2, from Lemma 5.2.3, we find that

$$\Delta(\beta^{(3)}_{(u,4)}, 4)\Delta(\beta^{(4)}_3, 4) \geq \Delta(\beta^{(3)}_{(u,3)}, 4)\Delta(\beta^{(4)}_4, 4)\,.$$

And, if $u = 3$ or 4, from Lemma 5.2.4, we have that

$$\Delta(\beta^{(3)}_{(1,u)}, 4)\Delta(\beta^{(4)}_2, 4) \geq \Delta(\beta^{(3)}_{(1,2)}, 4)\Delta(\beta^{(4)}_u, 4)\,.$$

From this, we conclude that

$$b_{u,1} = \Delta(\beta^{(3)}_{(j,\ell)}, 4)\Delta(\beta^{(4)}_k, 4) \geq \Delta(\beta^{(3)}_{(j,k)}, 4)\Delta(\beta^{(4)}_\ell, 4) = b_{u,2}\,,$$

for $u = 1, \ldots, 4$. Hence, by Lemma 5.2.2

$$\sum_{h=1}^{2} a_{u,h} b_{u,h} \geq \frac{1}{2}\left(\sum_{h=1}^{2} a_{u,h}\right)\left(\sum_{h=1}^{2} b_{u,h}\right) = \lambda_j \left(\sum_{h=1}^{2} b_{u,h}\right) \geq 0.$$

This proves the required result. ∎

The last two results were given in Torrado et al. [453]. They also conjectured that

(i) $\Delta(\beta^{(n-1)}_{(k,\ell)}, n)\Delta(\beta^{(n)}_j, n) \geq \Delta(\beta^{(n-1)}_{(j,k)}, n)\Delta(\beta^{(n)}_\ell, n)$,

(ii) $\Delta(\beta^{(n-1)}_{(j,\ell)}, n)\Delta(\beta^{(n)}_k, n) \geq \Delta(\beta^{(n-1)}_{(j,k)}, n)\Delta(\beta^{(n)}_\ell, n)$,

(iii) $\Delta(\beta^{(n-1)}_{(k,\ell)}, n)\Delta(\beta^{(n)}_j, n) \geq \Delta(\beta^{(n-1)}_{(j,\ell)}, n)\Delta(\beta^{(n)}_k, n)$, for $j = 1$ and $k = 2$ if $\beta^{(n-1)}_{(j,\ell)} - \beta^{(n)}_k < 0$,

(iv) $\Delta(\beta^{(n-1)}_{(k,\ell)}, n)\Delta(\beta^{(n)}_{j}, n) \geq \Delta(\beta^{(n-1)}_{(j,\ell)}, n)\Delta(\beta^{(n)}_{k}, n)$, for $j \neq 1$ or $k \neq 2$,

where $\beta^{(n-1)}_{m_k} = \frac{\lambda_j+\lambda_\ell}{2}$ and $\beta^{(n)}_{m_j} = \lambda_j$, with $m_k = (j,\ell)$ and $m_j = j$, respectively. Assuming that these conjectures hold, it would be then possible to prove that, for all n and for all $\lambda_i > 0$,

$$D_{n-1:n} \leq_{\mathrm{hr}} D_{n:n} \quad \text{and} \quad D^*_{n-1:n} \leq_{\mathrm{hr}} D^*_{n:n},$$

using the same methodology as those in Theorem 5.3.6 and Theorem 4.1 in Torrado et al. [453], respectively. To the best of our knowledge, this is still an open problem. Another interesting, open problem is to study whether the hazard rate ordering among simple spacings implies the hazard rate ordering between normalized spacings and vice versa.

Likelihood ratio orderings of spacings of heterogeneous exponential random variables have been also studied in the literature. Given two random variables X and Y, with density functions f and g, respectively, X is said to be smaller than Y in the *likelihood ratio order*, denoted by $X \leq_{\mathrm{lr}} Y$, if $g(t)/f(t)$ is increasing in t. Kochar and Korwar [249] obtained the following result on likelihood ratio ordering between $D^*_{1:n}$ and $D^*_{i:n}$, for $1 < i \leq n$.

Theorem 5.3.7 (Theorem 3.5. in Kochar and Korwar [249])**.** *Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i , $i = 1, \ldots, n$. Then*

$$D^*_{1:n} \leq_{\mathrm{lr}} D^*_{i:n}\,, \quad \textit{for } i = 2, \ldots, n.$$

In a recent paper, Hu et al. [200], using the theory of permanents, established a result similar to Theorem 5.3.7 for the first and the second simple spacings of nonidentical independent exponential random variables.

Theorem 5.3.8 (Theorem 3.2. in Hu et al.[200])**.** *Under the same assumptions as those in Theorem 5.3.7, then*

$$D_{1:n} \leq_{\mathrm{lr}} D_{2:n}, \quad \textit{for all } \lambda_i > 0.$$

Furthermore, they proved that if $\lambda_i + \lambda_j \geq \lambda_k$ for all distinct i, j and k, then

$$D_{n-1:n} \leq_{\mathrm{lr}} D_{n:n} \quad \text{and} \quad D_{n:n+1} \leq_{\mathrm{lr}} D_{n:n}\,.$$

Wen et al. [474] conjectured that simple spacings of order statistics from independent exponential random variables with different scale parameters are ordered according to the likelihood ratio order, that is,

$$D_{i:n} \leq_{\mathrm{lr}} D_{i+1:n}, \quad \text{for } i = 1, \ldots, n-1. \tag{5.3.5}$$

And also, that, for $i{=}1,\ldots,n$, $D_{i:n}{\leq_{\mathrm{lr}}}D_{i+1:n+1}$, if $\lambda_{n+1}{\leq}\min\{\lambda_k, k{=}1,\ldots,n\}$ and $D_{i:n+1} \leq_{\mathrm{lr}} D_{i:n}$, if $\lambda_{n+1} \geq \max\{\lambda_k, k = 1,\ldots,n\}$. Hu et al. [200] gave some advances on this. Specifically, they proved that Eq. (5.3.5) holds for $n = 3$, if $\lambda_{n+1} \leq \lambda_k$, for $k = 1,\ldots,n$, then $D_{1:n} \leq_{\mathrm{lr}} D_{2:n+1}$ and $D_{n:n} \leq_{\mathrm{lr}} D_{n+1:n+1}$, and if $\lambda_{n+1} \geq \lambda_k$, for $k = 1,\ldots,n$, then $D_{2:n+1} \leq_{\mathrm{lr}} D_{2:n}$.

Next we proceed to provide results for the dispersive order. Given the two random variables X and Y with cumulative distributions functions F and G, respectively, we say that X is smaller than Y in the *dispersive order*, denoted by $X \leq_{\mathrm{disp}} Y$, if $F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$ whenever $0 \leq \alpha \leq \beta \leq 1$. That is, the dispersive order requires that the difference between any two quantiles of X to be smaller than the difference between corresponding quantiles of Y.

Bagai and Kochar [24] proved that if $X \leq_{\mathrm{hr}} Y$ and either F or G is DHR (decreasing failure rate), then $X \leq_{\mathrm{disp}} Y$. It is known that spacings of independent heterogeneous exponential random variables have DHR distributions (cf. Kochar and Korwar [249]) and that the likelihood ratio order implies the hazard rate order. Combining these observations and some above results, we have the following corollaries for normalized and simple spacings, respectively.

Corollary 5.3.9. *Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i, $i = 1,\ldots,n$. Then*

(i) $D^*_{1:n} \leq_{\mathrm{disp}} D^*_{i:n}$, *for $i = 2,\ldots,n$.*

(ii) $D^*_{1:n} \leq_{\mathrm{disp}} D^*_{2:n} \leq_{\mathrm{disp}} D^*_{3:n}$, *for all n.*

(iii) $D^*_{1:4} \leq_{\mathrm{disp}} D^*_{2:4} \leq_{\mathrm{disp}} D^*_{3:4} \leq_{\mathrm{disp}} D^*_{4:4}$.

Corollary 5.3.10. *Under the same assumptions as those in Corollary 5.3.9, then*

(i) $D_{1:n} \leq_{\mathrm{disp}} D_{2:n} \leq_{\mathrm{disp}} D_{3:n}$, *for all n.*

(ii) $D_{1:4} \leq_{\mathrm{disp}} D_{2:4} \leq_{\mathrm{disp}} D_{3:4} \leq_{\mathrm{disp}} D_{4:4}$.

(iii) $D_{n-1:n} \leq_{\mathrm{disp}} D_{n:n}$, *if $\lambda_i + \lambda_j \geq \lambda_k$ for all distinct i, j, and k.*

5.4 Stochastic Orderings of Spacings from Two Samples

In the literature, there exist different works about the problem of stochastic comparisons of the spacings of two samples. Kochar [247] and Khaledi and Kochar [229] considered two sequences of i.i.d. random variables $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ from two random variables X and Y, with cdf's F and G and density functions f and g, respectively. It was shown that if either F or G is DHR, then, whenever $j \geq i$ and $j - i \geq m - n$,

$$X \leq_{\text{hr}} Y \Rightarrow D^*_{i:n} \leq_{\text{st}} C^*_{i:n}$$

and

$$X \leq_{\text{lr}} Y \Rightarrow D^*_{i:n} \leq_{\text{hr}} C^*_{i:n},$$

where $C^*_{i:n} = (n-i+1)(Y_{i:n} - Y_{i-1:n})$ and $D^*_{i:n} = (n-i+1)(X_{i:n} - X_{i-1:n})$ are the i-th normalized spacing from Y_i's and X_i's, respectively, with $Y_{0:n} \equiv 0$ and $X_{0:n} \equiv 0$. Hu and Wei [203] proved the following result for generalized and simple spacings from two samples of i.i.d. random variables.

Theorem 5.4.1 (Theorems 3.1. and 3.3. in Hu and Wei [203])**.** *If F or G is DFR then*

$$X \leq_{\text{hr}} Y \Rightarrow D_{j,i:n} \leq_{\text{st}} C_{\ell,k:m},$$

whenever $k \geq i$ and $\ell - j \geq k - i \geq m - n$. Moreover, if F and G are IMRL and DMRL, respectively, then

$$X \leq_{\text{mrl}} Y \Rightarrow D_{i:n} \leq_{\text{st}} C_{j:m},$$

for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

Let $m(t) = \mathrm{E}[X - t \mid X > t]$ denote the *mean residual life* (mrl) function of X. Then X or F is said to be IMRL (DMRL) if $m(t)$ is increasing (decreasing) in t. If X and Y have mrl functions $m_X(t)$ and $m_Y(t)$ such that $m_X(t) \leq m_Y(t)$ for all t, then X is said to be smaller than Y in the *mean residual life order* (denoted by $X \leq_{\text{mrl}} Y$).

Many researchers have showed that the i-th normalized spacing of a sample of size n from heterogeneous exponential population is stochastically larger than the i-th normalized spacing of a sample of size n whose distribution is the average of the distributions in the heterogeneous case, according to different stochastic orderings.

Theorem 5.4.2 (Theorem 3.2. in Pledger and Proschan [377]). *Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from an exponential distribution with common hazard rate $\overline{\lambda} = \sum_{i=1}^{n} \lambda_i / n$. Then*

$$C^*_{i:n} \leq_{\text{st}} D^*_{i:n},$$

for $i = 1, \ldots, n$, where $C^*_{i:n} = (n-i+1)(Y_{i:n} - Y_{i-1:n})$ and $D^*_{i:n} = (n-i+1)(X_{i:n} - X_{i-1:n})$ are the i-th normalized spacing from Y_i's and X_i's, respectively, with $Y_{0:n} \equiv 0$ and $X_{0:n} \equiv 0$.

Kochar and Korwar [249] extended this result from stochastic ordering to likelihood ratio ordering.

Theorem 5.4.3 (Theorem 3.5. in Kochar and Korwar [249]). *Under the same assumptions as those in Theorem 5.4.2, then*

$$C^*_{i:n} \leq_{\text{lr}} D^*_{i:n}, \quad \text{for } i = 1, \ldots, n.$$

Kochar and Rojo [255] further strengthened Theorem 5.4.3 to multivariate likelihood ratio order. For details about the definition and properties of multivariate likelihood ratio order, see Shaked and Shanthikumar [426].

Theorem 5.4.4 (Theorem 2.1. in Kochar and Rojo [255]). *Under the same assumptions as those in Theorem 5.4.2, then*

$$\left(C^*_{1:n}, \ldots, C^*_{n:n}\right) \leq_{\text{lr}} \left(D^*_{1:n}, \ldots, D^*_{n:n}\right).$$

Kochar and Xu [261] provided necessary and sufficient conditions for stochastically comparing according to hazard the rate and likelihood ratio orderings when $Y_1, Y_2, \ldots, Y_n$ is a random sample of size n from an exponential distribution with common hazard rate λ which can differ from $\overline{\lambda}$.

Theorem 5.4.5 (Theorems 8 and 9 in Kochar and Xu [261]). *Let $X_1, X_2, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i , $i = 1, \ldots, n$. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample of size n from an exponential distribution with common hazard rate λ. Then, for $i \geq 2$,*

$$C_{i:n} \leq_{\mathrm{lr}} D_{i:n} \Longleftrightarrow (n-i+1)\,\lambda \geq \frac{\sum\limits_{j\in \boldsymbol{r}_n} p_j \left(\sum\limits_{j=i}^{n} \lambda(r_j)\right)^2}{\sum\limits_{j\in \boldsymbol{r}_n} p_j \left(\sum\limits_{j=i}^{n} \lambda(r_j)\right)},$$

and

$$C_{i:n} \leq_{\mathrm{hr}} D_{i:n} \Longleftrightarrow (n-i+1)\,\lambda \geq \sum_{j\in \boldsymbol{r}_n} p_j \left(\sum_{j=i}^{n} \lambda(r_j)\right),$$

for $i = 1, \ldots, n$, *where*

$$p_j = \frac{\prod\limits_{k=1}^{n} \lambda_k}{\prod\limits_{k=1}^{n} \left(\sum\limits_{j=k}^{n} \lambda(r_j)\right)}.$$

Using Lemma 2.1 in Păltănea [365], it is easy to show that the condition on λ is a necessary and sufficient condition for a hazard rate order between $C_{i:n}$ and $D_{i:n}$.

Some researchers have investigated the effect on the survival function, the hazard rate function, and other characteristics of the time to failure of the spacings when we switch the vector $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ to another vector $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$. Torrado and Lillo [451] proved that if the simple spacings are ordered in the likelihood ratio ordering, then the normalized spacings are also ordered, and vice versa. To the best of our knowledge, this has not been proved in the one sample problem.

Theorem 5.4.6 (Theorems 3.1. in Torrado and Lillo [451])**.** *Let* $X_1, \ldots, X_n$ *and* $Y_1, \ldots, Y_n$ *be two sequences of independent but not necessarily identically distributed random variables. Then*

$$C_{i:n} \leq_{\mathrm{lr}} D_{i:n} \Longleftrightarrow C^*_{i:n} \leq_{\mathrm{lr}} D^*_{i:n},$$

for $i = 1, \ldots, n$.

Proof: It is easy to see that $D^*_{i:n} = \varphi_i(D_{i:n})$, where $\varphi_i(x) = (n-i+1)x$ is an increasing function. If $C_{i:n} \leq_{\mathrm{lr}} D_{i:n}$, then from Theorem 1.C.8. in [426], we get that $C^*_{i:n} \leq_{\mathrm{lr}} D^*_{i:n}$, and vice versa, since $\varphi^{-1}(x)$ is also an increasing function. ∎

The theory of majorization has been used extensively in the literature in order to compare stochastically both order statistics and spacings from two sequences of independent but not identically distributed random variables. Let $\{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$ denote the increasing arrangement of the components of the vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$. The vector $\boldsymbol{x}$ is said to be majorized by the vector $\boldsymbol{y}$, denoted by $\boldsymbol{x} \overset{m}{\prec} \boldsymbol{y}$, if

$$\sum_{i=1}^{j} x_{(i)} \geq \sum_{i=1}^{j} y_{(i)}, \quad \text{for } j = 1, \ldots, n-1 \quad \text{and} \quad \sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}.$$

A real valued function φ defined on a set $\mathcal{A} \in \mathbb{R}^n$ is said to be *Schur-convex* (*Schur-concave*) on $\mathcal{A}$ if $\boldsymbol{x} \prec \boldsymbol{y}$ on $\mathcal{A}$ implies $\varphi(\boldsymbol{x}) \leq (\geq)\ \varphi(\boldsymbol{y})$. For more details see Marshall and Olkin [308].

Pledger and Proschan [377] have shown with the help of an example that for $n = 3$, the survival function of the last spacing $D_{3:3}$ is not Schur-convex. However, Kochar and Korwar [249] proved that the survival function of $D^*_{2:n}$ is Schur-convex in $(\lambda_1, \ldots, \lambda_n)$ for any n and, for $n = 2$, the hazard rate of $D^*_{2:2}$ is Schur-concave in (λ_1, λ_2). Kochar and Rojo [255] strengthened this result from hazard rate ordering to likelihood ratio ordering.

Theorem 5.4.7 (Theorem 3.2. in Kochar and Rojo [255])**.** *Let X_1 and X_2 be two independent exponential random variables with hazard rates λ_1 and λ_2, respectively. Let Y_1 and Y_2 be another set of independent exponential random variables with respective hazard rates θ_1 and θ_2. Then, for $(\theta_1, \theta_2) \prec (\lambda_1, \lambda_2)$,*

$$C_{2:2} \leq_{\mathrm{lr}} D_{2:2}\,.$$

Now we can establish the likelihood ratio ordering between simple spacings from two heterogeneous exponential samples. But first, we need a lemma which is a consequence of Lemma 5.2.1.

Lemma 5.4.8 (Lemma 3.2. in Torrado and Lillo [451])**.** *Let $\Delta(\beta^{(i)}_{m_j}, n)$ be as defined in Eq. (5.2.5). Suppose that m_1 and m_2 are two subsets of $\{1, \ldots, n\}$ of size $n - i + 1$ $(1 < i \leq n)$ and having all but one element in common. Denote the different elements in m_1 by a_1 and those in m_2 by a_2. Then*

$$\beta^{(i)}_{m_1}\, \Delta(\beta^{(i)}_{m_1}, n) \geq \beta^{(i)}_{m_2}\, \Delta(\beta^{(i)}_{m_2}, n) \qquad \text{if} \quad \lambda_{a_2} \geq \lambda_{a_1}.$$

Theorem 5.4.9 (Theorem 3.3. in Torrado and Lillo [451]). *Let $X_1,\dots,X_n$ be independent exponential random variables such that X_i has hazard rate λ_i, for $i=1,\dots,n$, and $Y_1,\dots,Y_n$ be independent exponential random variables such that Y_i has hazard rate θ_i, for $i=1,\dots,n$. If*

$$\sum_{j=1}^{n-i+1} \theta_{(j)} \geq (n-i+1)\overline{\lambda},$$

where $\overline{\lambda}=\sum_{i=1}^{n}\lambda_i/n$ and $\left\{\theta_{(1)},\dots,\theta_{(n)}\right\}$ denote the increasing arrangement of θ_i, for $i=1,\dots,n$. Then

$$C_{i:n} \leq_{\mathrm{lr}} D_{i:n},$$

for $i=1,\dots,n$, where $D_{i:n}$ and $C_{i:n}$ are the i-th simple spacing from X_i's and Y_i's, respectively.

Corollary 5.4.10 (Corollary 3.7. in Torrado and Lillo [451]). *Let $X_1,\dots,X_n$ be independent exponential random variables such that X_i has hazard rate λ_i, for $i=1,\dots,n$, and $Y_1,\dots,Y_n$ be a random sample of size n from an exponential distribution with common hazard rate θ. Then,*

(a) $C_{i:n} \leq_{\mathrm{lr}} D_{i:n}$, if $\overline{\lambda} \leq \theta$,

(b) $D_{i:n} \leq_{\mathrm{lr}} C_{i:n}$, if $(n-i+1)\theta \leq \sum_{j=1}^{n-i+1}\lambda_{(j)}$,

for $i=1,\dots,n$.

Note that Theorem 5.4.3 of Kochar and Korwar can be seen as a particular case of Corollary 5.4.10(a), when $\theta=\overline{\lambda}$.

The following result is of interest because it provides upper and lower bounds for the survival and the hazard rate functions, since the likelihood ratio order implies the usual stochastic and the hazard rate orders.

Proposition 5.4.11 (Proposition 3.8. in Torrado and Lillo [451]). *Let $X_1,\dots,X_n$ be independent exponential random variables such that X_i has hazard rate λ_i, for $i=1,\dots,n$, $Y_1,\dots,Y_n$ be a random sample of size n from an exponential distribution with common hazard rate $\lambda_{(n)}=\max\left\{\lambda_1,\dots,\lambda_n\right\}$, and $Z_1,\dots,Z_n$ be a random sample of size n from an exponential distribution with common hazard rate $\lambda_{(1)}=\min\left\{\lambda_1,\dots,\lambda_n\right\}$. Then*

$$C_{i:n} \leq_{\text{lr}} D_{i:n} \leq_{\text{lr}} H_{i:n}, \quad \textit{for } i = 1, \ldots, n,$$

where $C_{i:n}$, $D_{i:n}$, $H_{i:n}$ denote the i-th simple spacings of Y_i's, X_i's, and Z_i's, respectively.

5.5 Stochastic Orderings of Spacings from Multiple-Outlier Models

In this section, we consider the special case when $X_1, \ldots, X_n$ are independent exponential random variables such that X_i has hazard rate λ for $i = 1, \ldots, p$ and X_j has hazard rate λ_* for $j = p+1, \ldots, n$, where two samples are independent. These models are called multiple-outlier exponential models and have applications in robustness theory, though much of the early work in this direction concentrated only on the case when there is one outlier in the sample (single-outlier model); see, e.g., Barnett and Lewis [40] and Balakrishnan [25]. The simple spacings and normalized spacings from a multiple-outlier exponential model are, respectively, defined by

$$D_{i:n}(p, q; \lambda, \lambda_*) = X_{i:n} - X_{i-1:n}$$

and

$$D^*_{i:n}(p, q; \lambda, \lambda_*) = (n - i + 1)\, D_{i:n}(p, q; \lambda, \lambda_*),$$

for $i = 1, \ldots, n$, with $X_{0:n} \equiv 0$, $q = n - p \geq 1$ and $p \geq 1$. When $p = n - 1$, the multiple-outlier exponential model reduces to a single-outlier exponential model. Khaledi and Kochar [231] demonstrated the conjecture of Kochar and Korwar [249] in the case that the random variables $X_1, \ldots, X_n$ follow a single-outlier model with parameters λ and λ_*, that is, when $\lambda_1 = \cdots = \lambda_{n-1} = \lambda$ and $\lambda_n = \lambda_*$.

Theorem 5.5.1 (Khaledi and Kochar [231])**.** *Let $X_1, X_2, \ldots, X_n$ follow a single-outlier exponential model with parameters λ and λ_*. Then*

$$D^*_{i:n}(n-1, 1; \lambda, \lambda_*) \leq_{\text{hr}} D^*_{i+1:n}(n-1, 1; \lambda, \lambda_*),$$

for $i = 1, \ldots, n-1$.

In a multiple-outlier exponential model, by the theory of permanents, Wen et al. [474] established a likelihood ratio ordering between consecutive simple spacings.

Theorem 5.5.2 (Theorems 1.1. and 1.2. in Wen et al. [474]). *Let $X_1, X_2, \ldots, X_n$ follow a multiple-outlier exponential model with parameters λ and λ_*. Then,*

(i) $D_{i:n}(p, q; \lambda, \lambda_*) \leq_{\mathrm{lr}} D_{i+1:n}(p, q; \lambda, \lambda_*)$

(ii) $D_{i:n}(p, q; \lambda, \lambda_*) \leq_{\mathrm{lr}} D_{i+1:n+1}(p+1, q; \lambda, \lambda_*)$*, if* $\lambda \leq \lambda_*$

(iii) $D_{i:n+1}(p, q+1; \lambda, \lambda_*) \leq_{\mathrm{lr}} D_{i:n}(p, q; \lambda, \lambda_*)$*, if* $\lambda \leq \lambda_*$

(iv) $D_{i:n}(p, q; \lambda, \lambda_*) \leq_{\mathrm{lr}} D_{i:n}(p+1, q-1; \lambda, \lambda_*)$*, if* $\lambda \leq \lambda_*$

for $p, q \geq 1$ *and* $i = 1, \ldots, n-1$.

Chen and Hu [87] strengthened Theorem 5.5.2(ii)–(iv) to multivariate likelihood ratio ordering. Recently, Torrado and Lillo [451] proved the analogue of Theorem 5.5.2(iv) as a special case of Theorem 5.4.9 when $\lambda \geq \lambda_*$.

Theorem 5.5.3 (Theorem 4.4 in Torrado and Lillo [451]). *Under the same assumptions as those in Theorem 5.5.2, if* $\lambda \geq \lambda_*$*,* $p \geq 1$ *and* $q \geq 1$*, then*

$$D_{i:n}(p-k_2, q+k_2; \lambda, \lambda_*) \geq_{\mathrm{lr}} D_{i:n}(p, q; \lambda, \lambda_*) \geq_{\mathrm{lr}} D_{i:n}(p+k_1, q-k_1; \lambda, \lambda_*),$$

where $1 \leq k_1 \leq q$*,* $1 \leq k_2 \leq p$ *and* $i = 1, \ldots, n$.

For two single-outlier exponential models Khaledi and Kochar [232] showed the following result.

Theorem 5.5.4 (Theorem 5.2 in Khaledi and Kochar [232]). *Let $X_1, X_2, \ldots, X_n$ follow a single-outlier exponential model with parameters λ and λ_* and let $Y_1, Y_2, \ldots, Y_n$ another single-outlier exponential model with parameters θ and θ_*. If $(\theta_*, \theta, \ldots, \theta) \leq^{\mathrm{m}} (\lambda_*, \lambda, \ldots, \lambda)$, then*

$$C_{i:n}(n-1, 1; \theta, \theta_*) \leq_{\mathrm{hr}} D_{i:n}(n-1, 1; \lambda, \lambda_*), \quad \textit{for } i = 1, \ldots, n\,.$$

Hu et al. [202] investigated stochastic comparisons of simple spacings from two multiple-outlier exponential models. They used the theory of permanents to prove the following result.

Theorem 5.5.5 (Theorem 4.1 in Hu et al. [202]). *Let $X_1, X_2, \ldots, X_n$ follow a multiple-outlier exponential model with parameters λ and λ_**

and let $Y_1, Y_2, \ldots, Y_n$ another multiple-outlier exponential model with parameters θ and λ_* . If $\lambda \le \lambda_* \le \theta$ then

$$\Big(C_{1:n}(p,q;\theta,\lambda_*),\ldots,C_{n:n}(p,q;\theta,\lambda_*)\Big) \\ \le_{\mathrm{lr}} \Big(D_{1:n}(p,q;\lambda,\lambda_*),\ldots,D_{n:n}(p,q;\lambda,\lambda_*)\Big),$$

with $p, q \ge 2$.

Since the multivariate likelihood ratio order is closed under marginalization (see Shaked and Shanthikumar[426]), it holds that, for $\lambda \le \lambda_* \le \theta$,

$$C_{i:n}(p,q;\theta,\lambda_*) \le_{\mathrm{lr}} D_{i:n}(p,q;\lambda,\lambda_*), \qquad \text{for } i = 1,\ldots,n. \tag{5.5.1}$$

It is easy to check that Eq. (5.5.1) is a special case of Theorem 5.4.9 (see Torrado and Lillo [451] for details).

Using again Theorem 5.4.9, we give below a similar result to Eq. (5.5.1) when the number of exponential random variables with hazard rate λ and λ_* can be changed.

Theorem 5.5.6 (Theorem 4.2. in Torrado and Lillo [451])**.** *Under the same assumptions as those in Theorem 5.5.5. If $\lambda \le \lambda_* \le \theta$, then*

(i) $C_{i:n}\left(p,q;\theta,\lambda_*\right) \le_{\mathrm{lr}} D_{i:n}\left(p+k_1, q-k_1;\lambda,\lambda_*\right)$, *with* $0 \le k_1 \le q$

(ii) $C_{i:n}\left(p,q;\theta,\lambda_*\right) \le_{\mathrm{lr}} D_{i:n}\left(p-k_2, q+k_2;\lambda,\lambda_*\right)$, *with* $0 \le k_2 \le p$

where $q = n - p \ge 1$, $p \ge 1$.

It is worthwhile to mention that some researchers have studied stochastic orderings of m-spacings, $D_{i:n}^{(m)} = X_{i+m-1:n} - X_{i-1:n}$ with $i = 1,\ldots,n-m+1$ and $X_{0:n} \equiv 0$. Specifically, Misra and van der Meulen [326] investigated the likelihood ratio ordering of m-spacings of order statistics based on n independent observations of a random variable. Xu et al. [481] studied the likelihood ratio ordering of m-spacings when $X_1,\ldots,X_n$ follow a multiple-outlier exponential model.

5.6 Conclusions

This work is devoted to review stochastic comparisons of spacings from one and two samples of heterogeneous exponential random variables. In the first part of this report, we have shown results about stochastic orderings among both normalized spacings and simple spacings from one sample. In this case, there exit different open problems such as the conjectures of Sect. 5.3 and the conjectures of Torrado et al. [453] and Wen et al. [474]. The second part of this work concerns stochastic comparisons between spacings from two samples of exponential random variables with different scale parameters. It is still an open problem to study stochastic orderings of spacings when the scale parameters $(\lambda_1, \ldots, \lambda_n)$ and $(\theta_1, \ldots, \theta_n)$ are ordered in the majorization ordering. We also have shown stochastic comparisons among spacings of multiple-outlier exponential models. It would be interesting to study stochastic properties of spacings from other models, such us proportional random variables (PRV) and proportional hazard rates (PHR) models.

Chapter 6
Sample Range of Two Heterogeneous Exponential Variables

Peng Zhao and Xiaohu Li

Abstract: In this paper, we discuss ordering properties of sample range from two independent heterogeneous exponential variables in terms of the likelihood ratio order and the hazard rate order (dispersive order). It is shown, among others, that the weakly majorization order between two parameter vectors is equivalent to the likelihood ratio order between sample ranges and that the p-larger order between two parameter vectors implies the hazard rate order (dispersive order) between sample ranges. In the case of exponential sample range, we thus highlight the close connection that exists between some classical

---

P. Zhao (✉)
School of Mathematics and Statistics, Jiangsu Normal University,
Xuzhou 221116, China
e-mail: zhaop07@gmail.com

X. Li
School of Mathematical Sciences, Xiamen University, Xiamen 361005, China
e-mail: mathxhli@hotmail.com

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_6,

stochastic orders and majorization-type orders. Numerical examples are also provided to illustrate the theoretic results established here.

6.1 Introduction

Spacings are of great interest in goodness-of-fit tests, reliability theory, auction theory, actuarial science, life testing, operations research, and many other areas. One may refer to Balakrishnan and Rao [29, 30] for some goodness-of-fit tests based on functions of sample spacings. Let $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$ denote the order statistics of random variables $X_1, X_2, \ldots, X_n$. Then, the kth order statistic $X_{k:n}$ is just the lifetime of a $(n-k+1)$-out-of-n system, which is a very popular structure of redundancy in fault-tolerant systems that have been studied extensively. In particular, $X_{n:n}$ and $X_{1:n}$ correspond to the lifetimes of parallel and series systems, respectively.

Because of the nice mathematical form and the unique memoryless property, the exponential distribution has widely been used in many fields including reliability analysis. One may refer to Barlow and Proschan [38] and Balakrishnan and Basu [26] for an encyclopedic treatment to developments on the exponential distribution. There is a large number of papers in the literature on stochastic comparisons of exponential sample spacings, see Kochar and Xu [261] for a review on this topic. Recently, some researchers investigated stochastic properties of sample ranges from heterogeneous exponential samples.

Let us first recall some notions of stochastic orders and majorization and related orders that are pertinent to the present discussion. Throughout this paper, the term *increasing* is used for *monotone non-decreasing* and *decreasing* is used for *monotone non-increasing*. Assume two random variables X and Y have densities f_X and f_Y, distribution functions F_X and F_Y, and $\overline{F}_X = 1 - F_X$ and $\overline{F}_Y = 1 - F_Y$ as survival functions, respectively. Then, X is said to be smaller than Y in the *likelihood ratio order* (denoted by $X \le_{\mathrm{lr}} Y$) if $f_Y(x)/f_X(x)$ is increasing in x; X is said to be smaller than Y in the *hazard rate order* (denoted by $X \le_{\mathrm{hr}} Y$) if $\overline{F}_Y(x)/\overline{F}_X(x)$ is increasing in x; X is said to be smaller than Y in the *reversed hazard rate order* (denoted by $X \le_{\mathrm{rh}} Y$) if $F_Y(x)/F_X(x)$ is increasing in x; X is said to be smaller than Y in the *usual stochastic order* (denoted by $X \le_{\mathrm{st}} Y$) if $\overline{F}_Y(x) \ge \overline{F}_X(x)$. It is well known that the following chain of implications hold (see Shaked and Shanthikumar [426]):

$$X \le_{\rm lr} Y \Longrightarrow X \le_{\rm hr} [\le_{\rm rh}] Y \Longrightarrow X \le_{\rm st} Y,$$

Sometimes, it is also of interest to compare variability of probability distributions. The following order, called *dispersive order*, is defined as follows.

Definition 6.1.1. X is said to be less dispersed than Y (denoted by $X \le_{\rm disp} Y$) if

$$F^{-1}(v) - F^{-1}(u) \le G^{-1}(v) - G^{-1}(u)$$

for $0 < u \le v < 1$, where F^{-1} and G^{-1} are the right continuous inverses of F and G, respectively.

The notion of *majorization* is quite useful in establishing various inequalities. Let $x_{(1)} \le \cdots \le x_{(n)}$ be the increasing arrangement of the components of the vector $\boldsymbol{x} = (x_1, \ldots, x_n)$. A vector $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is said to majorize another vector $\boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ (written as $\boldsymbol{x} \overset{\rm m}{\succeq} \boldsymbol{y}$) if $\sum_{i=1}^{j} x_{(i)} \le \sum_{i=1}^{j} y_{(i)}$ for $j = 1, \ldots, n-1$ and $\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$; a vector $\boldsymbol{x} \in \mathbb{R}^n$ is said to weakly supmajorize another vector $\boldsymbol{y} \in \mathbb{R}^n$ (written as $\boldsymbol{x} \overset{\rm w}{\succeq} \boldsymbol{y}$) if $\sum_{i=1}^{j} x_{(i)} \le \sum_{i=1}^{j} y_{(i)}$ for $j = 1, \ldots, n$; a vector $\boldsymbol{x} \in \mathbb{R}_+^n$ is said to be p-larger than another vector $\boldsymbol{y} \in \mathbb{R}_+^n$ (written as $\boldsymbol{x} \overset{\rm p}{\succeq} \boldsymbol{y}$) if $\prod_{i=1}^{j} x_{(i)} \le \prod_{i=1}^{j} y_{(i)}$ for $j = 1, \ldots, n$. Apparently, $\boldsymbol{x} \overset{\rm m}{\succeq} \boldsymbol{y}$ implies $\boldsymbol{x} \overset{\rm w}{\succeq} \boldsymbol{y}$, and $\boldsymbol{x} \overset{\rm p}{\succeq} \boldsymbol{y}$ is equivalent to $\log(\boldsymbol{x}) \overset{\rm w}{\succeq} \log(\boldsymbol{y})$, where $\log(\boldsymbol{x})$ is the vector of logarithms of the coordinates of $\boldsymbol{x}$. Also, $\boldsymbol{x} \overset{\rm m}{\succeq} \boldsymbol{y}$ implies $\boldsymbol{x} \overset{\rm p}{\succeq} \boldsymbol{y}$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}_+^n$.

For more details on majorization and p-larger orders and their applications, one may refer to Marshall, Olkin, and Arnold [312] and Bon and Păltănea [75].

Suppose $X_1, \ldots, X_n$ are independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$. Let $Y_1, \ldots, Y_n$ be a random sample of size n from an exponential distribution with common hazard rate λ. Kochar and Rojo [255] showed, for $\lambda \ge \bar{\lambda} = \sum_{i=1}^{n} \lambda_i/\text{n}$, that

$$X_{n:n} - X_{1:n} \ge_{\rm st} Y_{n:n} - Y_{1:n}. \qquad (6.1.1)$$

Zhao and Li [490] built the following characterization result:

$$\lambda \ge \lambda^* \Longleftrightarrow X_{n:n} - X_{1:n} \ge_{\rm st} Y_{n:n} - Y_{1:n}, \qquad (6.1.2)$$

where

$$\lambda^* = \left(\frac{\prod_{i=1}^{n}\lambda_i}{\bar{\lambda}}\right)^{1/(n-1)}.$$

On the other hand, Kochar and Xu [256] improved the result in Eq. (6.1.1) from the usual stochastic order to the reversed hazard rate order as

$$X_{n:n} - X_{1:n} \geq_{\text{rh}} Y_{n:n} - Y_{1:n}.$$

Afterward, Genest et al. [174] further proved, for $\lambda = \bar{\lambda}$, that

$$X_{n:n} - X_{1:n} \geq_{\text{lr}} Y_{n:n} - Y_{1:n}$$

and

$$X_{n:n} - X_{1:n} \geq_{\text{disp}} Y_{n:n} - Y_{1:n}.$$

Recently, Mao and Hu [302] further presented the following equivalent characterizations:

$$\lambda \geq \bar{\lambda} \Longleftrightarrow X_{n:n} - X_{1:n} \geq_{\text{lr}} Y_{n:n} - Y_{1:n} \Longleftrightarrow X_{n:n} - X_{1:n} \geq_{\text{rh}} Y_{n:n} - Y_{1:n}.$$

Here, we will discuss the various ordering properties of sample range from two independent heterogeneous exponential variables. Let X_1, X_2 be independent exponential random variables with X_i having hazard rate λ_i, $i = 1, 2$. Let X_1^*, X_2^* be another set of independent exponential random variables with X_i^* having hazard rate λ_i^*. It is then shown, under the condition $\lambda_1 \leq \lambda_1^* \leq \lambda_2^* \leq \lambda_2$, that

$$(\lambda_1, \lambda_2) \overset{\text{w}}{\succeq} (\lambda_1^*, \lambda_2^*) \Longleftrightarrow X_{2:2} - X_{1:2} \geq_{\text{lr}} X_{2:2}^* - X_{1:2}^*$$

and

$$(\lambda_1, \lambda_2) \overset{\text{p}}{\succeq} (\lambda_1^*, \lambda_2^*) \Longrightarrow X_{2:2} - X_{1:2} \geq_{\text{hr}} [\geq_{\text{disp}}] X_{2:2}^* - X_{1:2}^*.$$

We also establish, under the condition $\lambda_1 \leq \lambda_1^* \leq \lambda_2 \leq \lambda_2^*$, that

$$\lambda_2 - \lambda_1 \geq \lambda_2^* - \lambda_1^* \Longrightarrow X_{2:2} - X_{1:2} \geq_{\text{lr}} X_{2:2}^* - X_{1:2}^*$$

and

$$\frac{\lambda_2}{\lambda_1} \geq \frac{\lambda_2^*}{\lambda_1^*} \Longrightarrow X_{2:2} - X_{1:2} \geq_{\text{hr}} [\geq_{\text{disp}}] X_{2:2}^* - X_{1:2}^*.$$

Some immediate consequences of these results are also pointed out. As a matter of fact, the above results reveal a correspondence between the various stochastic orders of sample range from two heterogeneous exponential variables and majorization-type orders of the vectors of hazard rates.

6.2 Likelihood Ratio Ordering

Let (X_1, X_2) be a vector of independent exponential random variables with respective hazard rates λ_1 and λ_2, and (X_1^*, X_2^*) be another vector of independent exponential random variables with respective hazard rates λ_1^* and λ_2^*. In what follows we shall proceed by distinguishing three cases for parameters:

6.2.1 Case 1: $\max(\lambda_1, \lambda_2) \leq \min(\lambda_1^*, \lambda_2^*)$

Theorem 6.2.1. *Let (X_1, X_2) be a vector of independent exponential random variables with respective hazard rates λ_1 and λ_2, and (X_1^*, X_2^*) be another vector of independent exponential random variables with respective hazard rates λ_1^* and λ_2^*. If $\lambda_1 \leq \lambda_2 \leq \lambda_1^* \leq \lambda_2^*$, then have*

$$R(X_1, X_2) = X_{2:2} - X_{1:2} \geq_{\rm lr} X_{2:2}^* - X_{1:2}^* = R(X_1^*, X_2^*).$$

Proof: The density function of $R(X_1, X_2)$ can be written as

$$f_{R(X_1,X_2)}(t) = \frac{\lambda_1\lambda_2}{\lambda_1 + \lambda_2}\left[e^{-\lambda_1 t} + e^{-\lambda_2 t}\right], \quad t \geq 0.$$

The required result follows immediately by noting that the ratio

$$\frac{e^{-\lambda_1 t} + e^{-\lambda_2 t}}{e^{-\lambda_1^* t} + e^{-\lambda_2^* t}} = \frac{1}{e^{-(\lambda_1^*-\lambda_1)t} + e^{-(\lambda_2^*-\lambda_1)t}} + \frac{1}{e^{-(\lambda_1^*-\lambda_2)t} + e^{-(\lambda_2^*-\lambda_2)t}} \tag{6.2.1}$$

is increasing in $t \in \mathbb{R}_+$. ∎

6.2.2 Case 2: $\min(\lambda_1, \lambda_2) \leq \min(\lambda_1^*, \lambda_2^*) \leq \max(\lambda_1^*, \lambda_2^*) \leq \max(\lambda_1, \lambda_2)$

Theorem 6.2.2. *Suppose $\lambda_2 = \lambda_2^* = \lambda$. If $\lambda_1 \leq \min(\lambda, \lambda_1^*)$, then,*

$$R(X_1, X_2) = X_{2:2} - X_{1:2} \geq_{\rm lr} X_{2:2}^* - X_{1:2}^* = R(X_1^*, X_2^*).$$

Proof: It suffices for us to show the ratio

$$\Delta(t) = \frac{f_{R(X_1,X_2)}(t)}{f_{R(X_1^*,X_2^*)}(t)} \propto \frac{e^{-\lambda_1 t} + e^{-\lambda t}}{e^{-\lambda_1^* t} + e^{-\lambda t}}$$

is increasing in $t \in \mathbb{R}_+$.

In view of the assumption $\lambda_1 \leq \min(\lambda, \lambda_1^*)$, it holds that

$$
\begin{aligned}
& \Delta'(t)\left[e^{-\lambda_1^* t}+e^{-\lambda t}\right]^2 \\
= \quad & \left[-\lambda_1 e^{-\lambda_1 t}-\lambda e^{-\lambda t}\right]\left[e^{-\lambda_1^* t}+e^{-\lambda t}\right] \\
& -\left[-\lambda_1^* e^{-\lambda_1^* t}-\lambda e^{-\lambda t}\right]\left[e^{-\lambda_1 t}+e^{-\lambda t}\right] \\
= \quad & (\lambda_1^*-\lambda_1)e^{-(\lambda_1+\lambda_1^*)t}+(\lambda_1^*-\lambda)e^{-(\lambda_1^*+\lambda)t}+(\lambda-\lambda_1)e^{-(\lambda+\lambda_1)t} \\
\geq \quad & (\lambda_1^*-\lambda_1)\left[e^{-(\lambda_1+\lambda_1^*)t}+e^{-(\lambda+\lambda_1^*)t}\right] \\
\geq \quad & 0,
\end{aligned}
$$

and thus the required result follows. ■

Theorem 6.2.3. *Suppose $\lambda_1 \leq \lambda_1^* \leq \lambda_2^* \leq \lambda_2$. Then, we have*

$$(\lambda_1,\lambda_2) \overset{\mathrm{w}}{\succeq} (\lambda_1^*,\lambda_2^*) \Longleftrightarrow R(X_1,X_2) \geq_{\mathrm{lr}} R(X_1^*,X_2^*).$$

Proof: Sufficiency Suppose $(\lambda_1,\lambda_2) \overset{\mathrm{w}}{\succeq} (\lambda_1^*,\lambda_2^*)$. We then have $\lambda_1 \leq \lambda_1^* \leq \lambda_2^* \leq \lambda_2$ and $\lambda_1+\lambda_2 \leq \lambda_1^*+\lambda_2^*$. The desired result follows for the case when $\lambda_1+\lambda_2 = \lambda_1^*+\lambda_2^*$ by Theorem 3.2 of Kochar and Rojo [255]. Assume $\lambda_1+\lambda_2 < \lambda_1^*+\lambda_2^*$. There then exists some λ such that $\lambda+\lambda_2 = \lambda_1^*+\lambda_2^*$ and $\lambda_1 < \lambda \leq \lambda_1^*$. Let $Y_{2:2}-Y_{1:2}$ be the sample range from two independent exponentials with respective hazard rates λ and λ_2. From Theorem 3.2 of Kochar and Rojo [255] once again, we have $Y_{2:2}-Y_{1:2} \geq_{\mathrm{lr}} R(X_1^*,X_2^*)$. Moreover, we have $R(X_1,X_2) \geq_{\mathrm{lr}} Y_{2:2}-Y_{1:2}$ from Theorem 6.2.2, and so we can conclude that $R(X_1,X_2) \geq_{\mathrm{lr}} R(X_1^*,X_2^*)$.

Necessity Assume $R(X_1,X_2) \geq_{\mathrm{lr}} R(X_1^*,X_2^*)$. Taking derivative of the ratio on the left side of Eq. (6.2.1) and letting $t=0$, we get $\lambda_1+\lambda_2 \leq \lambda_1^*+\lambda_2^*$ and thus $(\lambda_1,\lambda_2) \overset{\mathrm{w}}{\succeq} (\lambda_1^*,\lambda_2^*)$ holds. ■

As a consequence of Theorem 6.2.3, we have the following corollary.

Corollary 6.2.4. *Let (X_1,X_2) be a vector of independent exponential random variables with respective hazard rates λ_1 and λ_2, and (X_1^*,X_2^*) be another vector of independent exponential random variables with common hazard rate λ. Then, we have*

$$\lambda \geq \frac{\lambda_1+\lambda_2}{2} \Longrightarrow R(X_1,X_2) \geq_{\mathrm{lr}} R(X_1^*,X_2^*).$$

Proof: The case when $\lambda \leq \max(\lambda_1,\lambda_2)$ can be directly obtained from Theorem 6.2.3. We only need to discuss the case when $\lambda >$

$\max(\lambda_1, \lambda_2)$. Let $Z_\lambda\,[Z_\mu]$ be the sample range of a random sample of size 2 from an exponential distribution with common hazard rate $\lambda\,[\mu]$. Assume $\lambda < \mu$. It can be readily seen from Theorem 6.2.1 that $Z_\lambda \geq_{\rm lr} Z_\mu$. Based on this fact, we can conclude that the required result is also valid for the case with $\lambda > \max(\lambda_1, \lambda_2)$. ■

6.2.3 Case 3: $\min(\lambda_1, \lambda_2) \leq \min(\lambda_1^*, \lambda_2^*) \leq \max(\lambda_1, \lambda_2) \leq \max(\lambda_1^*, \lambda_2^*)$

Theorem 6.2.5. *Suppose $\lambda_1 \leq \lambda_1^* \leq \lambda_2 \leq \lambda_2^*$. If $\lambda_2 - \lambda_1 = \lambda_2^* - \lambda_1^*$, then*

$$R(X_1, X_2) \geq_{\rm lr} R(X_1^*, X_2^*).$$

Proof: Without loss of generality, let us assume that $\lambda_1 \leq \lambda_2$, $\lambda_1^* \leq \lambda_2^*$ and $\lambda_1 \leq \lambda_1^*$. It suffices to show that the ratio

$$\zeta(t) = \frac{f_{R_2(X)}(t)}{f_{R_2(X^*)}(t)} = \frac{e^{-\lambda_1 t} + e^{-\lambda_2 t}}{e^{-\lambda_1^* t} + e^{-\lambda_2^* t}}$$

is increasing in $t \in \mathbb{R}_+$. Note that

$$\begin{aligned}
& \zeta'(t)\left[e^{-\lambda_1^* t} + e^{-\lambda_2^* t}\right]^2 \\
= \; & \left[-\lambda_1 e^{-\lambda_1 t} - \lambda_2 e^{-\lambda_2 t}\right]\left[e^{-\lambda_1^* t} + e^{-\lambda_2^* t}\right] \\
& - \left[-\lambda_1^* e^{-\lambda_1^* t} - \lambda_2^* e^{-\lambda_2^* t}\right]\left[e^{-\lambda_1 t} + e^{-\lambda_2 t}\right] \\
= \; & (\lambda_1^* - \lambda_1) e^{-(\lambda_1^* + \lambda_1)t} + (\lambda_2^* - \lambda_2) e^{-(\lambda_2^* + \lambda_2)t} + (\lambda_2^* - \lambda_1) e^{-(\lambda_2^* + \lambda_1)t} \\
& + (\lambda_1^* - \lambda_2) e^{-(\lambda_1^* + \lambda_2)t} \\
= \; & (\lambda_1^* - \lambda_1)\left[e^{-(\lambda_1^* + \lambda_1)t} + e^{-(\lambda_1^* + \lambda_2)t}\right] + 2(\lambda_1^* - \lambda_1) e^{-(\lambda_1 + \lambda_2^*)} \\
\geq \; & 0,
\end{aligned}$$

and thus the required result follows. ■

Theorem 6.2.6. *Suppose $\lambda_1 \leq \lambda_1^* \leq \lambda_2 \leq \lambda_2^*$. Then, we have*

$$\lambda_2 - \lambda_1 \geq \lambda_2^* - \lambda_1^* \Longrightarrow R(X_1, X_2) \geq_{\rm lr} R(X_1^*, X_2^*).$$

Proof: For the case with $\lambda_2 - \lambda_1 = \lambda_2^* - \lambda_1^*$, the desired result follows directly from Theorem 6.2.5.

Let us assume $\lambda_2 - \lambda_1 > \lambda_2^* - \lambda_1^*$. There then exists some λ such that $\lambda_2 - \lambda_1 = \lambda_2^* - \lambda$ and $\lambda_1 < \lambda \leq \lambda_1^*$. Let $Z_{2:2} - Z_{1:2}$ be the sample range from two independent exponentials with respective hazard rates λ and λ_2^*. From Theorem 6.2.5 once again, we have $R(X_1, X_2) \geq_{\rm lr} Z_{2:2} - Z_{1:2}$.

On the other hand, by Theorem 6.2.2, we have $R(X_1^*, X_2^*) \leq_{\rm lr} Z_{2:2} - Z_{1:2}$. And thus we can conclude that $R(X_1, X_2) \geq_{\rm lr} R(X_1^*, X_2^*)$. ■

Example 6.2.7.

(a) Set $\lambda_1 = 2$, $\lambda_2 = 2.8$, $\lambda_1^* = 3$ and $\lambda_2^* = 3.5$. It may be easily verified that the assumption in Theorem 6.2.1 is satisfied. Then $R(X_1, X_2) \geq_{\rm lr} R(X_1^*, X_2^*)$. This coincides with what is displayed in Fig. 6.1a.

(b) Set $\lambda_1 = 2$, $\lambda_2 = 3.5$, $\lambda_1^* = 2.8$ and $\lambda_2^* = 3$. The assumption in Theorem 6.2.3 is satisfied. It can be seen from Fig. 6.1b that the likelihood ratio is increasing which coincides with the theoretic result in Theorem 6.2.3.

(c) Set $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_1^* = 2.8$ and $\lambda_2^* = 3.5$. The assumption in Theorem 6.2.6 is satisfied. By observing Fig. 6.1c, we find that the likelihood ratio is increasing which coincides with the theoretic result in Theorem 6.2.6.

(d) Set $\lambda_1 = 2$, $\lambda_2 = 3.7$, $\lambda_1^* = 2.5$ and $\lambda_2^* = 3$. The assumption in Theorem 6.2.3 is violated, and Fig. 6.1d shows that the likelihood ratio function has a locally decreasing trend, which means $R(X_1, X_2) \ngeq_{\rm lr} R(X_1^*, X_2^*)$ and $R(X_1, X_2) \nleq_{\rm lr} R(X_1^*, X_2^*)$.

(e) Set $\lambda_1 = 2$, $\lambda_2 = 4.5$, $\lambda_1^* = 2.05$ and $\lambda_2^* = 6$. The assumption in Theorem 6.2.6 is violated, and Fig. 6.1e shows that the likelihood ratio function is not monotone, which means the likelihood ratio order does not hold in this case.

■

6.3 Hazard Rate and Dispersive Orderings

In this section, we establish some stochastic ordering results similar to those presented in the preceding section for the hazard rate (dispersive) order.

6.3.1 Case 1: $\min(\lambda_1, \lambda_2) \leq \min(\lambda_1^*, \lambda_2^*) \leq \max(\lambda_1^*, \lambda_2^*) \leq \max(\lambda_1, \lambda_2)$

Lemma 6.3.1 (Marshall, Olkin and Arnold, 2011 [312])**.** *Let $I \subset \mathbb{R}$ be an open interval and let $\phi : I^n \to \mathbb{R}$ be continuously differentiable. Then, ϕ is Schur-convex (Schur-concave) on I^n if and only if ϕ is symmetric on I^n and for all $i \neq j$,*

$$(z_i - z_j)\left[\frac{\partial}{\partial z_i}\phi(\boldsymbol{z}) - \frac{\partial}{\partial z_j}\phi(\boldsymbol{z})\right] \geq [\leq] 0 \quad \text{for all } \boldsymbol{z} \in I^n.$$

Theorem 6.3.2. *If $\min(\lambda_1, \lambda_2) \leq \min(\lambda_1^*, \lambda_2^*)$ and $\lambda_1\lambda_2 = \lambda_1^*\lambda_2^*$, then*

$$R(X_1, X_2) \geq_{\text{hr}} R(X_1^*, X_2^*).$$

Proof: Without loss of generality, let us assume that $\lambda_1 \leq \lambda_2$ and $\lambda_1^* \leq \lambda_2^*$ and, hence, $\lambda_1 \leq \lambda_1^* \leq \lambda_2^* \leq \lambda_2$. The hazard rate function of $R(X_1, X_2)$ can be written as

$$r_{R(X_1,X_2)}(t) = \frac{\lambda_1\lambda_2(e^{-\lambda_1 t} + e^{-\lambda_2 t})}{\lambda_2 e^{-\lambda_1 t} + \lambda_1 e^{-\lambda_2 t}}, \quad t \in \mathbb{R}_+,$$

and now we need to show, under the conditions $\lambda_1 \leq \lambda_1^*$ and $\lambda_1\lambda_2 = \lambda_1^*\lambda_2^*$, that

$$\frac{e^{-\lambda_1 t} + e^{-\lambda_2 t}}{\lambda_2 e^{-\lambda_1 t} + \lambda_1 e^{-\lambda_2 t}} \leq \frac{e^{-\lambda_1^* t} + e^{-\lambda_2^* t}}{\lambda_2^* e^{-\lambda_1^* t} + \lambda_1^* e^{-\lambda_2^* t}}$$

for all $t \in \mathbb{R}_+$, which is actually equivalent to

$$\frac{e^{-x_1} + e^{-x_2}}{x_2 e^{-x_1} + x_1 e^{-x_2}} \leq \frac{e^{-x_1^*} + e^{-x_2^*}}{x_2^* e^{-x_1^*} + x_1^* e^{-x_2^*}}$$

under the conditions $x_1 \leq x_1^*$ and $x_1 x_2 = x_1^* x_2^*$. Denote $y_1 = \log x_1, y_2 = \log x_2, y_1^* = \log x_1^*$ and $y_2^* = \log x_2^*$. We have the following relation:

$$(y_1, y_2) \overset{\text{m}}{\succeq} (y_1^*, y_2^*).$$

Then, it suffices to show that the symmetrical differentiable function $H : (\infty, \infty)^2 \to (0, \infty)$ given by

$$H(y_1, y_2) = \frac{e^{-e^{y_1}} + e^{-e^{y_2}}}{e^{y_2} e^{-e^{y_1}} + e^{y_1} e^{-e^{y_2}}}$$

is Schur-concave.

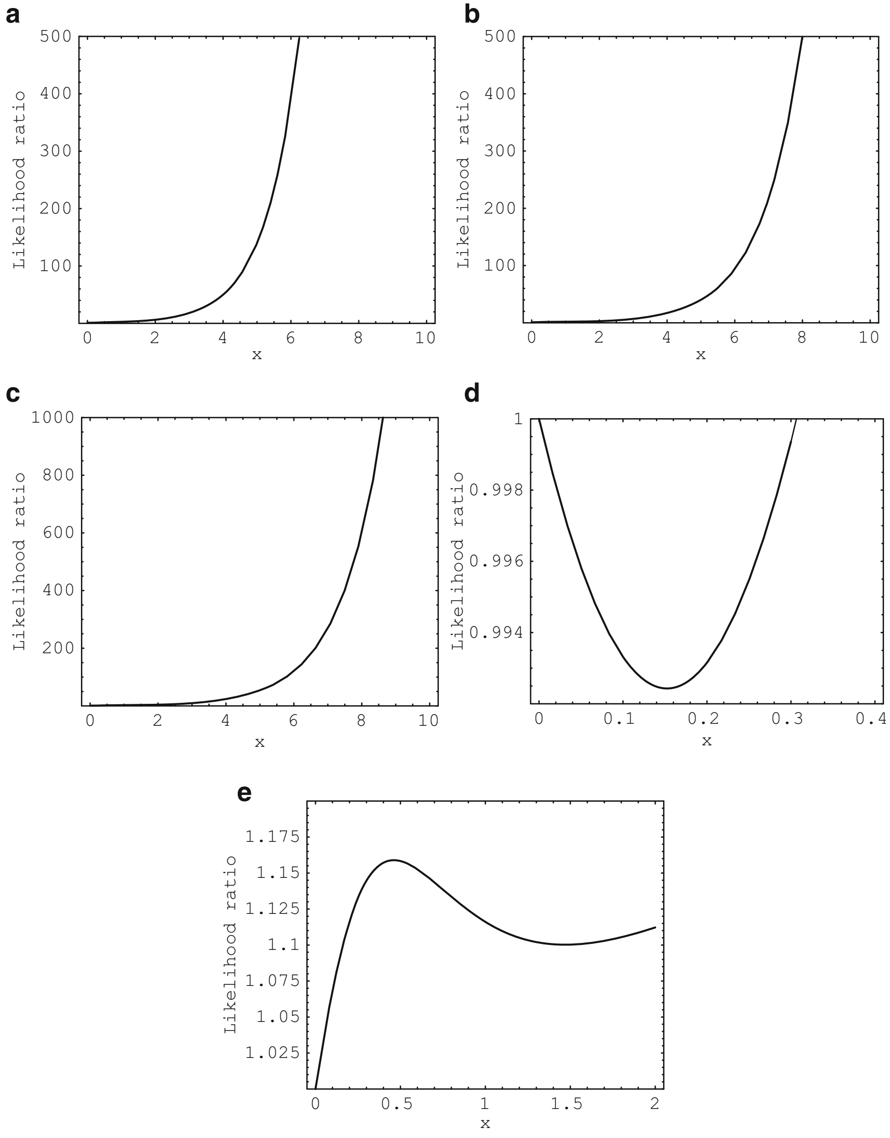


Figure 6.1: Likelihood ratio functions in Example 6.2.7 (**a**) $\lambda_1 = 2$, $\lambda_2 = 2.8$, $\lambda_1^* = 3$ and $\lambda_2^* = 3.5$ (**b**) $\lambda_1 = 2$, $\lambda_2 = 3.5$, $\lambda_1^* = 2.8$ and $\lambda_2^* = 3$ (**c**) $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_1^* = 2.8$ and $\lambda_2^* = 3.5$ (**d**) $\lambda_1 = 2$, $\lambda_2 = 3.7$, $\lambda_1^* = 2.5$ and $\lambda_2^* = 3$ (**e**) $\lambda_1 = 2$, $\lambda_2 = 4.5$, $\lambda_1^* = 2.05$ and $\lambda_2^* = 6$

Note that

$$\begin{aligned}
&\frac{\partial H}{\partial y_1}(y_1, y_2)\left[e^{y_2}e^{-e^{y_1}} + e^{y_1}e^{-e^{y_2}}\right]^2 \\
&\quad = -e^{y_1}e^{-e^{y_1}}\left[e^{y_2}e^{-e^{y_1}} + e^{y_1}e^{-e^{y_2}}\right] \\
&\qquad - \left[-e^{y_1}e^{y_2}e^{-e^{y_1}} + e^{y_1}e^{-e^{y_2}}\right]\left[e^{-e^{y_1}} + e^{-e^{y_2}}\right] \\
&\quad = e^{y_1}e^{y_2}e^{-(e^{y_1}+e^{y_2})} - (e^{2y_1} + e^{y_1})e^{-(e^{y_1}+e^{y_2})} - e^{y_1}e^{-2e^{y_2}}.
\end{aligned}$$

Likewise,

$$\begin{aligned}
&\frac{\partial H}{\partial y_2}(y_1, y_2)\left[e^{-e^{y_1}} + e^{-e^{y_2}} - e^{-(e^{y_1}+e^{y_2})}\right]^2 \\
&\quad = e^{y_1}e^{y_2}e^{-(e^{y_1}+e^{y_2})} - (e^{2y_2} + e^{y_2})e^{-(e^{y_1}+e^{y_2})} - e^{y_2}e^{-2e^{y_1}}.
\end{aligned}$$

We observe that

$$\begin{aligned}
&\frac{\partial H}{\partial y_1}(y_1, y_2) - \frac{\partial H}{\partial y_2}(y_1, y_2) \\
&\overset{\text{sgn}}{=} \left[(e^{2y_2}+e^{y_2})-(e^{2y_1} + e^{y_1})\right]e^{-(e^{y_1}+e^{y_2})} + \left[e^{y_2}e^{-2e^{y_1}} - e^{y_1}e^{-2e^{y_2}}\right].
\end{aligned}$$

Since

$$(e^{2y_2} + e^{y_2}) - (e^{2y_1} + e^{y_1}) \overset{\text{sgn}}{=} y_2 - y_1$$

and

$$e^{y_2}e^{-2e^{y_1}} - e^{y_1}e^{-2e^{y_2}} \overset{\text{sgn}}{=} y_2 - y_1,$$

it holds that

$$(y_1 - y_2)\left[\frac{\partial H}{\partial y_1}(y_1, y_2) - \frac{\partial H}{\partial y_2}(y_1, y_2)\right] \le 0.$$

Now, upon applying Lemma 6.3.1, we can conclude that the function $H(y_1, y_2)$ is Schur-concave and hence the desired result follows immediately. ■

Theorem 6.3.3. *Suppose $\lambda_1 \le \lambda_1^* \le \lambda_2^* \le \lambda_2$. Then, we have:*

$$(\lambda_1, \lambda_2) \overset{\text{p}}{\succeq} (\lambda_1^*, \lambda_2^*) \Longrightarrow R(X_1, X_2) \ge_{\text{hr}} R(X_1^*, X_2^*).$$

Proof: Suppose $(\lambda_1, \lambda_2) \overset{\text{p}}{\succeq} (\lambda_1^*, \lambda_2^*)$. Clearly, we have $\lambda_1 \le \lambda_1^* \le \lambda_2^* \le \lambda_2$ and $\lambda_1\lambda_2 \le \lambda_1^*\lambda_2^*$. When $\lambda_1\lambda_2 = \lambda_1^*\lambda_2^*$, the desired result readily follows from Theorem 6.3.2. Now, we assume that $\lambda_1\lambda_2 < \lambda_1^*\lambda_2^*$. Let $\lambda' = \frac{\lambda_1^*\lambda_2^*}{\lambda_2}$. We then have $\lambda'\lambda_2 = \lambda_1^*\lambda_2^*$ and $\lambda_1 < \lambda' \le \lambda_1^*$. Let $W_{2:2} -$

$W_{1:2}$ be the sample range from two independent exponential random variables with respective hazard rates λ' and λ_2. From Theorem 6.3.2, it follows that $W_{2:2} - W_{1:2} \geq_{\text{hr}} R(X_1^*, X_2^*)$. By Theorem 6.2.2, we also have $R(X_1, X_2) \geq_{\text{hr}} W_{2:2} - W_{1:2}$, and hence we obtain the desired result $R(X_1, X_2) \geq_{\text{hr}} R(X_1^*, X_2^*)$. ∎

According to Theorem 3.B.20(a) of Shaked and Shanthikumar [426], for two nonnegative random variables X and Y such that $X \leq_{\text{hr}} Y$, if either of them is DFR, then $X \leq_{\text{disp}} Y$. On the other hand, the exponential sample range has DFR property from Corollary 2.1 of Kochar and Korwar [249]. Based on these facts, we have the following result.

Theorem 6.3.4. *Let (X_1, X_2) be a vector of independent exponential random variables with respective hazard rates λ_1 and λ_2, and (X_1^*, X_2^*) be another vector of independent exponential random variables with respective hazard rates λ_1^* and λ_2^*. Suppose $\lambda_1 \leq \lambda_1^* \leq \lambda_2^* \leq \lambda_2$. Then, we have:*

$$(\lambda_1, \lambda_2) \overset{\text{p}}{\succeq} (\lambda_1^*, \lambda_2^*) \Longrightarrow R_2(X) \geq_{\text{disp}} R_2(X^*).$$

Upon using Theorems 6.3.3 and 6.3.4 and following a similar argument to the proof of Corollary 6.2.4, we can reach the following result.

Corollary 6.3.5. *Let (X_1, X_2) be a vector of independent exponential random variables with respective hazard rates λ_1 and λ_2, and (X_1^*, X_2^*) be another vector of independent exponential random variables with common hazard rate λ. Then,*

$$\lambda \geq \sqrt{\lambda_1 \lambda_2} \Longrightarrow R(X_1, X_2) \geq_{\text{hr}} [\geq_{\text{disp}}] R(X_1^*, X_2^*).$$

6.3.2 Case 2: $\min(\lambda_1, \lambda_2) \leq \min(\lambda_1^*, \lambda_2^*) \leq \max(\lambda_1, \lambda_2) \leq \max(\lambda_1^*, \lambda_2^*)$

Theorem 6.3.6. *Suppose $\lambda_1 \leq \lambda_1^* \leq \lambda_2 \leq \lambda_2^*$. Then, we have*

$$\frac{\lambda_2}{\lambda_1} = \frac{\lambda_2^*}{\lambda_1^*} \Longrightarrow R(X_1, X_2) \geq_{\text{hr}} R(X_1^*, X_2^*).$$

Proof: Let $\lambda_2/\lambda_1 = \lambda_2^*/\lambda_1^* = c \geq 1$. The hazard rate function of $R(X_1, X_2)$ can be rewritten as

$$r_{R_2(X)}(t) = \frac{c\lambda_1(e^{-\lambda_1 t} + e^{-c\lambda_1 t})}{ce^{-\lambda_1 t} + e^{-c\lambda_1 t}}, \quad t \in \mathbb{R}_+,$$

and now we need to show that

$$\frac{\lambda_1(e^{-\lambda_1 t} + e^{-c\lambda_1 t})}{ce^{-\lambda_1 t} + e^{-c\lambda_1 t}} \le \frac{\lambda_1^*(e^{-\lambda_1^* t} + e^{-c\lambda_1^* t})}{ce^{-\lambda_1^* t} + e^{-c\lambda_1^* t}}$$

for all $t \in \mathbb{R}_+$. It suffices to show the function

$$f(x) = \frac{x(e^{-x} + e^{-cx})}{e^{-cx} + ce^{-x}}$$

is increasing in $x \in \mathbb{R}_+$ for $c \ge 1$. Observe that

$$\begin{aligned}
&f'(x)\left[e^{-cx} + ce^{-x}\right]^2 \\
&\quad = \left[(1-x)e^{-x}+(1-cx)e^{-cx}\right]\left[e^{-cx} + ce^{-x}\right] + cx\left[e^{-cx}+e^{-x}\right]^2 \\
&\quad = [1 + c - (c-1)^2 x]e^{-(1+c)x} + ce^{-2x} + e^{-2cx} \\
&\quad \overset{\text{sgn}}{=} 1 + c - (c-1)^2 x + ce^{(c-1)x} + e^{(1-c)x} \\
&\quad \ge c(c-1)x - (c-1)^2 x \\
&\quad \ge 0,
\end{aligned}$$

and thus the theorem. ■

Theorem 6.3.7. *Suppose $\lambda_1 \le \lambda_1^* \le \lambda_2 \le \lambda_2^*$. Then, we have*

$$\frac{\lambda_2}{\lambda_1} \ge \frac{\lambda_2^*}{\lambda_1^*} \Longrightarrow R(X_1, X_2) \ge_{\text{hr}} R(X_1^*, X_2^*).$$

Proof: The result follows for the case when $\lambda_2/\lambda_1 = \lambda_2^*/\lambda_1^*$ by Theorem 6.3.6. Assume $\lambda_2/\lambda_1 > \lambda_2^*/\lambda_1^*$. There exists some λ such that $\lambda_2/\lambda_1 = \lambda_2^*/\lambda$ and $\lambda_1 < \lambda \le \lambda_1^*$. Let $V_{2:2} - V_{1:2}$ be the sample range from two independent exponentials with respective hazard rates λ and λ_2^*. From Theorem 6.3.6 once again, we have $R(X_1, X_2) \ge_{\text{hr}} V_{2:2} - V_{1:2}$. On the other hand, from Theorem 6.2.2, it follows that $R(X_1^*, X_2^*) \le_{\text{lr}} V_{2:2} - V_{1:2}$, and thus we obtain the desired result. ■

Similar to Theorem 6.3.4, we also have the following result for the dispersive order.

Theorem 6.3.8. *Suppose $\lambda_1 \le \lambda_1^* \le \lambda_2 \le \lambda_2^*$. Then, we have*

$$\frac{\lambda_2}{\lambda_1} \ge \frac{\lambda_2^*}{\lambda_1^*} \Longrightarrow R(X_1, X_2) \ge_{\text{disp}} R(X_1^*, X_2^*).$$

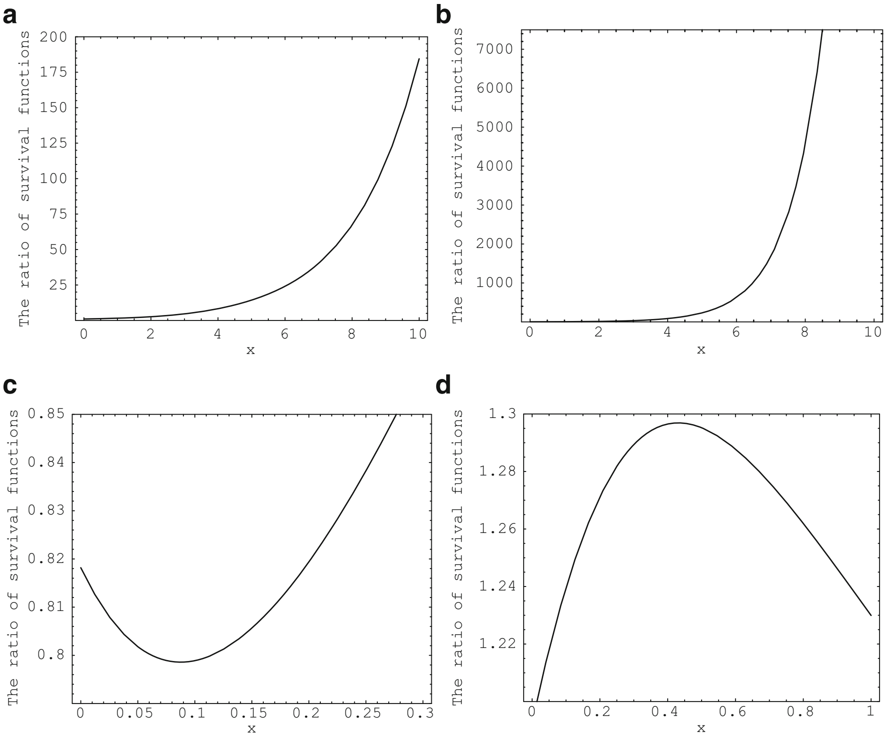


Figure 6.2: The ratio of survival functions in Example 6.3.9 (**a**) $\lambda_1 = 2$, $\lambda_2 = 3.7$, $\lambda_1^* = 2.5$ and $\lambda_2^* = 3$ (**b**) $\lambda_1 = 2$, $\lambda_2 = 2.4$, $\lambda_1^* = 3$ and $\lambda_2^* = 3.5$ (**c**) $\lambda_1 = 2$, $\lambda_2 = 10$, $\lambda_1^* = 3$ and $\lambda_2^* = 3.5$ (**d**) $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_1^* = 2.005$ and $\lambda_2^* = 5$

Example 6.3.9.

(a) Set $\lambda_1 = 2$, $\lambda_2 = 3.7$, $\lambda_1^* = 2.5$ and $\lambda_2^* = 3$. In this case, as is displayed in Fig. 6.1d, the likelihood ratio order does not hold. The assumption in Theorem 6.3.3, however, is satisfied, and the hazard rate order holds as displayed in Fig. 6.2a.

(b) Set $\lambda_1 = 2$, $\lambda_2 = 2.4$, $\lambda_1^* = 3$ and $\lambda_2^* = 3.5$. The assumption in Theorem 6.3.7 is satisfied, and Fig. 6.2b shows that the ratio of survival functions is increasing, which is consistent with the result of Theorem 6.3.7.

(c) Set $\lambda_1 = 2$, $\lambda_2 = 10$, $\lambda_1^* = 3$ and $\lambda_2^* = 3.5$. The assumption in Theorem 6.3.3 is violated, and Fig. 6.1c shows that the ratio of survival functions is not monotone, which means the hazard rate order does not hold in this case.

(d) Set $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_1^* = 2.005$ and $\lambda_2^* = 5$. The assumption in Theorem 6.3.7 is violated, and Fig. 6.1d shows that the hazard rate order does not hold in this case.

■

Acknowledgments

Peng Zhao and Xiaohu Li are supported by National Natural Science Foundation of China (11001112, 11171278) and the Priority Academic Program Development of Jiangsu Higher Education Institutions.

Part III

Stochastic Orders in Reliability

Chapter 7

On Bivariate Signatures for Systems with Independent Modules

Gaofeng Da and Taizhong Hu

Abstract: Gertsbakh et al. (Journal of Applied Probability, 49, 416–429, 2012) proposed the concept of bivariate signature for a 3-state system. In this paper, we first give an equivalent definition for the bivariate signature in the sense of order statistics of component lifetimes and establish the formula for computing the bivariate signature of the dual of a 3-state system. A useful method for computing the bivariate signature based on the combinatorial meaning of the bivariate tail signature is presented. By this method, we derive formulas for computing the bivariate signatures of some systems consisting of independent modules. Some examples are also presented to illustrate our main results.

G. Da (✉) • T. Hu
Department of Statistics and Finance, School of Management,
University of Science and Technology of China, Hefei, Anhui 230026, China
e-mail: dagf@ustc.edu.cn, thu@ustc.edu.cn

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_7,

7.1 Introduction

The concept of signature, introduced by Samaniego [407] for coherent systems with independent and identically distributed (i.i.d.) components, has been a useful tool in studying theoretical behaviors of systems. The *signature* of a coherent system containing n components with i.i.d. lifetimes was defined as the n-dimensional probability vector with the ith element $p_i = \mathrm{P}\{T = X_{i:n}\}$, where T denotes the lifetime of the system, and $X_{i:n}$ is the ith order statistic of the component lifetimes $X_1, \ldots, X_n$. Equivalently, under the i.i.d. assumption, p_i can be expressed as the ratio $N_i/n!$, where N_i is the number of orderings of failure times of all components for which the ith failure causes the system failure. A fundamental property on the signature is that one can express the system reliability at time t in terms of the signature and of the survival functions of lifetimes' order statistics (e.g., see Samaniego [408]). Based on this property, signature has been widely used to evaluate the system reliability and to compare performance of different system structures and so on; see Kochar et al. [253], Navarro et al. [343, 344, 348], Navarro and Eryilmaz [341], Navarro and Rychlik [345, 346], and related references therein.

Often, it is rather difficult to compute the signature of a system by making use of its definition, especially when the structure of the system is complex or the number of the components is large. A method proposed by Boland [65] sometimes makes the computation of signatures relatively easier. The method is mainly based on determining the number of path sets of systems. Denote by r_i the number of path sets of size i of a system with n components. Then the system signatures can be computed by

$$p_i = \frac{r_{n-i+1}}{\binom{n}{i-1}} - \frac{r_{n-i}}{\binom{n}{i}}, \qquad i = 1, \ldots, n.$$

Also,

$$r_i = \binom{n}{i} \bar{P}_{n-i}, \tag{7.1.1}$$

where $\bar{P}_j = \sum_{k=j+1}^{n} p_k$, $j = 0, \ldots, n-1$, are called as the tail signatures of the system. Note that the method just associates with the combination rather than the permutation of the system components. One may refer to Boland [65], Eryilmaz and Zuo [151], Da et al. [102, 103] for computing signatures of some systems by employing this method.

In the last few years, there are some extensions on the notion of signature from various perspectives in the literature. For example, Navarro et al. [350] and Navarro et al. [349] extended the signature to the cases of systems with exchangeable components and of systems with independent heterogeneous components, respectively; Marichal and Mathonet [303] and Marichal et al. [304] studied the signature for the systems with arbitrary dependent components.

Recently, Levitin et al. [279] and Gertsbakh et al. [182] extended the notion of signature to multivariate case in the frame of multi-state systems with binary components. Typically, they mainly focused on the *bivariate signatures* for 3-state coherent systems. Consider a system with three states, e.g., 0, 1, 2, which contains n binary components. Gertsbakh et al. [182] (see also Levitin et al. [279]; Gertsbakh and Shpungin, [179, 180]) defined the bivariate signature for such systems under some regularity assumptions (see Sect. 7.2) as an $n \times n$ matrix, whose (i, j)-element is the ratio $s_{i,j} = N_{i,j}/n!$, where $N_{i,j}$ is the number of orderings of failure times of all components such that ith failure leads to the change of the system state from 2 to 1, and jth failure leads to the change of the system state from 1 to 0. The bivariate signature has important applications especially in the field of network reliability. As described in Levitin et al. [279], the model corresponds to the disintegration of an initially connected network with n edges into isolated parts due to edge failures (nodes are supposed to be functioning forever): the state 2 may correspond to the situation of complete connection, state 1 corresponds to the network with several isolated parts, and state 0 to the presence of the more isolated parts or complete disconnection. Similar to that of the univariate signature, a fundamental property for the bivariate signature (Theorem 7.2.3) was established by Gertsbakh et al. [182].

In the context of 2-state coherent systems, by using different methods and in terms of the signatures of its modules, Da et al. [103] and Gertsbakh et al. [181] derived formulas for computing the signature of a system consisting of independent modules in the structure of series and/or parallel. Likewise, Gertsbakh et al. [182] obtained formulas for computing the bivariate signature of a 3-state system consisting of independent modules in the structures of generalized series and/or parallel by using the fundamental property mentioned above.

In the present paper, we further study the bivariate signature and its computation for systems with independent modules. The paper is organized as follows. In Sect. 7.2, we give an equivalent definition of the bivariate signature in the sense of order statistics of component lifetimes and reprove the fundamental property established by Gertsbakh et al. [182]. Also, we derive a formula for computing the bivariate signature of the dual of a system, which extends a similar result on univariate signature established by Kochar et al. [253]. In Sect. 7.3, a useful method for computing the bivariate signature based on combinatorial meaning of the *bivariate tail signature* is presented. By using this method, in Sect. 7.4, we derive formulas for computing the bivariate signatures of some systems consisting of independent modules such as generalized series and parallel systems, systems with componentwise redundancy, and an important class of 3-state systems. In particular, the important class of 3-state systems provides us with a way to get two different 3-state systems but which have the same signature. Finally, in Sect. 7.5, some numerical examples are provided to illustrate our main results.

7.2 Bivariate Signatures

Consider a 3-state monotone system containing n binary components with its structure function

$$\phi: \ \{0,1\}^n \to \{0,1,2\},$$

where state "2" is called the *perfection state* of the system, "0" is called the *complete failure state*, and "1" is the *partial failure state*. Further, the system is supposed to satisfy the following regularity assumptions:

1. The system is *coherent*, that is, ϕ is increasing, and each component is *relevant*.

2. The failure of one component at most changes the state by one unit.

For the details of such systems and their applications in the network reliability one can refer to Levitin et al. [279] and Gertsbakh et al. [182]. For such systems, Gertsbakh et al. [182] gave a formal definition of the bivariate signature in the following purely combinatorial way.

Denote by Π_n the set of permutations of $\{1,2,\ldots,n\}$. According to the regularity assumptions, for each $\boldsymbol{\pi} = (\pi_1,\ldots,\pi_n) \in \Pi_n$, there are two integers $\ell_1 = \ell_1(\boldsymbol{\pi})$ and $\ell_2 = \ell_2(\boldsymbol{\pi})$ such that

$$\phi(0_{\pi_1}, \ldots, 0_{\pi_{\ell_k - 1}}, 1_{\pi_{\ell_k}}, 1_{\pi_{\ell_k+1}}, \ldots, 1_{\pi_n}) = 3 - k, \tag{7.2.1}$$

$$\phi(0_{\pi_1}, \ldots, 0_{\pi_{\ell_k - 1}}, 0_{\pi_{\ell_k}}, 1_{\pi_{\ell_k+1}}, \ldots, 1_{\pi_n}) = 2 - k \tag{7.2.2}$$

for $k = 1, 2$.

For $1 \le i < j \le n$, denote

$$A_{i,j} := \{\pi \in \Pi_n : \ \ell_1(\boldsymbol{\pi}) = i, \ell_2(\boldsymbol{\pi}) = j\}. \tag{7.2.3}$$

Definition 7.2.1. The signature of system ϕ is defined as an $n \times n$ probability matrix with (i, j)-element given by

$$s_{i,j} = \begin{cases} \frac{|A_{i,j}|}{n!} & \text{for } \ 1 \le i < j \le n, \\ 0 & \text{otherwise,} \end{cases}$$

where $|A_{i,j}|$ denotes the cardinality of $A_{i,j}$.

Note that the definition above completely avoids any reference to lifetimes of the components. Now, we give an equivalent definition in view of order statistics of component lifetimes of the system, which can be seen as a direct generalization of the definition of the univariate signature of coherent systems. This new definition may be useful for further research on the bivariate signature, for example, one can consider the case of systems with non-i.i.d. components.

Definition 7.2.2. Assume that $X_1, \ldots, X_n$, the lifetimes of the n components of the system ϕ, are i.i.d. absolutely continuous random variables. Let T_1 and T_2 denote the time of the system from the state 2 to the state 1 and from the state 2 to the state 0, respectively. Then the signature of the system is defined as an $n \times n$ probability matrix with (i, j)-element given by

$$s_{i,j} = \begin{cases} \mathrm{P}\{T_1 = X_{i:n}, T_2 = X_{j:n}\} & \text{for } \ 1 \le i < j \le n, \\ \\ 0 & \text{otherwise,} \end{cases} \tag{7.2.4}$$

where $X_{1:n}, \ldots, X_{n:n}$ are the order statistics of $X_1, \ldots, X_n$.

Gertsbakh et al. [182] also presented a very important notion, bivariate tail signature, denoted by $\bar{S}_{i,j}$. In the context of Definition 7.2.2, it can be written as

$$\bar{S}_{i,j} = \mathrm{P}\{T_1 > X_{i:n}, T_2 > X_{j:n}\} = \sum_{l=i+1}^{n-1} \sum_{m=(l+1)\vee(j+1)}^{n} s_{l,m}. \tag{7.2.5}$$

Clearly, $\bar{S}_{i,j} = \bar{S}_{i,i}$ for $j < i$. Also,

$$\bar{S}_i^{(1)} := \bar{S}_{i,1} = \sum_{l=i+1}^{n-1} \sum_{m=l+1}^{n} s_{l,m}, \quad i = 1, \ldots, n,$$

and

$$\bar{S}_j^{(2)} := \bar{S}_{0,j} = \sum_{l=1}^{n-1} \sum_{m=(l+1)\vee(j+1)}^{n} s_{l,m}, \quad j = 1, \ldots, n,$$

are so-called marginal tail signatures of a 3-state system. For the sake of discrimination, $\bar{\boldsymbol{S}}^{(1)} = (\bar{S}_1^{(1)}, \ldots, \bar{S}_n^{(1)})$ and $\bar{\boldsymbol{S}}^{(2)} = (\bar{S}_1^{(2)}, \ldots, \bar{S}_n^{(2)})$ are termed as the type I and type II marginal tail signatures, respectively. For the meaning of the marginal tail signature and further discussion on this topic, one refers to Remark 7 in Gertsbakh et al. [182].

Moreover, the signature can be obtained from the tail signature as follows

$$s_{i,j} = \bar{S}_{i-1,j-1} - \bar{S}_{i,j-1} - \bar{S}_{i-1,j} + \bar{S}_{i,j}, \tag{7.2.6}$$

for $1 \le i < j \le n$ (see Gertsbakh et al. [182]). It should be remarked that it is more convenient to discuss the tail signature than the signature itself in many situations.

The following result states that the joint survival function of the degradation times of a 3-state system can be represented in terms of joint survival functions of order statistics of component lifetimes and of the bivariate signature of the system, which can be seen as a direct generalization of the representation in terms of the univariate signature. The result was proved by Gertsbakh et al. [182]. Here we reprove it briefly.

Theorem 7.2.3. *Let $X_1, \ldots, X_n$ be i.i.d. component lifetimes of a 3-state coherent system of order n, and let T_1 and T_2 be the time of the system from the state 2 to the state 1 and from the state 2 to the state 0, respectively. Then,*

$$\begin{aligned} \mathrm{P}\{T_1 > t_1, T_2 > t_2\} &= \sum_{i=0}^{n-1} \sum_{j=i}^{n} \bar{S}_{i,j} \cdot \frac{n!}{i!(j-i)!(n-j)!} F^i(t_1) \\ &\quad [F(t_2) - F(t_1)]^{j-i} \overline{F}^{n-j}(t_2), \end{aligned} \tag{7.2.7}$$

for $t_1 \le t_2$, where F is the common distribution function of all X_i.

Proof: According to the regularity assumptions, T_1 and T_2 will necessarily take on different values of $X_{1:n}, \ldots, X_{n:n}$. By using the i.i.d. assumption on the component lifetimes, we have

$$\begin{aligned}
& \mathrm{P}\{T_1 > t_1, T_2 > t_2\} \\
= & \sum_{i<j} \mathrm{P}\{T_1{>}t_1, T_2{>}t_2 |\ T_1{=}X_{i:n}, T_2{=}X_{j:n}\}\ \mathrm{P}\{T_1{=}X_{i:n}, T_2{=}X_{j:n}\} \\
= & \sum_{i<j} s_{i,j} \cdot \mathrm{P}\{X_{i:n} > t_1, X_{j:n} > t_2\}.
\end{aligned}$$

The joint survival function of ith and jth order statistics from i.i.d. random variables is

$$\begin{aligned}
\mathrm{P}\{X_{i:n} > t_1, X_{j:n} > t_2\} \quad = & \sum_{k=0}^{i-1} \sum_{m=k}^{j-1} \frac{n!}{k!(m-k)!(n-m)!} F^k(t_1) \\
& [F(t_2) - F(t_1)]^{m-k} \overline{F}^{n-m}(t_2)
\end{aligned}$$

for $0 < t_1 < t_2$. Thus, interchanging the order of summations yields

$$\begin{aligned}
& \mathrm{P}\{T_1 > t_1, T_2 > t_2\} \\
= & \sum_{k=0}^{n} \sum_{m=k}^{n} \sum_{i=k+1}^{n-1} \sum_{j=(i+1)\vee(m+1)}^{n} s_{i,j} \cdot \frac{n!}{k!(m-k)!(n-m)!} \\
& \times F^k(t_1)[F(t_2) - F(t_1)]^{m-k} \overline{F}^{n-m}(t_2) \\
= & \sum_{k=0}^{n} \sum_{m=k}^{n} \bar{S}_{k,m} \cdot \frac{n!}{k!(m-k)!(n-m)!} F^k(t_1) \\
& [F(t_2) - F(t_1)]^{m-k} \overline{F}^{n-m}(t_2).
\end{aligned}$$

This completes the proof. ■

Similar to traditional coherent systems, as is clear from Theorem 7.2.3, the degradation times of a 3-state system with i.i.d. components depend on the structure of the system only through the bivariate signatures $s_{i,j}$, $1 \leq i < j \leq n$. In other words, if two systems have the same signature, then the stochastic behaviors of the degradation times of them are identical. It is natural to ask whether two different systems can have the same signature. The answer is yes; see the discussion in Sect. 7.4.3

Undoubtedly, the number of 3-state systems of order n can be very large and it is very hard to get all signatures of them. So it is necessary to find some ways to reduce the computational burden. Obviously, to establish the relationship between the signatures of a system and of its dual system is an efficient way for this since it can cut this burden in half.

The dual of a binary system ϕ is defined as

$$\phi^D(\boldsymbol{x}) = 1 - \phi(\mathbf{1} - \boldsymbol{x}) \text{ for all } \boldsymbol{x} \in \{0,1\}^n,$$

where $\mathbf{1} = (1, \ldots, 1)$. According to Kochar et al. [253], the signature $(p_1, \ldots, p_n)$ of system ϕ and the signature $(p_1^D, \ldots, p_n^D)$ of the dual system ϕ^D have the following relationship:

$$p_i^D = p_{n-i+1}, \quad i = 1, \ldots, n.$$

So, what can we say about the bivariate signature? By the definition of duality for multi-state systems, proposed by EI-Neweihi et al. [145], the dual system of a 3-state system with ϕ has structure function

$$\phi^D(\boldsymbol{x}) = 2 - \phi(\mathbf{1} - \boldsymbol{x}) \text{ for all } \boldsymbol{x} \in \{0,1\}^n.$$

The dual system of a 3-state system has also three states.

Theorem 7.2.4. *Let $s_{i,j}$, $1 \le i < j \le n$, be the signatures of a 3-state system consisting of n i.i.d. components. Then the signatures of its dual system are given by*

$$s_{i,j}^D = s_{n-j+1,n-i+1} \quad \textit{for all } 1 \le i < j \le n.$$

Proof: Similar to Eq. (7.2.3), we define

$$A_{r_1,r_2}^D := \{\pi \in \mathcal{P}_n | \ \ell_1'(\pi) = r_1, \ell_2'(\pi) = r_2\}.$$

for the dual system ϕ^D. It suffices to prove that $|A_{r_1,r_2}| = |A_{n-r_2+1,n-r_1+1}^D|$.

For any $\boldsymbol{\pi} \in A_{r_1,r_2}$, from Eqs. (7.2.1) and (7.2.2), it follows that

$$\phi(0_{\pi_1}, \ldots, 0_{\pi_{r_i-1}}, 1_{\pi_{r_i}}, 1_{\pi_{r_i+1}}, \ldots, 1_{\pi_n}) = 3 - i$$

and

$$\phi(0_{\pi_1}, \ldots, 0_{\pi_{r_i-1}}, 0_{\pi_{r_i}}, 1_{\pi_{r_i+1}}, \ldots, 1_{\pi_n}) = 2 - i$$

for $i = 1, 2$. By the definition of the duality, we have

$$\phi^D(0_{\pi_n}, \ldots, 0_{\pi_{r_i+1}}, 0_{\pi_{r_i}}, 1_{\pi_{r_i-1}}, \ldots, 1_{\pi_1}) = i - 1$$

and

$$\phi^D(0_{\pi_n}, \ldots, 0_{\pi_{r_i+1}}, 1_{\pi_{r_i}}, 1_{\pi_{r_i-1}}, \ldots, 1_{\pi_1}) = i$$

for $i = 1, 2$. Now, we replace i with $3-j$ in the last two equations and then get

$$\phi^D(0_{\pi_n}, \ldots, 0_{\pi_{r_{3-j}+1}}, 0_{\pi_{r_{3-j}}}, 1_{\pi_{r_{3-j}-1}}, \ldots, 1_{\pi_1}) = 2 - j$$

and

$$\phi^D(0_{\pi_n}, \ldots, 0_{\pi_{r_{3-j}+1}}, 1_{\pi_{r_{3-j}}}, 1_{\pi_{r_{3-j}-1}}, \ldots, 1_{\pi_1}) = 3 - j$$

for $j = 1, 2$, which implies that, for any $\boldsymbol{\pi} \in A_{r_1,r_2}$,

$$\boldsymbol{\pi}' = (\pi_n, \ldots, \pi_1) \in A^D_{n-r_2+1,n-r_1+1}.$$

Similarly, the reverse one of each permutation in $A^D_{n-r_2+1,n-r_1+1}$ also belongs to A_{r_1,r_2}. This means there is a one-to-one relationship between A_{r_1,r_2} and $A^D_{n-r_2+1,n-r_1+1}$, and hence $|A_{r_1,r_2}| = |A^D_{n-r_2+1,n-r_1+1}|$. We complete the proof. ■

7.3 A Useful Method to Compute Bivariate Signatures

Clearly, if we compute the bivariate signature of a given 3-state system by using its definition, we have to check every permutation of the ordering of component lifetimes of the system. This is really a hard work. In this section, we present a useful method to compute the bivariate signature of a system, which can be regarded as an extension of the method (from Boland [65]) for computing univariate signatures mentioned in Sect. 7.1. It should be pointed out that the method is based on the discussion on the combinatorial meaning of bivariate tail signatures $\bar{S}_{i,j}$ given in Sect. 2.1 of Gertsbakh et al. [182].

For a given 3-state system, let us use $C = (c_1, \ldots, c_n)$ to represent the n components of the system. A subset P [L] of C is called a *path set* [*perfection path set*] if the functioning of the components in P [L] implies that the system is not in the complete failure state (is in the perfection state). Denote by $d_{i,j}$ the number of the pairs of a perfect path set of size i and a path set of size j such that the latter is a subset of the former, that is,

$$d_{i,j} = \#\{(L,P): \ L \in \mathcal{L}_i, P \in \mathcal{P}_j, P \subset L\},$$

where $\mathcal{L}_i$ and $\mathcal{P}_j$ represent the classes of all perfection path sets of size i and of all path sets of size j, respectively. We define

$$a_{i,j} := \frac{d_{i,j}}{\binom{n}{i-j,j}} \tag{7.3.1}$$

for $j = 1, \ldots, i$ and $i = 2, \ldots, n$, and $d_{i,j} = 0$ otherwise, where

$$\binom{m}{k,l} = \frac{m!}{k! \cdot l! \cdot (m-k-l)!}$$

is a multinomial coefficient. In fact, $a_{i,j}$ is the proportion of $d_{i,j}$ among all the pairs of two subsects of C with respective sizes i and j such that the second is a subset of the first.

Recall that T_1 and T_2 denote the degenerate times of a 3-state system, and F is the distribution function of component lifetimes of the system. Then we can write

$$\begin{aligned} & \mathrm{P}\{T_1 > t_1, T_2 > t_2\} \qquad (7.3.2) \\ = & \sum_{i=2}^{n}\sum_{j=1}^{i} a_{i,j} \cdot \binom{n}{i-j,j} F^{n-i}(t_1)[F(t_2)-F(t_1)]^{i-j}\bar{F}^j(t_2) \\ = & \sum_{i=1}^{n}\sum_{j=0}^{i} a_{i,j} \cdot \binom{n}{i-j,j} F^{n-i}(t_1)[F(t_2)-F(t_1)]^{i-j}\bar{F}^j(t_2) \\ = & \sum_{i=0}^{n-1}\sum_{j=i}^{n} a_{n-i,n-j} \cdot \binom{n}{i,j-i} F^{i}(t_1)[F(t_2)-F(t_1)]^{j-i}\bar{F}^{n-j}(t_2). \end{aligned}$$

According to Eqs. (7.2.7) and (7.3.2), we have the following proposition, which establishes the useful relationship between $a_{i,j}$ and the tail signatures $\bar{S}_{i,j}$.

Proposition 7.3.1. *For* $0 \le i \le j \le n$, $\bar{S}_{i,j} = a_{n-i,n-j}$.

Namely, Proposition 7.3.1 establishes a relationship between bivariate tail signatures $\bar{S}_{i,j}$ and $d_{i,j}$ for a 3-state system. From the knowledge of bivariate tail signatures one can obtain more information on the system structure. For example, we can obtain two additional quantities associated with $d_{i,j}$, denoted by $d_{i,\bar{j}}$ and $d_{\bar{i},j}$, where $d_{i,\bar{j}}$ is the number of the pairs of a perfect path set of size i and a non-path set of size j such that the latter is a subset of the former, and $d_{\bar{i},j}$ is the number of the pairs of a non-perfect path set of size i and a path set of size j such that the latter is a subset of the former. These two quantities will be useful in the study of Sect. 7.4. It is easily to see that

$$d_{i,\bar{j}} = |\mathcal{L}_i| \cdot \binom{i}{j} - d_{i,j} \tag{7.3.3}$$

and

$$d_{\bar{i},j} = |\mathcal{P}_j| \cdot \binom{n-j}{i-j} - d_{i,j}. \tag{7.3.4}$$

Further, according to Eq. (7.1.1), we have

$$|\mathcal{L}_i| = \binom{n}{i} \cdot \bar{S}^{(1)}_{n-i} \tag{7.3.5}$$

and

$$|\mathcal{P}_i| = \binom{n}{i} \cdot \bar{S}^{(2)}_{n-i}, \quad i = 1, \ldots, n. \tag{7.3.6}$$

7.4 The Signatures of Systems with Independent Modules

In this section, we restrict our attention to computing bivariate signatures of several classes of 3-state systems with independent modules, including the generalized series and parallel systems, the systems with componentwise redundancy, and an important class of 3-state systems. We employ the method presented in Sect. 7.3 to derive expressions for the bivariate signatures of these systems in terms of the signatures of their modules, respectively. By the way, in this section, some notations used in the proofs may not be explained, and one can promptly figure out their meanings according to Sect. 7.3.

7.4.1 Generalized Series and Parallel Systems

We restate the definitions of the generalized series and parallel systems given in Gertsbakh et al. [182] (see also Lisnianski and Levtin [283]). Consider two different 3-state systems with respective structure function:

$$\phi_1 : \{0,1\}^{n_1} \to \{0,1,2\},$$
$$\phi_2 : \{0,1\}^{n_2} \to \{0,1,2\}.$$

Denote $n = n_1 + n_2$. The "series system" of the two modules is defined as the system with structure function $\psi_S : \{0,1\}^n \to \{0,1,2\}$,

$$\psi_S(x_1, \ldots, x_{n_1}; y_1, \ldots, y_{n_2}) = \phi_1(x_1, \ldots, x_{n_1}) \wedge \phi_2(y_1, \ldots, y_{n_2}),$$

and the "parallel system" is defined as the system with structure function $\psi_P : \{0,1\}^n \to \{0,1,2\}$,

$$\psi_P(x_1, \ldots, x_{n_1}; y_1, \ldots, y_{n_2}) = \phi_1(x_1, \ldots, x_{n_1}) \vee \phi_2(y_1, \ldots, y_{n_2}).$$

Gertsbakh et al. [182] derived a formula for computing the bivariate tail signature of the generalized series system ψ_S in terms of the tail signatures of modules ϕ_1 and ϕ_2 by using Eq. (7.2.7) and pointed out that one can obtain such a formula for the cumulative signature of the generalized parallel system ψ_P in a similar way. In this section, we make use of the method presented in Sect. 7.3 to derive such formulas for the generalized series and parallel systems, and we can find that our method makes the derivation more easier.

Denote by $\bar{S}_{i,j}(\psi_S)$, $\bar{S}_{i,j}(\psi_P)$, $\bar{S}_{i,j}(\phi_1)$, and $\bar{S}_{i,j}(\phi_2)$ the tail signatures of the generalized series system ψ_S, the generalized parallel system ψ_P, the module ϕ_1 and the module ϕ_2, respectively.

Theorem 7.4.1. *For $0 \le i \le j \le n$,*

$$\bar{S}_{i,j}(\psi_S) = \frac{d_{n-i,n-j}(\psi_S)}{\binom{n}{i,\, j-i}},$$

where

$$d_{u,v}(\psi_S) = \sum_{l=l_0}^{n_1 \wedge u} \sum_{m=m_0}^{l \wedge v} \binom{n_1}{l-m,\, m} \binom{n_2}{n_2-u+l,\, v-m} \times \bar{S}_{n_1-l,n_1-m}(\phi_1) \cdot \bar{S}_{n_2-u+l,n_2-v+m}(\phi_2),$$

$l_0 = 2 \vee (u - n_2)$ and $m_0 = 1 \vee (v - u + l)$.

Proof: Fix u and v. We note that the fact that, for any pair (L, P) of the series system ψ_S such that $L \in \mathcal{L}_u(\psi_S)$, $P \in \mathcal{P}_v(\psi_S)$, and $P \subset L$, the components of the module ϕ_1 in L $[P]$ and the components of the module ϕ_2 in L [P] constitute a perfect path set [a path set] of size l_1 [m_1] for module ϕ_1 and a perfect path set [a path set] of size l_2 [m_2] for module ϕ_2, respectively, where $l_1 + l_2 = u$, $m_1 + m_2 = v$, $m_1 \le l_1$ and $m_2 \le l_2$. Thus, the number of the pairs (L, P) of the system ψ_S is given by

$$d_{u,v}(\psi_S) = \sum_{l=2\vee(u-n_2)}^{n_1\wedge u} \sum_{m=1\vee(v-u+l)}^{l\wedge v} d_{l,m}(\phi_1) \cdot d_{u-l,v-m}(\phi_2).$$

Then the desired result follows from Eq. (7.3.1) and Proposition 7.3.1. ∎

Theorem 7.4.2. *For $0 \le i \le j \le n$, we have*

$$\bar{S}_{i,j}(\psi_P) = \frac{d_{n-i,n-j}(\psi_P)}{\binom{n}{i,\, j-i}},$$

where

$$\begin{aligned} d_{u,v}(\psi_P) \;=\; & \sum_{k=1}^{2} \sum_{l=l_{3-k}}^{n_k\wedge u} \sum_{m=m_0}^{l\wedge v} \bar{S}_{n_k-l,n_k-m}(\phi_k) \cdot \binom{n_k}{l-m,\, m} \\ & \binom{n_{3-k}}{n_{3-k}-u+l,\, v-m} \\ & -\sum_{l=l_2}^{n_1\wedge u} \sum_{m=m_0}^{l\wedge v} \binom{n_1}{l-m,\, m}\binom{n_2}{n_2-u+l,\, v-m} \\ & \qquad \times \bar{S}_{n_1-l,n_1-m}(\phi_1) \cdot \bar{S}_{n_2-u+l,n_2-v+m}(\phi_2) \\ & +\sum_{k=1}^{2} \sum_{l=l_{3-k}}^{n_k\wedge u} \sum_{m=m_0'}^{(l-1)\wedge v} \left[\bar{S}^{(1)}_{n_k-l}(\phi_k) \cdot \binom{n_k}{l}\binom{l}{m} \right. \\ & \left. -\bar{S}_{n_k-l,n_k-m}(\phi_k) \cdot \binom{n_k}{l-m,\, m}\right] \\ & \times \left[\bar{S}^{(2)}_{n_{3-k}-v+m}(\phi_{3-k})\binom{n_{3-k}}{v-m}\binom{n_{3-k}-v+m}{u-v-l+m} \right. \\ & \left. -\bar{S}_{n_{3-k}-u+l,n_{3-k}-v+m}(\phi_{3-k}) \cdot \binom{n_{3-k}}{n_{3-k}-u+l,\, v-m}\right], \end{aligned}$$

and

$$l_1 = 2 \vee (u - n_1), \qquad l_2 = 2 \vee (u - n_2),$$

$$m_0 = 1 \vee (v - u + l), \qquad m_0' = 0 \vee (v - u + l),$$

$\bar{S}_a^{(1)}(\phi_k)$ and $\bar{S}_a^{(2)}(\phi_k)$ denote the type I and type II marginal tail signatures of module ϕ_k respectively, $a = 1, \ldots, n$, $k = 1, 2$.

Proof: Fix u and v. For any given pair (L, P) of two sets for the system ψ_P with $L \in \mathcal{L}_u(\psi_P)$, $P \in \mathcal{P}_v(\psi_P)$, and $P \subset L$, denote by L_k $[P_k]$ the components of module ϕ_k in L $[P]$, $k = 1, 2$. Then only one of the following statements holds:

1. L_1 is a perfection path set and P_1 is a path set for module ϕ_1 or L_2 is a perfection path set and P_2 is a path set for module ϕ_2.

2. L_1 is a perfection path set while P_1 is not a path set for module ϕ_1 and L_2 is not a perfection path set while P_2 is a partial path set for module ϕ_2.

3. L_1 is not a perfection path set while P_1 is a partial path set for module ϕ_1 and L_2 is a perfection path set while P_2 is not a path set for module ϕ_2.

Thus, the number of all of pairs (L, P) can be written as

$$d_{u,v}(\psi_P) = d_{u,v}^{(1)} + d_{u,v}^{(2)} + d_{u,v}^{(3)},$$

where $d_{u,v}^{(1)}$, $d_{u,v}^{(2)}$ and $d_{u,v}^{(3)}$ denote the numbers of pairs (L, P) satisfying (1), (2), and (3), respectively. Denote $l_k = |L_k|$ and $m_k = |P_k|$ for $k = 1, 2$. Obviously, $m_1 \leq l_1$, $m_2 \leq l_2$, $l_1 + l_2 = u$ and $m_1 + m_2 = v$. Then, it is not hard to get the following expressions:

$$\begin{aligned} d_{u,v}^{(1)} \quad = \quad & \sum_{l_1 = 2\vee(u-n_2)}^{n_1 \wedge u} \sum_{m_1 = 1\vee(v-u+l_1)}^{l_1 \wedge v} d_{l_1,m_1}(\phi_1) \cdot \binom{n_2}{u - l_1} \binom{u - l_1}{v - m_1} \\ & + \sum_{l_2 = 2\vee(u-n_1)}^{n_2 \wedge u} \sum_{m_2 = 1\vee(v-u+l_2)}^{l_2 \wedge v} d_{l_2,m_2}(\phi_2) \cdot \binom{n_1}{u - l_2} \binom{u - l_2}{v - m_2} \\ & - \sum_{l_1 = 2\vee(u-n_2)}^{n_1 \wedge u} \sum_{m_1 = 1\vee(v-u+l_1)}^{l_1 \wedge v} d_{l_1,m_1}(\phi_1) \cdot d_{u-l_1,v-m_1}(\phi_2), \end{aligned}$$

$$d_{u,v}^{(2)} = \sum_{l_1=2\vee(u-n_2)}^{n_1\wedge u} \sum_{m_1=0\vee(v-u+l_1)}^{(l_1-1)\wedge v} d_{l_1,\overline{m}_1}(\phi_1) \cdot d_{\overline{u-l_1},v-m_1}(\phi_2),$$

and

$$d_{u,v}^{(3)} = \sum_{l_2=2\vee(u-n_1)}^{n_2\wedge u} \sum_{m_2=0\vee(v-u+l_2)}^{(l_2-1)\wedge v} d_{\overline{u-l_2},v-m_2}(\phi_1) \cdot d_{l_2,\overline{m}_2}(\phi_2).$$

Now, by Eqs. (7.3.1) and (7.3.3)–(7.3.6) and Proposition 7.3.1 and after some simplifications, the desired formula is obtained. ■

7.4.2 Redundancy Systems

In reliability engineering, redundancy is well known since it is often used to improve the reliability of a coherent system. There are two types of redundancy. One is systemwise redundancy and the other is componentwise redundancy. Consider a coherent system with n components. Suppose one has an opportunity to enhance its performance by incorporating redundancy of n identical spares for the components. Systemwise redundancy is to place the n spares as an identical system in parallel with the original system, and componentwise redundancy is to place every spare separately in parallel with every component in the original system. A well-known principle is that componentwise redundancy is more effective than systemwise redundancy. One may refer to Barlow and Proschan [39] and Boland and EI-Newehi [66] for details on redundancy of traditional coherent systems and to EI-Neweihi et al. [145] for multi-state coherent systems.

Whether systemwise or componentwise redundancy, it is usually not easy to get its signature of a redundancy system since the number of its components is relatively large. So, can we compute the signature of a system with redundancy in terms of the signature of its original system? In the context of traditional coherent systems, Da et al. [103] have established a formula for computing the signature of a redundancy system in terms of the signature of its original system. In this section, we explore similar issues for 3-state systems. We note that the system with systemwise redundancy is just a special case of the generalized parallel systems discussed in Sect. 7.4.1. Thus, we next focus on the case of componentwise redundancy.

For a 3-state system containing n components with structure function ϕ, its componentwise redundancy system can be defined as the system with structure function

$$\psi_R(x_1, \ldots, x_n; y_1, \ldots, y_n) = \phi(x_1 \vee y_1, \ldots, x_n \vee y_n).$$

Denote by $\bar{S}_{i,j}(\psi_R)$ and $\bar{S}_{i,j}(\phi)$ the signatures of the system with componentwise redundancy ψ_R and of the original system ϕ. The following theorem gives an expression for $\bar{S}_{i,j}(\psi_R)$ in terms of $\bar{S}_{i,j}(\phi)$ only:

Theorem 7.4.3. *For $0 \le i \le j \le n$, we have*

$$\bar{S}_{i,j}(\psi_R) = \frac{d_{2n-i,2n-j}(\psi_R)}{\binom{2n}{i, j-i}},$$

where

$$\begin{aligned} d_{u,v}(\psi_R) \;&=\; \sum_{l=l_0}^{u \wedge n} \sum_{m=m_0}^{l \wedge v} \sum_{k=k_0}^{k_1} \binom{n}{l-m, m} \binom{2m-v}{k} \cdot c(v, m) \\ &\quad \cdot c(u-v-k, l-m) \cdot \bar{S}_{n-l,n-m}(\phi), \end{aligned}$$

$$l_0 = \lfloor \frac{u-1}{2} \rfloor + 1, \qquad m_0 = (v-u+l) \vee (\lfloor \frac{v-1}{2} \rfloor + 1),$$

$$k_0 = u - v - 2(l-m), \qquad k_1 = (2m - v) \wedge (u - v - l + m),$$

$$c(a, b) = 2^{2b-a} \binom{b}{a-b},$$

and $\lfloor x \rfloor$ denotes the largest integer not greater than x.

Proof: We regard the redundancy system as being composed of n modules, each being a parallel system containing two components—the original component and its identical back up one. Fix u and v. For each pair (L, P) with $L \in \mathcal{L}_u(\psi_R)$, $P \in \mathcal{P}_v(\psi_R)$, and $P \subset L$, there exist a unique $l \in \{\lfloor \frac{u-1}{2} \rfloor + 1, \ldots, u \wedge n\}$ and a unique $m \in \{(v-u+l) \vee (\lfloor \frac{v-1}{2} \rfloor + 1), \ldots, l \wedge v\}$ such that

1. The components in L are from l different modules, and the original components of the l modules constitute a perfect path set of size l for the original system.

2. The components in P are from m different modules, and the original components of the m modules constitute a path set of size m for the original system.

3. The path set of size m is a subset of the perfection path set of size l of the original system, where $m \geq v - u + l$ ensures that the v components can be from m different modules.

Thus, the collection of all pairs (L, P) with $L \in \mathcal{L}_u(\psi_R)$, $P \in \mathcal{P}_v(\psi_R)$, and $P \subset L$ can be written as

$$\bigcup_{l=l_0}^{u \wedge n} \bigcup_{m=m_0}^{l \wedge v} W_{l,m;u,v},$$

where $l_0 = \lfloor \frac{u-1}{2} \rfloor + 1$, $m_0 = (v-u+l) \vee (\lfloor \frac{v-1}{2} \rfloor + 1)$, and $W_{l,m;u,v}$ denotes the collection of all pairs (L, P) satisfying (1), (2), and (3). Note that for all (l, m), $W_{l,m;u,v}$ are pairwise mutually exclusive. Hence

$$d_{u,v}(\psi_R) = \sum_{l=l_0}^{u \wedge n} \sum_{m=m_0}^{l \wedge v} |W_{l,m;u,v}|.$$

Furthermore, observed that $|W_{l,m;u,v}|$ can be represented as

$$|W_{l,m;u,v}| = N_{l,m;u,v} \cdot d_{l,m}(\phi),$$

where, for given l and m, $N_{l,m;u,v}$ denotes the number of all pairs (L, P) satisfying (1), (2), and (3) above. Clearly, one can obtain $N_{l,m;u,v}$ as follows: Choose v components from the m different modules as P such that each of the m modules has at least one component to be chosen; then choose additionally $u-v$ as $L \backslash P$ from the l modules such that each of the other $l - m$ modules has at least one component to be chosen. Let k denote the number of the components of the m modules to be chosen in the second choice, and note that it should be $k_0 \leq k \leq k_1$ so as to ensure such choices can be realized, where $k_0 = u - v - 2(l - m)$ and $k_1 = (2m - v) \wedge (u - v - l + m)$. Since the number of all possible outcomes in choosing a components from b different modules such that each of the modules has at least one component to be chosen is

$$c(a, b) = 2^{2b-a} \cdot \binom{b}{a-b}$$

(see the proof of Theorem 3.4 of Da et al. [103]). Thus, $|N_{l,m;u,v}|$ can be computed as

$$|N_{l,m;u,v}| = c(v, m) \sum_{k=k_0}^{k_1} \binom{2m-v}{k} \cdot c(u - v - k, l - m).$$

Thus, the desired result now follows from Eq. (7.3.1) and Proposition 7.3.1. ∎

7.4.3 An Important Class of 3-State Systems

In this section, we consider an important class of 3-state systems which is defined as follows. Let

$$\phi_1 : \{0,1\}^{n_1} \to \{0,1\} \quad \text{and} \quad \phi_2 : \{0,1\}^{n_2} \to \{0,1\}$$

denote the structure functions of two different traditional coherent systems, respectively, and suppose there are not common components between the two systems. We consider a new 3-state system of order $n = n_1 + n_2$, whose structure function is defined as

$$\psi(x_1,\ldots,x_{n_1}; y_1,\ldots,y_{n_2}) = \phi_1(x_1,\ldots,x_{n_1}) + \phi_2(y_1,\ldots,y_{n_2}). \tag{7.4.1}$$

It is obvious that this new 3-state system satisfies all regularity assumptions given in Sect. 7.2. It should be pointed out that although we are not yet able to give more practical backgrounds for such systems, the construction is important from a theoretical point of view and it shows a method to generate a 3-state system by traditional coherent systems at least.

Now, our problem is how to compute the bivariate signature of the 3-state system ψ if the signatures of systems ϕ_1 and ϕ_2 are given. Let $(p_1,\ldots,p_{n_1})$ and $(q_1,\ldots,q_{n_2})$ be the signature vectors of two systems ϕ_1 and ϕ_2, respectively. Let $\bar{S}_{i,j}$, $0 \le i \le j \le n$, be the bivariate tail signatures of system ψ. Next we give an expression for $\bar{S}_{i,j}$ in terms of $(p_1,\ldots,p_{n_1})$ and $(q_1,\ldots,q_{n_2})$.

First we present a useful lemma.

Lemma 7.4.4. *For system ϕ_1, the number of the pairs (L,P) such that $L \in \mathcal{P}_l$, $P \notin \mathcal{P}_m$, $|P| = m$ and $P \subset L$ is given by*

$$r_{l,\overline{m}}(\phi_1) = \binom{n_1}{l-m,m} \cdot \left(\bar{P}_{n_1-l} - \bar{P}_{n_1-m}\right)$$

for $1 \le m \le l \le n_1$, where $\bar{P}_j = \sum_{k=j+1}^{n} p_k$, $j = 0,\ldots,n_1-1$.

Proof: Denote by $r_i\ (\phi_1)$ the number of path sets of size i for system ϕ_1. For $1 \le m \le l \le n_1$, the number of the pairs (L,P) with $L \in \mathcal{P}_l$, $P \in \mathcal{P}_m$, and $P \subset L$ is given by

$$r_m(\phi_1) \cdot \binom{n_1-m}{l-m}.$$

Then we can immediately get that

$$r_{l,\overline{m}}(\phi_1) = r_l(\phi_1) \cdot \binom{l}{m} - r_m(\phi_1) \cdot \binom{n_1 - m}{l - m}.$$

Thus, the desired result follows from Eq. (7.1.1). ■

Theorem 7.4.5. *The bivariate tail signatures of system ψ defined in Eq. (7.4.1) are given by*

$$\bar{S}_{i,j}(\psi) = \frac{d_{n-i,n-j}(\psi)}{\binom{n}{i, j-i}}, \qquad 0 \le i \le j \le n,$$

where

$$\begin{aligned} d_{u,v}(\psi) \ = \ & \sum_{l=1}^{n_1 \wedge u} \sum_{m=m_0}^{l \wedge v} \binom{n_1}{l-m, m} \binom{n_2}{u-l-v+m, v-m} \\ & \left(\bar{P}_{n_1-m} - \bar{P}_{n_1-l}\right) \left(\bar{Q}_{n_2-u+l} - \bar{Q}_{n_2-v+m}\right) \\ & + \sum_{l=1}^{n_1 \wedge u} \binom{u}{v} \binom{n_1}{l} \binom{n_2}{u-l} \bar{P}_{n_1-l} \cdot \bar{Q}_{n_2-u+l} \end{aligned}$$

with $m_0 = 0 \vee (v - u + l)$.

Proof: Fix u and v. By Eq. (7.3.3), the number of all pairs (L, P) with $L \in \mathcal{L}_u(\psi)$, $P \in \mathcal{P}_v(\psi)$ and $P \subset L$ can be computed as

$$d_{u,v}(\psi) = |\mathcal{L}_u(\psi)| \cdot \binom{u}{v} - d_{u,\overline{v}}(\psi),$$

where $d_{u,\overline{v}}(\psi)$ represents the number of all pairs (L, P) with $L \in \mathcal{L}_u(\psi)$, $P \notin \mathcal{P}_v(\psi)$, $|P| = v$, and $P \subset L$. For such a pair, denote by L_1 and L_2 [P_1 and P_2] the components of systems ϕ_1 and ϕ_2 in L [P], respectively. From the structure of system ψ, it follows that L_1 and L_2 are the path sets for systems ϕ_1 and ϕ_2, respectively, and neither P_1 is a path set for system ϕ_1 nor P_2 is a path set for system ϕ_2. Thus, $d_{u,\overline{v}}(\psi)$ can be obtained by the following formula:

$$d_{u,\overline{v}}(\psi) = \sum_{l=1}^{n_1 \wedge u} \sum_{m=0 \vee (v-u+l)}^{l \wedge v} r_{l,\overline{m}}(\phi_1) \cdot r_{u-l,\overline{v-m}}(\phi_2).$$

On the other hand, $|\mathcal{L}_u(\psi)|$ can be represented as

$$|\mathcal{L}_u(\psi)| = \sum_{l=1}^{n_1 \wedge u} r_l(\phi_1) \cdot r_{u-l}(\phi_2).$$

Thus, the desired result follows from Eq. (7.1.1), Lemma 7.4.4, and Proposition 7.3.1. ■

In the end, we point out an important theoretical point for the construction in Eq. (7.4.1). It is known that for traditional coherent systems, two different systems may possess a common signature (e.g., Kochar et al. [253]; Samaniego [408]), does this still hold for 3-state systems? The answer is yes. Since from Theorem 7.4.5, the bivariate tail signature (bivariate signature) of the 3-state system ψ only depends on the signatures of systems ϕ_1 and ϕ_2; we can construct two different 3-state systems which possess a common bivariate signature as follows. Let φ_1 and φ_2 be structure functions of two traditional coherent systems which are different but possess a common univariate signature and φ_0 be the structure function of another traditional coherent system. Consider two 3-state systems with respective structure functions given by $\varphi_0 + \varphi_1$ and $\varphi_0 + \varphi_2$. Obviously, these two systems are different, but according to Theorem 7.4.5, they have the same signature. A numerical example is given in the next section.

7.5 Some Examples

In this section, we compute the bivariate signatures for some 3-state systems by using the formulas given in Sect. 7.4. We continue to use the notations in the previous section.

Example 7.5.1. Consider two 3-state systems with respective structure functions ϕ_1 and ϕ_2 given by

$$\phi_1(x_1, x_2, x_3, x_4) = (x_1 + (x_2 \vee x_3) + x_4 - 1) \vee 0,$$

$$\phi_2(x_1, x_2, x_3, x_4) = (x_1 + (x_2 \wedge x_3) + x_4 - 1) \vee 0.$$

From Example 2 of Gertsbakh et al. [182], the bivariate tail signatures of the two systems are given by

$$\bar{S}(\phi_1) = \begin{pmatrix} 1 & 1 & \frac{5}{6} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \bar{S}(\phi_2) = \begin{pmatrix} 1 & 1 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now according to Theorems 7.4.1 and 7.4.2, the tail signatures of the generalized series system ψ_S and parallel system ψ_P composed of systems ϕ_1 and ϕ_2 can be computed as

$$\bar{S}(\psi_S) = \begin{pmatrix} 1 & 1 & \frac{23}{28} & \frac{1}{2} & \frac{1}{7} & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{6} & \frac{3}{70} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\bar{S}(\psi_P) = \begin{pmatrix} 1 & 1 & 1 & 1 & \frac{33}{35} & \frac{9}{14} & \frac{1}{4} & 0 \\ 1 & 1 & 1 & 1 & \frac{33}{35} & \frac{9}{14} & \frac{1}{4} & 0 \\ \frac{5}{7} & \frac{5}{7} & \frac{5}{7} & \frac{5}{7} & \frac{73}{105} & \frac{17}{35} & \frac{27}{140} & 0 \\ \frac{5}{14} & \frac{5}{14} & \frac{5}{14} & \frac{5}{14} & \frac{5}{14} & \frac{39}{140} & \frac{17}{140} & 0 \\ \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{1}{7} & \frac{31}{420} & 0 \\ \frac{1}{28} & \frac{1}{28} & \frac{1}{28} & \frac{1}{28} & \frac{1}{28} & \frac{1}{28} & \frac{1}{28} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Further, by Eq. (7.2.6), the bivariate signatures of the systems ψ_S and ψ_P are given by

$$s(\psi_S) = \begin{pmatrix} 0 & \frac{5}{28} & \frac{5}{21} & \frac{7}{30} & \frac{1}{10} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{12} & \frac{13}{105} & \frac{3}{70} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$s(\psi_P) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{4}{105} & \frac{19}{210} & \frac{1}{10} & \frac{2}{35} & 0 \\ 0 & 0 & 0 & \frac{2}{105} & \frac{11}{84} & \frac{19}{140} & \frac{1}{14} & 0 \\ 0 & 0 & 0 & 0 & \frac{11}{140} & \frac{37}{120} & \frac{1}{21} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{29}{420} & \frac{4}{105} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{28} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

■

Example 7.5.2. In Example 1 of Gertsbakh et al. [182], they considered a bridge network with four nodes and five edges and computed the signature for this system. In the present example, let us compute the signature for the bridge system with componentwise redundancy by using the formula given in Theorem 7.4.3. Denote by ϕ the structure function of the bridge system (original system). Then, from Gertsbakh et al. [182], the bivariate tail signature of ϕ is

$$\bar{S}(\phi) = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ \frac{4}{5} & \frac{4}{5} & \frac{4}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Now, according to Theorem 7.4.3 and Eq. (7.2.6), we can get the bivariate tail signature and the bivariate signature of the system with componentwise redundancy ψ_R as follows:

$$\bar{S}(\psi_R) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{8}{9} & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{8}{9} & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{8}{9} & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \frac{8}{9} & 0 \\ \frac{104}{105} & \frac{104}{105} & \frac{104}{105} & \frac{104}{105} & \frac{104}{105} & \frac{104}{105} & \frac{104}{105} & \frac{104}{105} & \frac{1388}{1575} & 0 \\ \frac{20}{21} & \frac{20}{21} & \frac{20}{21} & \frac{20}{21} & \frac{20}{21} & \frac{20}{21} & \frac{20}{21} & \frac{20}{21} & \frac{268}{315} & 0 \\ \frac{88}{105} & \frac{88}{105} & \frac{88}{105} & \frac{88}{105} & \frac{88}{105} & \frac{88}{105} & \frac{88}{105} & \frac{88}{105} & \frac{16}{21} & 0 \\ \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & \frac{8}{15} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$s(\psi_R) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{525} & \frac{4}{525} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{4}{525} & \frac{16}{525} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{315} & \frac{4}{45} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{105} & \frac{8}{35} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{8}{15} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

■

In the following example, we construct two different 3-state systems which possess a common signature according to discussion in Sect. 7.4.3.

Example 7.5.3. Let ϕ_0, ϕ_1 and ϕ_2 be the structure functions of three 2-state coherent systems given by

$$\begin{aligned} \phi_0(x_1, x_2, x_3) &= x_1 \wedge (x_2 \vee x_3), \\ \phi_1(x_1, x_2, x_3, x_4) &= (x_1 \vee x_2) \wedge (x_2 \vee x_4) \wedge (x_3 \vee x_4), \\ \phi_2(x_1, x_2, x_3, x_4) &= (x_1 \vee x_2) \wedge (x_1 \vee x_3) \wedge (x_1 \vee x_4) \wedge (x_2 \vee x_3 \vee x_4). \end{aligned}$$

From Samaniego [408] (see also Kochar et al. [253]), systems ϕ_1 and ϕ_2 have the same signature $(0, 1/2, 1/2, 0)$, and the signature of ϕ_0 is $(1/3, 2/3, 0)$. We now consider two 3-state systems with respective structure functions

$$\psi_1(x_1, \ldots, x_7) = \phi_0(x_1, x_2, x_3) + \phi_1(x_4, x_5, x_6, x_7),$$

and

$$\psi_2(x_1, \ldots, x_7) = \phi_0(x_1, x_2, x_3) + \phi_2(x_4, x_5, x_6, x_7).$$

Clearly, the two 3-state systems are different. However, from Theorem 7.4.5, they have the same bivariate signature given by

$$s(\psi_1) = s(\psi_2) = \begin{pmatrix} 0 & 0 & \frac{1}{35} & \frac{3}{70} & \frac{3}{70} & \frac{1}{35} & 0 \\ 0 & 0 & \frac{2}{35} & \frac{1}{10} & \frac{11}{105} & \frac{1}{14} & 0 \\ 0 & 0 & 0 & \frac{1}{7} & \frac{9}{70} & \frac{17}{210} & 0 \\ 0 & 0 & 0 & 0 & \frac{4}{35} & \frac{2}{35} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \qquad \blacksquare$$

Acknowledgments

The research is supported by the Fundamental Research Funds for the Central Universities and the NNSF of China (Nos. 11071232,70821001, 71090401).

Chapter 8
Stochastic Comparisons of Cumulative Entropies

Antonio Di Crescenzo and Maria Longobardi

Abstract: The cumulative entropy is an information measure which is alternative to the differential entropy and is connected with a notion in reliability theory. Indeed, the cumulative entropy of a random lifetime X can be expressed as the expectation of its mean inactivity time evaluated at X. After a brief review of its main properties, in this paper, we relate the cumulative entropy to the cumulative inaccuracy and provide some inequalities based on suitable stochastic orderings. We also show a characterization property of the dynamic version of the cumulative entropy. In conclusion, a stochastic comparison between the empirical cumulative entropy and the empirical cumulative inaccuracy is investigated.

A. Di Crescenzo (✉)
Dipartimento di Matematica, Università di Salerno,
Via Ponte don Melillo, I-84084 Fisciano, Italy
e-mail: adicrescenzo@unisa.it

M. Longobardi
Dipartimento di Matematica e Applicazioni, Università di Napoli Federico II, Via Cintia, I-80126 Napoli, Italy
e-mail: maria.longobardi@unina.it

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_8,

8.1 Introduction

In the last 40 years stochastic orders have attracted an increasing number of authors, who used them in several areas of probability and statistics, with applications in many fields, such as reliability theory, queueing theory, survival analysis, operations research, mathematical finance, risk theory, management science and biomathematics. Indeed, stochastic orders are often invoked not only to provide useful bounds and inequalities but also to compare stochastic systems. A landmark in this area is the book by Shaked and Shanthikumar [426], which represents an essential reference for a large number of researchers dealing with stochastic orderings. To give an idea of its broad impact we notice that up to now it has received more than 2,000 citations in the literature.

The aim of this paper is twofold: to give a brief review on the properties of an information measure recently introduced by the authors and to provide some new results, including simple examples of applications of stochastic orders to related notions of information theory.

It is well known that the basic way to establish if one random variable is "larger" than another is based on the comparison of their distributions functions. Formally, given two random variables X and Y, we say that X is smaller than Y in the usual stochastic order, denoted by $X \leq_{\mathrm{st}} Y$, if and only if

$$\mathrm{E}[\phi(X)] \leq \mathrm{E}[\phi(Y)] \tag{8.1.1}$$

for all increasing functions $\phi : \mathbb{R} \to \mathbb{R}$ for which the expectations exist (see Sect. 1.A.1 of [426]). Equivalently, $X \leq_{\mathrm{st}} Y$ if and only if $\mathrm{P}\{X \leq t\} \geq \mathrm{P}\{Y \leq t\}$ for all $t \in \mathbb{R}$. Another stochastic order that will be used in this paper is the decreasing convex order, denoted by $X \leq_{\mathrm{dcx}} Y$, which holds if and only if Eq. (8.1.1) is true for all decreasing convex functions $\phi : \mathbb{R} \to \mathbb{R}$ for which the expectations exist. We remark that the notion of dcx order is counterintuitive, in the sense that if $X \leq_{\mathrm{dcx}} Y$, then X is "larger" than Y in some stochastic sense (see Sect. 4.A.1 of [426]).

Let us now recall some preliminary notions of information theory. The concept of entropy was introduced by Claude Shannon [432] as a measure of the uncertainty associated with a discrete random variable. Formally, for a random variable X with possible values $\{x_1, \ldots, x_n\}$ and probability mass function $p(\cdot)$, the *entropy* is given by

$$H(X) = -\mathrm{E}[\log_b p(X)] = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i), \tag{8.1.2}$$

where b, the base of the logarithm, is usually equal to 2, e, or 10. Entropy is the minimum descriptive complexity of a random variable X, in the sense that it quantifies the expected value of the information contained in a realization of X. For a thorough description of its role in coding theory, compression schemes and other fields of information theory see [96], for instance. A comprehensive description of information-theoretic methodologies, based on focal measures such as Shannon entropy and Kullback–Leibler information, is given in [142].

A suitable extension of the Shannon entropy to the absolutely continuous case is the so-called differential entropy, which is a shift-independent functional given by

$$H(X) = -\mathrm{E}[\log f_X(X)] = -\int_{-\infty}^{\infty} f_X(x) \log f_X(x)\,\mathrm{d}x, \tag{8.1.3}$$

where $\log = \log_e$ and $f_X(x)$ is the probability density function of an absolutely continuous random variable X having support in $\mathbb{R}$. However, although the analogy is between definitions (8.1.2) and (8.1.3), the differential entropy is an inaccurate extension of the Shannon discrete entropy. Indeed, the latter is not invariant under changes of variables and can even become negative.

Various alternatives for the entropy of a continuous distribution have been proposed in the literature. In Sect. 4 of [208] the following notion is suggested:

$$H_m(X) = -\int_{-\infty}^{\infty} f_X(x) \log \frac{f_X(x)}{m(x)}\,\mathrm{d}x, \tag{8.1.4}$$

where $m(x)$ is a suitable invariant measure. More recently, the "measure problem" involving Eq. (8.1.4) has been encountered in [313]. Another example of information notion is due to [414], who proposed a measure that, unlike entropy, can be easily and consistently extended to the continuous probability distributions on interval $[a, b]$ and, unlike differential entropy, is always positive and invariant with respect to linear transformations of coordinates. A "length-biased" shift-dependent measure of uncertainty that stems from the differential entropy is a weighted entropy (see [127]):

$$H^w(X) = -\mathrm{E}[X \log f_X(X)] = -\int_{0}^{+\infty} x\, f_X(x) \log f_X(x)\,\mathrm{d}x, \tag{8.1.5}$$

which assigns larger weights to larger values of a non-negative random variable X.

Moreover, the cumulative residual entropy is defined as (see [385])

$$\mathcal{E}(X) = -\int_{-\infty}^{+\infty} \bar{F}_X(x) \log \bar{F}_X(x)\,\mathrm{d}x, \tag{8.1.6}$$

where $\bar{F}_X(x) = \mathrm{P}\{X > x\}$ is the cumulative residual distribution, or survival function, of a random variable X. Various applications of Eq. (8.1.6) are given in [20, 467–469].

In Sect. 8.2 we recall an information measure, named "cumulative entropy", defined by substituting the survival function $\bar{F}_X(x)$ with the distribution function of X in Eq. (8.1.6). Evaluations of the cumulative entropy for some distributions over finite and infinite domains are explicitly given. We also present various properties of such measure. In particular, we relate the cumulative entropy to the cumulative inaccuracy and recall that it can be expressed as the expectation of the mean inactivity time evaluated at X. Section 8.3 is devoted to provide some bounds and inequalities involving the cumulative entropy, for which use of stochastic orders is made. In Sect. 8.4 the dynamic version of the cumulative entropy is recalled, and a characterization property is provided. Finally, in Sect. 8.5, we illustrate some features of a simple estimator of the cumulative entropy based on the sample spacings. The empirical cumulative inaccuracy is also introduced, and a stochastic comparison between such empirical measures is provided.

Note that throughout this chapter, the terms "increasing" and "decreasing" are used in non-strict sense.

8.2 Cumulative Entropy

An information measure similar to Eq. (8.1.6) is the cumulative entropy, defined as (see [128])

$$\mathcal{CE}(X) = -\int_{-\infty}^{+\infty} F_X(x) \log F_X(x)\,\mathrm{d}x, \tag{8.2.1}$$

where $F_X(x) = \mathrm{P}\{X \leq x\}$ is the cumulative distribution function of a random variable X. The measure $\mathcal{CE}(X)$ is defined similarly to the differential entropy (8.1.3). However, since the argument of the logarithm is a probability, we have

$$0 \leq \mathcal{CE}(X) \leq +\infty,$$

whereas $H(X)$ may be negative. Moreover, $\mathcal{CE}(X) = 0$ if and only if X is a constant. From Eqs. (8.1.6) and (8.2.1) it follows that the cumulative entropy and the cumulative residual entropy are related by the following relation (see [131]):

$$\mathcal{E}(X) + \mathcal{CE}(X) = \int_{-\infty}^{+\infty} h(x)\,\mathrm{d}x,$$

where

$$h(x) = -[F_X(x)\log F_X(x) + \bar{F}_X(x)\log \bar{F}_X(x)], \qquad x \in \mathbb{R}$$

is the partition entropy of X evaluated at x (see [78]).

The cumulative entropy is evaluated in Table 8.1 for various examples of even probability density functions of standard random variables.

We point out that if $Y = aX + b$, with $a \in \mathbb{R}$, $a \neq 0$ and $b \in \mathbb{R}$, then

Table 8.1: Cumulative entropies for some standard random variables with even densities

$f_X(x)$	Support	$\mathcal{CE}(X)$
$\dfrac{9\sqrt{3}}{10\sqrt{10}}x^2$	$-\sqrt{5/3} < x < \sqrt{5/3}$	0.789790
$\dfrac{1}{2\sqrt{3}}$	$-\sqrt{3} < x < \sqrt{3}$	0.866025
$\dfrac{3}{4\sqrt{5}}\left(1 - \dfrac{x^2}{5}\right)$	$-\sqrt{5} < x < \sqrt{5}$	0.885688
$\dfrac{15}{784\sqrt{7}}\left(7 - x^2\right)^2$	$-\sqrt{7} < x < \sqrt{7}$	0.892215
$\dfrac{1}{6}\left(\sqrt{6} - \lvert x\rvert\right)$	$-\sqrt{6} < x < \sqrt{6}$	0.892953
$\dfrac{3}{20\sqrt{10}}\left(\lvert x\rvert - \sqrt{10}\right)^2$	$-\sqrt{10} < x < \sqrt{10}$	0.900979
$\dfrac{1}{\sqrt{2}}\,\mathrm{e}^{-\sqrt{2}\,\lvert x\rvert}$	$-\infty < x < \infty$	0.901835
$\dfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-x^2/2}$	$-\infty < x < \infty$	0.903197

Table 8.2: Cumulative entropies for some non-negative variables with mean 1 and variance 1

$F_X(x)$	Support: $0 < x < +\infty$	$\mathcal{CE}(X)$
$\dfrac{\Gamma\left(3, \frac{2}{x}\right)}{\Gamma(3)}$	(Inverse-gamma distribution)	0.474543
$\dfrac{x^3(20 + x(15 + x(6 + x)))}{(1 + x)^6}$	(Beta prime distribution)	0.556511
$\dfrac{1}{2} + \dfrac{1}{2}\mathrm{erf}\left(\dfrac{\log x + 0.5 \log 2}{\sqrt{2 \log 2}}\right)$	(Lognormal distribution)	0.565746
$1 - e^{-x}$	(Exponential distribution)	0.644934

$$\mathcal{CE}(Y) = |a| \cdot \begin{cases} \mathcal{CE}(X) & \text{if } a > 0, \\ \mathcal{E}(X) & \text{if } a < 0. \end{cases}$$

Other features of $\mathcal{CE}(X)$, such as properties of its two-dimensional version and a normalized cumulative entropy defined as $\mathcal{NCE}(X) = \mathcal{CE}(X)/\mathrm{E}(X)$ for $0 < \mathrm{E}(X) < +\infty$, were discussed in [128].

Table 8.2 shows the cumulative entropy of some non-negative random variables having unity mean and variance.

We notice that an extension of the cumulative entropy has been proposed by Abbasnejad [1], namely, the failure entropy of order α defined as

$$\mathcal{FE}_\alpha(X) = -\frac{1}{\alpha - 1} \log \int_0^{+\infty} F_X^\alpha(x) \, \mathrm{d}x,$$

for $\alpha > 0$, $\alpha \neq 1$.

Furthermore, we recall that a weighted version of the cumulative entropy has been defined recently as (see [321])

$$\mathcal{CE}^w(X) = -\int_0^{+\infty} x \, F_X(x) \log F_X(x) \, \mathrm{d}x,$$

in analogy with the weighted entropy (8.1.5).

8.2.1 Connections to Reliability Theory

Let us now recall various connections between the cumulative entropy and concepts in reliability theory.

Let X be a non-negative random variable that represents the random lifetime of a reliability system. Denote by $[X \mid B]$ a random variable whose distribution is identical to that of X conditional on an event B. The residual lifetime $[X-t\,|\,X>t]$, $t>0$, describes the time length between the failure time X and the inspection time t, given that at time t the system is still active. One of the most used functions to describe the aging process of a system is the mean residual life of X, given by

$$\mathrm{mrl}(t)=\mathrm{E}[X-t\,|\,X>t]=\frac{1}{\bar F_X(t)}\int_t^{+\infty}\bar F_X(x)\mathrm{d}x,\quad \forall\, t\ge 0:\ \bar F_X(t)>0, \tag{8.2.2}$$

which uniquely determines the distribution function of X. Its properties in the description of systems composed by finite mixtures are pinpointed in [342]. Properties of the mean residual life function in a renewal process and relationships with other relevant functions of reliability theory are examined in [351].

Information measures have been proposed in the past as a tool to explore the information content in random lifetimes. We recall [141], where a new partial ordering among life distributions in terms of their uncertainties is introduced and is used to assess the notion of a "better system". See also [139, 140], where a direct approach to measure uncertainty in the residual lifetime distribution has been addressed. Further developments involving new properties of the proposed measure in connection to order statistics and record values are then derived in [19].

Theorem 2.1 of [20] shows that the cumulative residual entropy (8.1.6) can be expressed in terms of Eq. (8.2.2) as

$$\mathcal{E}(X)=\mathrm{E}[\mathrm{mrl}(X)]. \tag{8.2.3}$$

A similar result holds for the cumulative entropy. We recall that, given that at time t a system has been found inactive, $[t-X\,|\,X\le t]$, $t>0$, describes the inactivity time of the system, i.e., the time elapsing between the inspection time t and the failure time X. The inactivity time is thus dual to the residual lifetime $[X-t\,|\,X>t]$. The mean inactivity time of X, given by

$$\tilde\mu_X(t)=\mathrm{E}[t-X\,|X\le t]=\frac{1}{F_X(t)}\int_0^t F_X(x)\,\mathrm{d}x,\qquad \forall\, t\ge 0:\ F_X(t)>0, \tag{8.2.4}$$

has been studied in reliability theory in [5, 6, 323], for instance. Similarly to Eq. (8.2.3), Theorem 3.1 of [128] shows that the cumulative entropy can be expressed as the expectation of the mean inactivity time evaluated at X, i.e.,

$$\mathcal{CE}(X) = \mathrm{E}[\tilde{\mu}_X(X)]. \tag{8.2.5}$$

We recall that the reversed hazard rate of a random lifetime X is given by (see [63])

$$\tau_X(t) = \frac{\mathrm{d}}{\mathrm{d}t} \log F_X(t) = \frac{f_X(t)}{F_X(t)}, \qquad t > 0: \; F_X(t) > 0. \tag{8.2.6}$$

The following decreasing convex function is defined as a double integral of the reversed hazard rate:

$$T_X^{(2)}(x) = -\int_x^{+\infty} \log F_X(z)\,\mathrm{d}z = \int_x^{+\infty} \left[\int_z^{+\infty} \tau_X(u)\,\mathrm{d}u\right] \mathrm{d}z, \quad x \geq 0. \tag{8.2.7}$$

Its derivative is closely related to the distribution function of X. Indeed, from Eq. (8.2.7), we have

$$\dot{T}_X^{(2)}(x) := \frac{\mathrm{d}}{\mathrm{d}x} T_X^{(2)}(x) = \log F_X(x) = -\int_z^{+\infty} \tau_X(u)\,\mathrm{d}u. \tag{8.2.8}$$

We recall that Proposition 3.1 of [128] provides the following alternative expression of the cumulative entropy of X:

$$\mathcal{CE}(X) = \mathrm{E}\left[T_X^{(2)}(X)\right], \tag{8.2.9}$$

with $T_X^{(2)}$ defined in Eq. (8.2.7).

Given two random lifetimes X and Y having distribution functions F_X and F_Y defined on $(0, \infty)$, let us now introduce the "cumulative inaccuracy"

$$K[F_X, F_Y] = -\int_0^{+\infty} F_X(u) \log F_Y(u)\,\mathrm{d}u, \tag{8.2.10}$$

as the cumulative analog of the measure of inaccuracy due to Kerridge [225]. Denoting the reversed hazard rate of Y as τ_Y, we set

$$T_Y^{(2)}(x) = -\int_x^{+\infty} \log F_Y(z)\,\mathrm{d}z = \int_x^{+\infty} \left[\int_z^{+\infty} \tau_Y(u)\,\mathrm{d}u\right] \mathrm{d}z, \quad x \geq 0. \tag{8.2.11}$$

Hereafter we give a probabilistic meaning of the cumulative inaccuracy in terms of Eqs. (8.2.7) and (8.2.11).

Proposition 8.2.1. *For non-negative absolutely continuous random variables X and Y, having distribution functions F_X and F_Y, we have*

$$K[F_X, F_Y] = \mathrm{E}\left[T_Y^{(2)}(X)\right], \qquad K[F_Y, F_X] = \mathrm{E}\left[T_X^{(2)}(Y)\right]. \quad (8.2.12)$$

The proof of Proposition 8.2.1 is omitted, being similar to that of Proposition 3.1 of [128].

We now aim to provide a connection between the information measures $\mathcal{CE}(X)$ and $K[\cdot,\cdot]$. Let X and Y be the random lifetimes of two systems which have finite unequal means and satisfy $X \geq_{\rm st} Y$ or $Y \geq_{\rm st} X$. Proposition 3.2 of [128] shows that if X is absolutely continuous and $\mathrm{E}[\tilde{\mu}_X(Y)]$ is finite, then

$$\mathcal{CE}(X) = \mathrm{E}[\tilde{\mu}_X(Y)] + \mathrm{E}[\tilde{\mu}'_X(Z)]\,[\mathrm{E}(X) - \mathrm{E}(Y)], \quad (8.2.13)$$

where $\tilde{\mu}'_X(t) = 1 - \tau_X(t)\tilde{\mu}_X(t)$, for all $t > 0$ such that $F_X(t) > 0$, and where Z has probability density function

$$f_Z(x) = \frac{F_Y(x) - F_X(x)}{\mathrm{E}(X) - \mathrm{E}(Y)}, \qquad x \geq 0. \quad (8.2.14)$$

Hereafter we state an identity similar to Eq. (8.2.13).

Proposition 8.2.2. *Let X and Y be non-negative random variables with finite unequal means and satisfying $X \geq_{\rm st} Y$ or $Y \geq_{\rm st} X$, with X absolutely continuous. If $K[F_Y, F_X]$ is finite, then*

$$\mathcal{CE}(X) = K[F_Y, F_X] + \mathrm{E}\left[\dot{T}_X^{(2)}(Z)\right][\mathrm{E}(X) - \mathrm{E}(Y)], \quad (8.2.15)$$

where $\dot{T}_X^{(2)}(\cdot)$ is given in Eq. (8.2.8) and where Z is an absolutely continuous non-negative random variable having probability density function (8.2.14).

Proof: It follows from identity (8.2.9), from the second of Eq. (8.2.12) and from the probabilistic analog of the mean value theorem given in [125]. ■

8.3 Inequalities and Stochastic Comparisons

In this section we shall focus on upper and lower bounds for the cumulative entropy and on some stochastic comparisons.

In [128] it has been proved that if X is a non-negative random variable, then

(i) $\mathcal{CE}(X) \geq C\,\mathrm{e}^{H(X)}$, where $C = \exp\left\{\int_0^1 \log(x\,|\log x|)\,\mathrm{d}x\right\} = 0.2065$

(ii) $\mathcal{CE}(X) \geq \int_0^{+\infty} F(x)\,\bar{F}(x)\,\mathrm{d}x$

(iii) $\mathcal{CE}(X) \geq -\int_\mu^{+\infty} \log F(z)\,\mathrm{d}z$

(iv) $\mathcal{CE}(X) \leq \mathrm{E}[X]$

(v) $\mathcal{CE}(X) \leq \mathrm{e}^{-1}\,b$

(vi) $\mathcal{CE}(X) \leq (b - \mathrm{E}[X])\,\left|\log\left(1 - \frac{\mathrm{E}[X]}{b}\right)\right|$

where bounds (v) and (vi) hold if X takes values in $[0, b]$, with b finite. The latter inequality can be generalized by means of the log-sum inequality (see, for instance, [384]). Indeed, Proposition 1 of [130] states that if X and Y take values in $[0, b]$, with b finite, and if $X \geq_{\mathrm{st}} Y$, then

$$\mathcal{CE}(X) \leq \mathcal{CE}(Y) + (b - \mathrm{E}[X])\,\left|\log \frac{b - \mathrm{E}[X]}{b - \mathrm{E}[Y]}\right|. \tag{8.3.1}$$

We remark that the inequality given in Eq. (8.3.1) is tighter than that given in Proposition 4.5 of [128], which holds under the same assumption, that is, $X \geq_{\mathrm{st}} Y$. When the stochastic ordering between X and Y is reversed, the following result holds:

Proposition 8.3.1. *If X and Y are non-negative random variables such that $X \leq_{\mathrm{st}} Y$, then*

$$K[F_Y, F_X] \leq \mathcal{CE}(X) \leq K[F_X, F_Y].$$

Proof: Since, by assumption, $F_X(t) \geq F_Y(t)$ for all $t \in \mathbb{R}$, the proof follows from Eqs. (8.2.1) and (8.2.10). ■

We remark that $X \leq_{\mathrm{st}} Y$ does not imply $\mathcal{CE}(X) \leq \mathcal{CE}(Y)$.

Proposition 8.3.2. *If X and Y are non-negative random variables such that $X \leq_{\mathrm{dcx}} Y$, then*

$$\mathcal{CE}(X) \leq K[F_Y, F_X].$$

Proof: Recalling the definition of decreasing convex order, Eq. (8.2.9) and the second of Eq. (8.2.12), the proof follows noting that Eq. (8.2.7) is a decreasing convex function. ■

We notice that Proposition 8.3.2 substitutes Proposition 4.6 of [128].

Example 8.3.3. Let X and Y have distribution functions $F_X(x) = \exp\{-c\,x^{-\gamma}\}$, $x > 0$, and $F_Y(x) = \exp\{-d\,x^{-\gamma}\}$, $x > 0$, with $c > 0$, $d > 0$ and $\gamma > 1$. From Eqs. (8.2.7) and (8.2.11) we have

$$T_X^{(2)}(x) = \frac{c}{\gamma - 1}\, x^{-\gamma+1}, \qquad T_Y^{(2)}(x) = \frac{d}{\gamma - 1}\, x^{-\gamma+1}, \qquad x > 0.$$

Hence, making use of Eq. (8.2.9) and the second of Eq. (8.2.12), we obtain

$$\mathcal{CE}(X) = \frac{c^{1/\gamma}}{\gamma}\,\Gamma\left(1 - \frac{1}{\gamma}\right), \qquad K[F_Y, F_X] = \frac{cd^{-1+1/\gamma}}{\gamma}\,\Gamma\left(1 - \frac{1}{\gamma}\right).$$

It immediately follows that if $c \geq d$, i.e., $X \leq_{\rm dcx} Y$, then $\mathcal{CE}(X) \leq K[F_Y, F_X]$, in agreement with Proposition 8.3.2. ■

We conclude this section by recalling two further inequalities stated in [128]:

– If X and Y are non-negative and independent random variables, then
$$\max\{\mathcal{CE}(X), \mathcal{CE}(Y)\} \leq \mathcal{CE}(X+Y).$$

– If $X_1, X_2, \ldots, X_n$ are non-negative i.i.d. random variables, then
$$\mathcal{CE}(n\,X_1) \geq \mathcal{CE}(\max\{X_1, X_2, \ldots, X_n\}).$$

8.4 Dynamic Cumulative Entropy

Dynamic information measures are often employed in system reliability to describe the effect of the age t on the uncertainty in random lifetimes. For instance, we recall the residual entropy [139] and the past entropy [126], defined as the differential entropy of $[X \mid X > t]$ and of $[X \mid X \leq t]$, respectively.

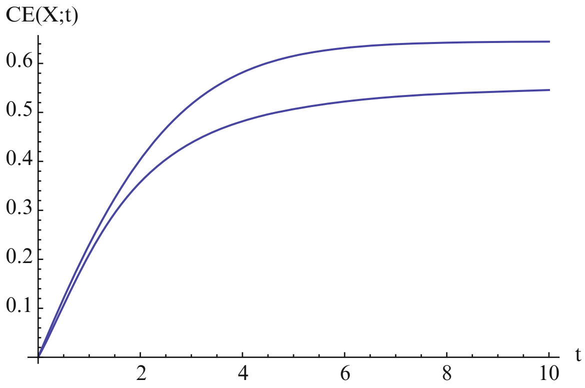


Figure 8.1: Dynamic cumulative entropy: beta prime (lower) and exponential distributions

The dynamic cumulative residual entropy was proposed by Asadi and Zohrevand [20] as the cumulative residual entropy of $[X \mid X > t]$, given by

$$\mathcal{E}(X;t) = -\int_t^{+\infty} \frac{\bar{F}_X(x)}{\bar{F}_X(t)} \log \frac{\bar{F}_X(x)}{\bar{F}_X(t)} \,\mathrm{d}x, \qquad t \geq 0. \tag{8.4.1}$$

Similarly to Eq. (8.4.1), the "dynamic cumulative entropy" was defined in [128] as the cumulative entropy of $[X \mid X \leq t]$, namely,

$$\mathcal{CE}(X;t) = -\int_0^t \frac{F_X(x)}{F_X(t)} \log \frac{F_X(x)}{F_X(t)} \,\mathrm{d}x, \qquad t > 0: \; F_X(t) > 0.$$

An alternative expression of $\mathcal{CE}(X;t)$ is given by

$$\mathcal{CE}(X;t) = -\frac{1}{F_X(t)} \int_0^t F_X(x) \log F_X(x) \mathrm{d}x + \tilde{\mu}_X(t) \log F_X(t), \quad t > 0 : F_X(t) > 0, \tag{8.4.2}$$

where $\tilde{\mu}_X(t)$ is the mean inactivity time defined in Eq. (8.2.4). We remark that $\mathcal{CE}(X;t)$ is non-negative for all t, with

$$\lim_{t \to 0^+} \mathcal{CE}(X;t) = 0, \qquad \lim_{t \to b^-} \mathcal{CE}(X;t) = \mathcal{CE}(X),$$

for any random variable X with support $(0,b)$, with $b \leq +\infty$. Figure 8.1 shows two cases where $\mathcal{CE}(X;t)$ is increasing in t. An instance of absolutely continuous distribution whose dynamic cumulative entropy is not increasing for all t is provided in Example 6.2 of

[128]. This paper also provides various properties of $\mathcal{CE}(X;t)$, such as lower and upper bounds, and the following two representations as conditional means:

$$\mathcal{CE}(X;t) = \mathrm{E}[\tilde{\mu}_X(X) \,|\, X \le t], \qquad t > 0,$$

and, when X is an absolutely continuous,

$$\mathcal{CE}(X;t) = \mathrm{E}[T_X^{(2)}(X;t) \,|\, X \le t], \qquad t > 0,$$

where

$$T_X^{(2)}(x;t) = -\int_x^t \log \frac{F(z)}{F(t)}\,\mathrm{d}z, \qquad t \ge x \ge 0. \tag{8.4.3}$$

Hereafter we give a characterization result for $\mathcal{CE}(X;t)$. To this purpose, we recall that (see Theorem 6.1 of [128]) $\mathcal{CE}(X;t)$ is increasing in t if and only if $\mathcal{CE}(X;t) \le \tilde{\mu}_X(t)$ for all $t > 0$ such that $F_X(t) > 0$.

Proposition 8.4.1. *If X is a non-negative absolutely continuous random variable and if $\mathcal{CE}(X;t)$ is increasing for all $t \ge 0$, then $\mathcal{CE}(X;t)$ uniquely determines $F_X(t)$.*

Proof: Differentiating Eq. (8.4.2) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{CE}(X;t) = \tau_X(t)\,[\tilde{\mu}_X(t) - \mathcal{CE}(X;t)]\,, \tag{8.4.4}$$

where $\tau_X(t)$ is given in Eq. (8.2.6). Hence, for any fixed t, the reversed hazard rate $\tau_X(t)$ is a positive solution of equation $g(x) = 0$, where

$$g(x) := x\,[\tilde{\mu}_X(t) - \mathcal{CE}(X;t)] - \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{CE}(X;t).$$

The assumption that $\mathcal{CE}(X;t)$ is increasing in t yields $\mathcal{CE}(X;t) \le \tilde{\mu}_X(t)$ for all t, so that $\lim_{x\to+\infty} g(x) = +\infty$ and $g(0) \le 0$, due to Eq. (8.4.4). Therefore, $g(x) = 0$ has a unique positive solution. Consequently $\tau_X(t)$, and hence $F_X(x)$, is uniquely determined by $\mathcal{CE}(X;t)$ under the assumption that such function is increasing in t. ■

We remark that Corollary 6.1 of [128] shows that $\mathcal{CE}(X;t)$ is increasing for all $t \ge 0$ if $\tilde{\mu}(t)$ is increasing for all $t \ge 0$. Such paper presents other results on the cumulative entropy, such as characterizations involving identities $\mathcal{CE}(X;t) = c\,\tilde{\mu}_X(t)$ and $\mathcal{CE}(X;t) = c\,\mu_X(t)$, where $\mu_X(t) = \mathrm{E}[X \,|\, X \le t]$ denotes the mean past lifetime of X. See also Sect. 4 of [340] for related results, such as an extension of a characterization of the power distribution that involves the cumulative entropy.

8.5 Empirical Cumulative Entropy

Let $X_1, X_2, \ldots, X_n$ be a random sample of non-negative, absolutely continuous i.i.d. random variables. A suitable estimator of $\mathcal{CE}(X)$ is the "empirical cumulative entropy", proposed in Sect. 7 of [128] as

$$\mathcal{CE}(\hat{F}_n) = -\int_0^{+\infty} \hat{F}_n(x) \log \hat{F}_n(x)\,\mathrm{d}x, \tag{8.5.1}$$

where

$$\hat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_i \le x\}}, \qquad x \in \mathbb{R}$$

is the empirical distribution of the sample. Denoting by $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ the sample order statistics and by

$$U_1 = X_{(1)}, \qquad U_i = X_{(i)} - X_{(i-1)}, \qquad i = 2, 3, \ldots, n$$

the corresponding sample spacings, it is not hard to prove that the empirical cumulative entropy can be expressed as

$$\mathcal{CE}(\hat{F}_n) = -\sum_{j=1}^{n-1} U_{j+1}\,\frac{j}{n}\log\frac{j}{n}. \tag{8.5.2}$$

Equation (8.5.2) shows that the empirical cumulative entropy is a positive linear combination of the sample spacings $U_2, \ldots, U_n$, where the outer spacings U_2 and U_n possess small weights, whereas the larger weight is given to the spacing U_{j+1} such that j is close to $\mathrm{e}^{-1}\,n \approx 0.3679\,n$. Equation (8.5.2) gives asymmetric weights to the sample spacings, so that the empirical cumulative entropy is asymmetric to the right. It is thus appropriate to measure variability in right-skewed distributions. A case study on neuronal firing data is provided in [131].

A discussion on $\mathcal{CE}(\hat{F}_n)$ in the case of random samples from uniform distribution and exponential distribution is given in [128]. Moreover, the following asymptotic results have been proved:

1. The standardized empirical cumulative entropy converges in distribution to a standard normal variable as $n \to +\infty$ [128].

2. $\mathcal{CE}(\hat{F}_n) \to \mathcal{CE}(X)$ a.s. as $n \to +\infty$ (see [129]).

We note that by use of identity $-u \log u \leq 1-u$, $0<u<1$, from Eq. (8.5.1), the following relation follows:

$$\mathcal{CE}(\hat{F}_n) \leq \overline{X} \quad \text{a.s.},$$

where $\overline{X}$ is the sample mean.

Let us now consider another random sample $Y_1, Y_2, \dots, Y_n$ of non-negative, absolutely continuous i.i.d. random variables and denote its empirical cumulative entropy by

$$\mathcal{CE}(\hat{G}_n) = -\int_0^{+\infty} \hat{G}_n(y) \log \hat{G}_n(y)\, \mathrm{d}y,$$

where $\hat{G}_n(y)$ is the empirical distribution of the sample. Moreover, in analogy with Eq. (8.2.10), we define the empirical cumulative inaccuracy as

$$K[\hat{F}_n, \hat{G}_n] = -\int_0^{+\infty} \hat{F}_n(u) \log \hat{G}_n(u)\, \mathrm{d}u.$$

It can be expressed as

$$K[\hat{F}_n, \hat{G}_n] = -\sum_{j=1}^{n-1} \int_{Y_{(j)}}^{Y_{(j+1)}} \hat{F}_n(u) \log \frac{j}{n}\, \mathrm{d}u, \tag{8.5.3}$$

where $Y_{(1)} < Y_{(2)} < \cdots < Y_{(n)}$ are the order statistics of the new sample. Let us denote by

$$N_j = \sum_{i=1}^{n} \mathbf{1}_{\{X_i \leq Y_{(j)}\}}, \qquad j = 1, 2, \dots, n,$$

the number of random variables of the first sample that are less than or equal to the jth order statistic of the second sample. Moreover, we rename by $X_{j,1} < X_{j,2} < \cdots$ the random variables of the first sample belonging to $(Y_{(j)}, Y_{(j+1)}]$, if any. From the above positions we thus have

$$\int_{Y_{(j)}}^{Y_{(j+1)}} \hat{F}_n(u)\, \mathrm{d}u = \frac{N_j}{n}\left[Y_{(j+1)} - Y_{(j)}\right] + \frac{1}{n} \sum_{r=1}^{N_{j+1}-N_j} \left[Y_{(j+1)} - X_{j,r}\right],$$

so that Eq. (8.5.3) becomes

$$K[\hat{F}_n, \hat{G}_n] = -\frac{1}{n} \sum_{j=1}^{n-1} \left[N_{j+1} Y_{(j+1)} - N_j Y_{(j)} - \sum_{r=1}^{N_{j+1}-N_j} X_{j,r} \right] \log \frac{j}{n}.$$

Cleary, $K[\hat{G}_n, \hat{F}_n]$ can be obtained by symmetry.

In analogy to Proposition 8.3.1, hereafter, we show that if the random variables of the two samples are stochastically ordered, then the empirical cumulative entropy and the empirical cumulative inaccuracies are suitably ordered.

Proposition 8.5.1. *If random variables X_i and Y_i satisfy condition $X_i \leq_{\rm st} Y_i$, then*

$$K[\hat{G}_n, \hat{F}_n] \leq_{\rm st} \mathcal{CE}(X) \leq_{\rm st} K[\hat{F}_n, \hat{G}_n].$$

Proof: Since $X_i \leq_{\rm st} Y_i$, from Theorem 1.A.3 of [426] we have that $\mathbf{1}_{\{X_i \leq x\}} \geq_{\rm st} \mathbf{1}_{\{Y_i \leq x\}}$, and thus $\hat{F}_n(x) \geq_{\rm st} \hat{G}_n(x)$, for all $x \in \mathbb{R}$. The proof then follows from the definitions of the involved notions. ■

Acknowledgments

This paper is dedicated to Moshe Shaked in admiration to his most profound contributions on stochastic orders. The corresponding author is grateful to Fabio Spizzichino for introducing Moshe Shaked to him in 1996, which started the collaboration behind the Di Crescenzo and Shaked [132] paper. Antonio Di Crescenzo and Maria Longobardi are partially supported by MIUR-PRIN 2008 "Mathematical models and computation methods for information processing and transmission in neuronal systems subject to stochastic dynamics".

Chapter 9

Decreasing Percentile Residual Life: Properties and Estimation

Alba M. Franco-Pereira, Jacobo de Uña, Rosa E. Lillo,

and Moshe Shaked

Abstract: In this work we describe the class of distribution functions with decreasing α-percentile residual life [DPRL(α)], $0 < \alpha < 1$. The close relationship between the aging notion DPRL(α) and increasing

---

A.M. Franco-Pereira (✉)
SiDOR Research Group, Universidad de Vigo, Vigo (Pontevedra), Spain
e-mail: alba.franco@uvigo.es

J. de Uña
Department of Statistics and Operations Research,
Universidad de Vigo, Vigoc (Pontevedra), Spain
e-mail: jacobo@uvigo.es

R.E. Lillo
Department of Statistics, Universidad Carlos III de Madrid, Getafe (Madrid), Spain
e-mail: lillo@est-econ.uc3m.es

M. Shaked
Department of Mathematics, University of Arizona, Tuscon, AZ 85721, USA.
e-mail: shaked@math.arizona.edu

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_9,

failure rate (IFR) will be discussed, as well as the relationship between the DPRL(α) and the percentile residual life orders. Besides, we introduce an estimator of the percentile residual life function, under the condition that it decreases, in the censored and the non-censored case. Finally, a real data illustration is provided.

9.1 Introduction

The residual life associated with a random variable is of interest in many areas of applied probability and statistics such as actuarial studies, biometry, survival analysis, economics, risk management, and reliability; see, e.g., [80, 336, 470]. Let X be a random variable, and let u_X be the right endpoint of its support. For any $t < u_X$, the *residual life* at time t is the random variable whose distribution is the conditional distribution of $X - t$ given that $X > t$. We denote it by

$$X_t = [X - t \big| X > t], \quad t < u_X.$$

Note that X_t is well defined for any $t < u_X$, even if t is not in the support of X. For example, even if X is a nonnegative random variable, we see that X_t is well defined for any $t < 0$. If F_X denotes the distribution function of X and $\overline{F}_X = 1 - F_X$ denotes the corresponding survival function, then the survival function of X_t is given by

$$\overline{F}_{X_t}(x) = \frac{\overline{F}_X(t+x)}{\overline{F}_X(t)}, \quad x \geq 0.$$

The expected value of X_t, which is called the mean residual life function that is associated with X, provides useful information about the underlying random variable. The median, or other percentiles, of the residual life of a random variable are, in some applications, useful alternatives to the mean residual life of that random variable. For any $t < u_X$, the α-percentile residual life function at t, $q_{X,\alpha}(t)$, is defined as the α-percentile or quantile of X_t. For $t \geq u_X$ we define $q_{X,\alpha}(t)$ to be zero. If F_X denotes the distribution function of X, then it can be shown that

$$q_{X,\alpha}(t) = F_X^{-1}(\alpha + (1-\alpha)F_X(t)) - t, \qquad t < u_X, \tag{9.1.1}$$

where $F_X^{-1}(p) = \inf\{x : F_X(x) \geq p\}$ is the so-called quantile function. Such a function describes, for example, the value that will be survived by $(1-\alpha)\%$ of items in reliability theory or of individuals in a medical study, among those that survived up to time t.

Statisticians have categorized life distributions according to different aging properties. These categories of distributions are useful for modeling situations where items deteriorate with age. A common approach is to stipulate the decreasingness of the mean residual life function (or of the harmonic mean residual life function) as a model of aging. This approach, however, sometimes has weaknesses that may prevent its use. For example, the mean residual life function may not exist. Or, even when it exists, it may have some practical shortcomings, especially in situations where the data are censored or when the underlying distribution is skewed or heavy-tailed. In such cases, either the empirical mean residual life function cannot be calculated or a single long-term survivor can have a marked effect upon it which will tend to be unstable due to its strong dependence on very long durations. Also, in an experiment, it is often impossible or impractical to wait until all items have failed. For those reasons, it is sometimes more convenient to consider the median residual life function or, more generally, the α-percentile residual life function instead, since it is less sensitive to outliers or censored data.

In this paper, we study random variables and distribution functions with decreasing α-percentile residual life functions [DPRL(α)], $0 < \alpha < 1$. Earlier papers, such as [167, 186, 215], studied some aspects of this class of distributions. This paper is organized as follows. In Sect. 9.2 we give the definition and some properties of the decreasing percentile residual life function. Besides, we analyze its relationship with the increasing hazard rate aging notion. Some characterizations of the percentile residual life orders, for DPRL(α) random variables, are reviewed in Sect. 9.3. In Sect. 9.4 we present an estimator of DPRL(α) function. It is then extended to the censored situation (Sect. 9.5). Finally we present the main conclusions and describe directions for further research.

9.2 Definitions and Basic Properties

Let $0 < \alpha < 1$ with $\overline{\alpha} = 1 - \alpha$. A random variable X is said to have (or to be) DPRL(α) if $q_{X,\alpha}(t)$ is decreasing in t. It is also possible to similarly define the notion of increasing α-percentile residual life

[IPRL(α)]. However, note that with our definition of $q_{X,\alpha}$, in order for a random variable to be IPRL(α), it is necessary that $u_X = \infty$.

Some useful equivalent conditions, which can help in practice to check the DPRL(α) condition, are given in the following proposition for absolutely continuous random variables with interval support (which may be finite or infinite). Its proof can be found in Franco–Pereira, Lillo, and Shaked ([167], Proposition 2.1).

Proposition 9.2.1. *Let X be an absolutely continuous random variable with interval support (l_X, u_X). Let f_X be the density function of X and $r_X \equiv f_X/\overline{F}_X$ its hazard rate function. Then, the following conditions are equivalent:*

(i) *X is* DPRL(α).

(ii) $\overline{\alpha} f_X(t) \leq f_X(\overline{F}_X^{-1}(\overline{\alpha}\,\overline{F}_X(t)))$ *for all* $t \in (l_X, u_X)$.

(iii) $\overline{\alpha} f_X(\overline{F}_X^{-1}(p)) \leq f_X(\overline{F}_X^{-1}(\overline{\alpha} p))$ *for all* $p \in (0, 1)$.

(iv) $r_X(t) \leq r_X(t + q_{X,\alpha}(t))$ *for all* $t \in (l_X, u_X)$.

From Proposition 9.2.1(iv) it follows that if r_X is increasing [i.e., if X has an increasing failure rate (IFR)] then X is DPRL(α) for any $\alpha \in (0, 1)$. On the other hand, if X is DPRL(α) for some $\alpha \in (0, 1)$, it is not necessary that X be IFR since, as it is shown in Example 3.1 in [167], IFR is a strictly stronger condition. In that example it was proved that given any $\varepsilon > 0$, even if X is DPRL(α) for every $\alpha \geq \varepsilon$, it is not necessary that X is IFR. Besides, Example 3.2 in the same paper shows that, given $\alpha \in (0, 1)$, it is possible to find a random variable X, and a $\beta \in (\alpha, 1)$, such that X is DPRL(α) but it is not DPRL(β). However, as we detail in next proposition, if the density function of X is decreasing on a specific region of its support, then, if X is DPRL(α), it does follow that X is DPRL(β) for $\beta > \alpha$.

Proposition 9.2.2. *Let X be an absolutely continuous random variable with interval support (l_X, u_X), such that $u_X < \infty$, and with density and survival functions f_X and $\overline{F}_X$, respectively. Let $\alpha \in (0, 1)$. If X is* DPRL(α) *and if f_X is increasing on $[\overline{F}_X^{-1}(\overline{\alpha}), u_X]$, then X is* DPRL(β) *for all $\beta > \alpha$.*

Note that if f_X is increasing on its support, then the monotonicity condition on f_X in Proposition 9.2.2 obviously holds. However, this observation does not tell us anything new because if f_X is increasing on its support, then X is IFR, and, as we noted at the beginning of this section, this implies that X is DPRL(α) for all $\alpha \in (0,1)$. The proof of Proposition 9.2.2 can be found in [167] (Proposition 3.3).

It is worthwhile to mention that Launer [273] has shown that a nonnegative random variable X, with a bathtub-shaped hazard rate function r_X, is DPRL(α) for all $\alpha \in (\alpha_0, 1)$ for some $\alpha_0 > 0$, provided there exists a $t_0 \geq 0$ such that $r_X(t_0) \geq r_X(0)$. More recently, in [165], this work was extended, providing characterization results for any kind of bathtub distributions. These characterizations are based on aging notions that link the percentile residual life function with the hazard rate function.

9.3 Relationship with the Percentile Residual Life Orders

A new family of stochastic orders indexed by α ($\alpha \in (0,1)$), which is based on the pointwise comparison of the percentile residual life functions of two random variables, was introduced and studied in detail in Franco–Pereira, Lillo, Romo, and Shaked [166]. These orders are the percentile residual life orders. Here we recall their definition.

Let X and Y be two random variables, let $\alpha \in (0,1)$, and let $q_{X,\alpha}$ and $q_{Y,\alpha}$ be their corresponding α-percentile residual life functions. If

$$q_{X,\alpha}(t) \leq q_{Y,\alpha}(t) \quad \text{for all } t, \tag{9.3.1}$$

then we say that X is smaller than Y in the *α-percentile residual life order*, and we denote it as $X \leq_{\alpha\text{-rl}} Y$. The α-percentile residual life orders were introduced in [216], but these orders were not extensively studied there. The focus of [216] was to test the hypothesis $H_0 : F_X = F_Y$ versus $H_1 : q_{X,\alpha} \leq q_{Y,\alpha}$ but it is should be noticed that, since the percentile residual life function does not characterize the distribution, this kind of tests may sometimes not be adequate.

Note that Eq. (9.3.1) defines a family of stochastic orders indexed by $\alpha \in (0,1)$. It follows from Eqs. (9.1.1) and (9.3.1) that if $X \leq_{\alpha\text{-rl}} Y$ then

$$u_X \leq u_Y,$$

where u_X and u_Y are the right endpoints of the corresponding supports.

In [167] some results relating the DPRL(α) and the α-percentile residual life orders were derived. In particular, under the DPRL(α)

aging notion, some closure properties of the α-percentile residual life order were obtained. Here we summarize some of the main results.

Theorem 9.3.1. *Let X be an absolutely continuous random variable with interval support. Then X is* DPRL(α) *if and only if any of the following equivalent conditions holds:*

(i) $X_t \geq_{\alpha\text{-rl}} X_{t'}$ *whenever* $t \leq t' < u_X$

(ii) $X \geq_{\alpha\text{-rl}} X_t$ *whenever* $0 \leq t < u_X$ (*when X is a nonnegative random variable*)

(iii) $X + t \leq_{\alpha\text{-rl}} X + t'$ *whenever* $t \leq t'$

In the literature there are results similar to Theorem 9.3.1, but which involve aging notions different from DPRL(α). For example, Theorems 1.A.30, 1.B.38, 3.B.24, 3.B.25, and 4.A.53 in [426], as well as a result in [49], give similar characterizations for the IFR aging notion. Also, Theorems 2.A.23, 2.B.17, 3.A.56, 3.C.13, and 4.A.51 in [426] give similar characterizations for the decreasing mean residual life (DMRL) aging notion.

The following result is an analog of Theorem 1.B.21 in [426] which involves the IFR aging notion and of Theorem 2.A.17 in [426] which involves the DMRL aging notion. The proof can be found in [167].

Theorem 9.3.2. *Let X be a positive, absolutely continuous,* DPRL(α) *random variable with interval support. Then,*

$$X \leq_{\alpha\text{-rl}} aX \quad \text{for all } a > 1.$$

Another situation in which the DPRL(α) aging notion arises as a natural condition will be described next. Theorem 9.3.3 below indicates a useful property of the order $\leq_{\alpha\text{-rl}}$ when one of the compared random variables is "larger in magnitude" than the other one; it is a generalization of the sufficiency part of Theorem 9.3.1(iii).

Theorem 9.3.3. *Let X be a continuous* DPRL(α) *random variable. Let Z be a nonnegative continuous random variable that is independent of X. Then,*

$$X \leq_{\alpha\text{-rl}} X + Z. \tag{9.3.2}$$

It is worthwhile to point out that if X in Theorem 9.3.3 is not DPRL(α) then the conclusion of that theorem need not hold. In order to see this, note that Theorem 9.3.1(iii) actually says that X is

DPRL(α) if and only if $X \leq_{\alpha\text{-rl}} X + a$ for every $a \geq 0$. Thus, if X in Theorem 9.3.3 is not DPRL(α) then there exists a degenerate Z such that Eq. (9.3.2) does not hold.

The DPRL(α) aging notion is also useful as a condition under which the order $\leq_{\alpha\text{-rl}}$ is preserved under certain random additions. This is stated next.

Theorem 9.3.4. *Let X and Y be two* DPRL(α) *random variables. Let Z be a random variable, independent of X and Y, with support in $[l, u]$, where $-\infty < l < u < \infty$. If $X + u \leq_{\alpha\text{-rl}} Y + l$, then*

$$X + Z \leq_{\alpha\text{-rl}} Y + Z.$$

9.4 Estimation

In many applications it is reasonable to assume that the system life is monotonically degrading or improving with age. In [167] the estimation of the mean residual life function under decreasing or increasing restrictions was studied and, following a similar approach, Franco–Pereira, Lillo, and Shaked [167] proposed an estimator for the decreasing percentile residual life function. In this section we first review this estimator of the percentile residual life function and its consistency.

Let $X_1, X_2, \ldots, X_n$ be independent random variables with a common distribution function F_X, and let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the corresponding order statistics. The resulting empirical distribution function is

$$F_{X,n}(t) = \frac{\#\{k : X_k \leq t, 1 \leq k \leq n\}}{n}, \quad t \in \mathbb{R},$$

and the corresponding left continuous inverse (i.e., the quantile function) is

$$F_{X,n}^{-1}(p) = X_{k:n} \quad \text{if} \quad \frac{k-1}{n} < p \leq \frac{k}{n}, \quad k = 1, 2, \ldots, n.$$

A natural empirical counterpart of $q_{X,\alpha}$ is the sample α-percentile residual life function, which is given by

$$\hat{q}_{X,n,\alpha}(t) = F_{X,n}^{-1}(\alpha + (1-\alpha)F_{X,n}(t)) - t, \quad t < X_{n:n}.$$

Note that $\hat{q}_{X,n,\alpha}$ is a piecewise linear function with jump discontinuities. It consists of line segments with slope equal to -1 with jump discontinuities (which gives rise to a rather ragged estimator). The estimator $\hat{q}_{X,n,\alpha}$ was introduced and studied in Csörgö and Csörgö (1987, [98]). Further properties of $\hat{q}_{X,n,\alpha}$ were obtained in [9, 33, 99, 100].

The estimator suggested in [167] is based on the fact that

$$q_{X,\alpha} \text{ is DPRL}(\alpha) \Longleftrightarrow q_{X,\alpha}(t) = \inf_{y \le t} q_{X,\alpha}(y). \tag{9.4.1}$$

Thus, the proposed estimator $\hat{q}^*_{X,n,\alpha}$ is given by

$$\hat{q}^*_{X,n,\alpha}(t) = I_{(t,\infty)}(X_{n:n}) \inf_{y \le t} \hat{q}_{X,n,\alpha}(y), \quad t \in \mathbb{R},$$

where $I_{(t,\infty)}$ denotes the indicator function of the indicated interval. Note that $\hat{q}^*_{X,n,\alpha}$ is the largest decreasing function that lies below the empirical $\hat{q}_{X,n,\alpha}$.

In Theorem 9.4.1 below we show that $\hat{q}^*_{X,n,\alpha}$ is a strongly uniform consistent estimator of $q_{X,\alpha}$. Its proof can be found in Franco–Pereira, Lillo, and Shaked ([167], Theorem 6.2).

Theorem 9.4.1. *Let X be a* DPRL(α) *random variable. If F_X has a continuous positive density function f_X such that $\inf_{0 \le p \le 1} f_X(F_X^{-1}(p)) > 0$, then $\hat{q}^*_{X,n,\alpha}$ is a strongly uniform consistent estimator of $q_{X,\alpha}$.*

To illustrate how the estimator looks like, consider a sample of size $n = 11$ with the ordered observed values $X_{1:11} = -5$, $X_{2:11} = X_{3:11} = -2$, $X_{4:11} = 1$, $X_{5:11} = X_{6:11} = 7$, $X_{7:11} = 11$, $X_{8:11} = 15$, $X_{9:11} = 16$, $X_{10:11} = 18$, and $X_{11:11} = 21$. Then there are $k = 9$ resulting ordered values with no ties:

$$Y_1 = -5,\ Y_2 = -2,\ Y_3 = 1,\ Y_4 = 7,\ Y_5 = 11,\ Y_6 = 15,\ Y_7 = 16,\ Y_8 = 18,\ \text{and } Y_9 = 21.$$

In Fig. 9.1 the empirical estimator $\hat{q}_{X,n,0.5}$ and the restricted estimator $\hat{q}^*_{X,n,0.5}$ are shown.

9.5 Extension of the Estimator to the Censored Case

The problem of censoring is common in survival analysis, and it occurs when the value of a observation is only partially known. For that reason, in Franco–Pereira and de Uña-Álvarez [164], the previous estimator of the percentile residual life function $\hat{q}_{X,n,\alpha}$ was extended to

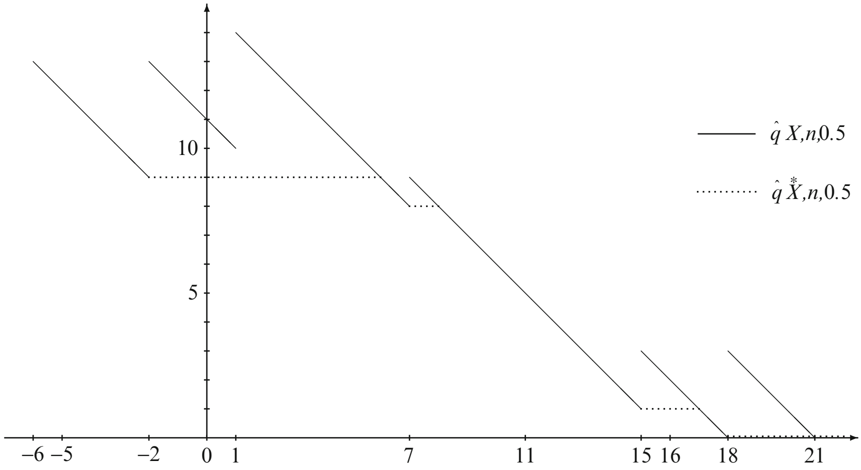


Figure 9.1: Illustration of the estimators $\hat{q}_{X,n,0.5}$ and $\hat{q}^*_{X,n,0.5}$

the censored case. In this section we review the main property of this new estimator (consistency) and show an application to a real data example.

Due to censoring, instead of the lifetime variables $X_1, X_2, \ldots, X_n$, one observes an independent and identically distributed sample (Z_1, δ_1), $\ldots$, (Z_n, δ_n) of the pair (Z, δ), where $Z_i = \min\{X_i, C_i)\}$ is the actual observed time, $\delta_i = I(X_i \le C_i)$ is the censoring indicator, and C_i is the potential censoring time. As usual, we assume that X_i and C_i are independent. In this setup, the nonparametric maximum-likelihood estimator of F_X is given by the Kaplan–Meier product-limit estimator

$$F_{X,n}(t) = 1 - \prod_{Z_{(i)} \le t} \left[1 - \frac{\delta_{[i]}}{n - i + 1} \right],$$

where $Z_{(1)} \le \cdots \le Z_{(n)}$ are the ordered Z-values, ties within lifetimes or within censoring times are ordered arbitrarily, and ties among lifetimes and censoring times are treated as if the former precedes the latter. Here, $\delta_{[i]}$ is the concomitant of the ith ordered statistics, that is, $\delta_{[i]} = \delta_j$ if $Z_{(i)} = Z_j$. Then, a natural nonparametric estimator of $q_{X,\alpha}(t)$, defined similarly as before, is

$$\hat{q}_{X,n,\alpha}(t) = F^{-1}_{X,n}\left(\alpha + (1-\alpha)F_{X,n}(t)\right) - t, \qquad t \le Z_{(n)},$$

where $F_{X,n}^{-1}(p) = \inf\{x : F_{X,n}(x) \geq p\}$ stands for the empirical quantile function that is associated with $F_{X,n}$. When $Z_{(n)}$ is uncensored, we have $F_{X,n}(Z_{(n)}) = 1$ and

$$\hat{q}_{X,n,\alpha}(Z_{(n)}) = F_{X,n}^{-1}(1) - Z_{(n)} = 0.$$

In this case, the estimator of the percentile residual life function $q_{X,\alpha}(t)$ is well defined for all t. However, when the maximum observed time is censored, we have $F_{X,n}(Z_{(n)}) < 1$ and the value $\hat{q}_{X,n,\alpha}(t)$ may not be well defined for large t. More explicitly, the function $\hat{q}_{X,n,\alpha}(t)$ is well defined only for $t \leq \tau_n$, where

$$\tau_n = \inf\left\{x : F_{X,n}(x) \geq \frac{F_{X,n}(Z_{(n)}) - \alpha}{1-\alpha}\right\}.$$

Certainly, for $t \leq \tau_n$, we have $F_{X,n}(t) \geq \left(F_{X,n}(Z_{(n)}) - \alpha\right)/(1-\alpha)$, and hence $\alpha + (1-\alpha)F_{X,n}(t) \geq F_{X,n}(Z_{(n)})$, from which we have that the set

$$\Theta_t = \{x : F_{X,n}(x) \leq \alpha + (1-\alpha)F_{X,n}(t)\}$$

is nonempty ($Z_{(n)}$ belongs to Θ_t). Therefore, $\hat{q}_{X,n,\alpha}(t)$ exists for $t \leq \tau_n$. As n grows, we have $\tau_n \to \tau \equiv F_X^{-1}((1-\alpha)^{-1}(F_X(b_H) - \alpha))$, where b_H is the upper limit of the support of Z. In words, it is not possible to estimate consistently the percentile residual life function beyond time τ. This may omit a portion of interest when b_H is smaller than the upper limit of the support of X. An analogous problem is found when recovering the cumulative distribution function $F_X(t)$ from the censored sample; in this case, consistency cannot be obtained for $t > b_H$. In this sense, the almost sure and in-probability uniform rates in Theorems 9.5.1 and 9.5.2 below, which hold on an interval $[0, T]$ where $T < b_H \wedge \tau$, are almost the most one can expect in this scenario.

If we assume that $q_{X,\alpha}(t)$ is monotone decreasing, then we have $q_{X,\alpha}(t) = \inf_{y \leq t} q_{X,\alpha}(y)$ and a natural estimator of the percentile residual life function is introduced [analogously as in Eq. (9.4.1)] through

$$\hat{q}^*_{X,n,\alpha}(t) = \inf_{y \leq t} \hat{q}_{X,n,\alpha}(y).$$

Some asymptotic properties of $\hat{q}^*_\alpha(t)$ are stated in the following results. Their proofs can be found in [164]. Specifically, a law of the iterated logarithm (LIL) and the $\sqrt{n}$-equivalence with respect to the unrestricted estimator are established. For finite sample sizes, however, the estimator $\hat{q}^*_{X,n,\alpha}(t)$ may outperform $\hat{q}_{X,n,\alpha}(t)$ since it incorporates the monotonicity information. See the simulation results in [164].

Put H for the distribution function of Z and $b_H = \inf\{t : H(t) = 1\}$ for the upper limit of the support of Z. Let $T < b_H \wedge \tau$, i.e., $T < b_H$ and $F_X^{-1}(\alpha + (1-\alpha)F_X(T)) < b_H$. Consider the following regularity conditions:

(C1) F_X is twice differentiable.

(C2) $f_X = F'_X$ is bounded away from zero on $\left[F_X^{-1}(\alpha), F_X^{-1}\right.$ $\left.(\alpha + (1-\alpha)F_X(T))\right]$.

Theorem 9.5.1 (LIL). *Under* (C1) *and* (C2) *we have, with probability* 1*,*

$$\sup_{0\le t\le T} \left|\hat{q}^*_{X,n,\alpha}(t) - q_{X,\alpha}(t)\right| = O\left(\left(\frac{\log\log n}{n}\right)^{1/2}\right).$$

Now, the $\sqrt{n}$-equivalence between the restricted and the unrestricted estimators of $q_{X,\alpha}(t)$ is established. From this second result, other asymptotic properties of the restricted estimator $\hat{q}^*_{X,n,\alpha}(t)$ (e.g., weak convergence) may be automatically obtained from those of $\hat{q}_{X,n,\alpha}(t)$, as, for instance, in [90]. See Corollaries 9.5.3 and 9.5.4 below. These are taken from [164].

Theorem 9.5.2. *Assume that, with* T *as in Theorem* 9.5.1*:*

(A1) $q'_{X,\alpha}(t)$ *exists and* $q'_{X,\alpha}(t) \le -c_1, 0 \le t \le T$*, for some* $c_1 > 0$*.*

(A2) $q''_{X,\alpha}(t)$ *exists and* $\sup_{0\le t\le T}\left|q''_{X,\alpha}(t)\right| \le c_2 < \infty$*.*

(A3) *conditions* (C1) *and* (C2) *above hold.*

Then we have

$$\sqrt{n}\sup_{0\le t\le T}\left|\hat{q}^*_{X,n,\alpha}(t) - \hat{q}_{X,n,\alpha}(t)\right| \to 0 \text{ in probability.}$$

The scaled product-limit α-percentile residual lifetime process $r^*_{X,n,\alpha}(t)$ is defined as

$$r^*_{X,n,\alpha}(t) = \sqrt{n} f_X(F_X^{-1}(\alpha + (1-\alpha)F_X(t)))[\hat{q}^*_{X,n,\alpha}(t) - q_{X,\alpha}(t)].$$

From our Theorem 9.5.2 and Theorem 6.1 in [90] we have the following result about the pointwise asymptotic distribution of $r^*_{X,n,\alpha}(t)$.

Corollary 9.5.3. *Under the assumptions of Theorem 9.5.2, the asymptotic distribution of* $r^*_{X,n,\alpha}(t)$ *is the normal with mean* 0 *and variance* $(1-\alpha)^2(1-F_X(t))^2[d(F_X^{-1}(\alpha+(1-\alpha)F_X(t)))-d(t)]$, *i.e.,*

$$r^*_{X,n,\alpha}(t) \longrightarrow_d N(0,(1-\alpha)^2(1-F_X(t))^2[d(F_X^{-1}(\alpha+(1-\alpha)F_X(t)))-d(t)])$$

as $n \to \infty$, *where* $\longrightarrow_d$ *denotes convergence in distribution,* $d(t) = \int_{-\infty}^{t}(1-H(s))^{-2}d\tilde{H}(s)$, *and* $\tilde{H}(t) = \int_{-\infty}^{t}\frac{(1-H(s))}{(1-F_X(s))}dF_X(s)$.

The following corollary follows from our Theorem 9.5.2 and Theorem 8.1 in [90]:

Corollary 9.5.4. *Under the assumptions of Theorem 9.5.2, we have, as* $n \to \infty$,

$$\sup_{0\le t\le T}\left|r^*_{X,n,\alpha}(t) - G_{X,\alpha}(t)\right| \overset{=}{a.s.} O\big(n^{-1/4}(\log n)^{1/2}(\log\log n)^{1/4}\big)$$

where $G_{X,\alpha}(t)$ *is the Gaussian process in* [90], *Eq.* (8.2), *namely,*

$$G_{X,\alpha}(t)=(1-\alpha)(1-F_X(t))\left[W(d(t))-W(d(F_X^{-1}(\alpha+(1-\alpha)F_X(t))))\right];$$

here W *is a standard Wiener process.*

9.5.1 A Real Data Example

For illustration purposes, we consider the primary biliary cirrhosis (PBC) data set reported and widely explained in Fleming and Harrington (1991), with $n = 312$ individuals. Here we replicate the study in [164]. The variable X denotes survival time (in days) for PBC patients. Censoring from the right is caused by the end of the follow-up period or by liver transplantation (187 censored times or about 60 % of censoring). It is known that the survival prognosis is greatly influenced by the level of edema, so we consider three different groups of patients according to this variable. The first group (edema=0) corresponds to patients with no edema; patients in second group (edema=0.5) had an untreated or a successfully treated edema, while the third group (edema=1) corresponds to patients with an unsuccessfully treated edema. In Table 9.1 we report the number of cases and deaths in each group, together with the median survival. From this table we see that an increasing value of edema is associated to a poorer survival prognosis.

Table 9.1: Number of cases and deaths in each group and median survival (in days)

Level of edema	Number of cases	Deaths	Median survival
0	263	89	3,584
0.5	29	17	1,576
1	20	19	299

In Fig. 9.2 we give the 0.25-percentile residual life function for the three groups of edema, when estimated by using the restricted or the unrestricted estimators. For the first group (edema=0), the unrestricted estimator suggests a decreasing shape; this is not surprising, since the convex cumulative hazard plot for this group (see Fig. 9.3) reveals an increasing hazard rate, which is a characteristic property of the decreasing percentile residual life populations, as we commented in Sect. 9.3. In this case, by using the monotone estimator, we get some smoothing of the curve which results in a nicer estimator. The other two groups offer a different situation, since the unrestricted estimator is not supporting in principle the monotonicity of the percentile residual life function. This could be explained by the existence of a nonincreasing hazard rate for the last two groups. Indeed, the corresponding Nelson–Aalen estimators (see, for instance, [77] for a review of the Nelson–Aalen estimators) depict a concave part (see Fig. 9.3). This suggests an increasing-decreasing shape for the hazard rate.

9.6 Discussion

In this paper, we review some properties of the decreasing percentile residual life aging notion, its relationship with the IFR aging notion and with the percentile residual life orders, and its estimation, including the random censorship setup. This estimator is convenient when investigating units which deteriorate with age. In such a case, the unrestricted estimator is not admissible, and the monotone estimator gives a proper modification of it. A LIL has been established. Besides, it has been demonstrated that the monotone estimator is $\sqrt{n}$-equivalent to the unrestricted one. As a consequence, the asymptotic normal distribution of the monotone estimator and its strong approximation to a Gaussian process have been established. Finally, a real data illustration has been provided.

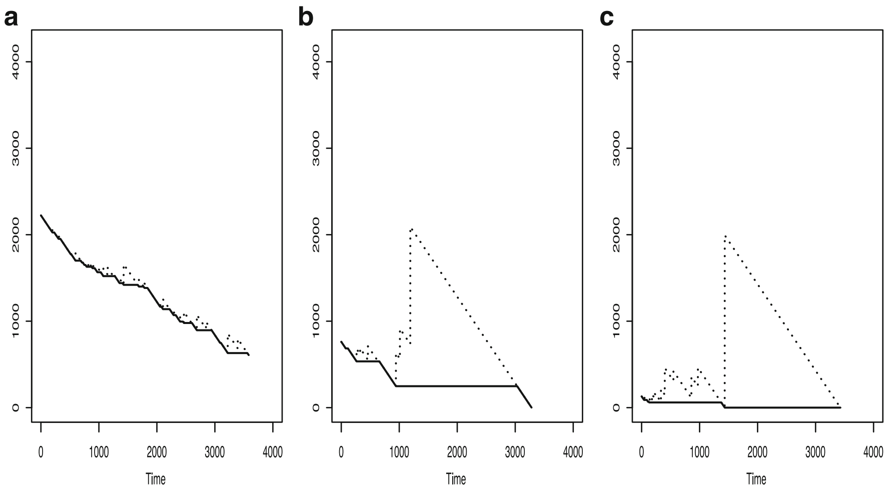


Figure 9.2: $\hat{q}_{X,n,0.25}$ (*dotted*) and $\hat{q}^{*}_{X,n,0.25}$ (*solid*) for the three groups of edema 0 (**a**) the group of edema 0.5 (**b**) the group of edema (**c**) the group of edema 1

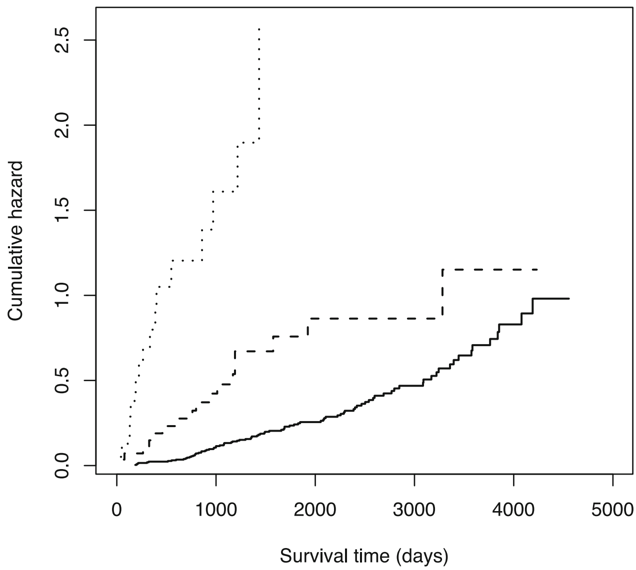


Figure 9.3: Nelson–Aalen's cumulative hazard estimates: 0 (*solid*), 0.5 (*dashed*), 1 (*dotted*)

A key question in practice is whether one should assume beforehand that the percentile residual life function is monotone. Our real data application shows that this is not always the case. It would be very interesting to develop goodness-of-fit tests for the monotonicity assumption. A possible way of doing that is through a proper distance between the restricted and the unrestricted estimators. This topic is currently under research. Finally, the application of the ideas in this paper for the estimation of monotone increasing residual life functions is possible, and completely analogous estimators are obtained in such a case.

Acknowledgments

Alba M. Franco-Pereira, Jacobo de Uña and Rosa E. Lillo are supported by the Grants MTM2011323204 (FEDER support included) of the Spanish Ministerio de Ciencia e Innovación, 10PXIB300068PR of the Xunta de Galicia, and CCG10-UC3M/HUM-5114 of the Comunidad de Madrid. Moshe Shaked is supported by NSA grant H98230-12-1-0222. INBIOMED project (DXPCTSUG, Ref. 2009/063) is also acknowledged.

Chapter 10
A Review on Convolutions of Gamma Random Variables

Baha-Eldin Khaledi and Subhash Kochar

Abstract: Due to its wide range of applications, the distribution theory of convolutions of gamma random variables has attracted the attention from many researchers. In this paper, we review some of the latest developments on this problem.

10.1 Introduction

The convolution of independent random variables has attracted considerable attention in the literature due to its typical applications in many applied areas. For example, in reliability theory, it is used to study the

B.-E. Khaledi (✉)
Department of Statistics, Razi University, Kermanshah, Iran
e-mail: bkhaledi@hotmail.com

S. Kochar
Fariborz Maseeh Department of Mathematics and Statistics,
Portland State University, Portland, OR 97201, USA
e-mail: kochar@pdx.edu

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_10,

lifetimes of redundant standby systems with independent components (cf. Bon and Păltănea [75]); in actuarial science, it is used to model the total claims on a number of policies in the individual risk model (cf. Kaas et al. [219]); in nonparametric goodness-of-fit tests, the limiting distributions of U-statistics are the convolutions of independent random variables (cf. Serfling [417], Sect. 5.2). As another example, let X_i denote the random value of ith shock on a system; then the system fails if the convolution of a number of X_i's exceeds the system's threshold (cf. Marshall and Olkin [311]). Therefore, study of lifetime of a standby system or a cumulative damage threshold model is based on stochastic properties of convolutions of random variables.

The gamma distribution is one of the most popular distributions in statistics, engineering, and reliability applications. In particular, gamma distribution plays a prominent role in actuarial science since most total insurance claim distributions have roughly the same shape as gamma distributions: skewed to the right, nonnegatively supported, and unimodal (cf. Furman [170]). As is well known, the gamma distribution includes exponential and chi-square, two important distributions, as special cases. Due to the complicated nature of the distribution function of gamma random variable, most of the work in the literature discusses only the convolutions of exponential random variables. Some relevant references are Khaledi [226], Boland et al. [71], Kochar and Ma [252], Bon and Păltănea [75], Zhao and Balakrishnan [488], and Kochar and Xu [259].

Let $X_1, \dots, X_n$ be a random sample from a gamma distribution with shape parameter $a > 0$, scale parameter $\lambda > 0$, and density function

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp\{-\lambda x\}, \quad x \geq 0.$$

We are interested in studying the stochastic properties of statistics of the form

$$W = \theta_1 X_1 + \theta_2 X_2 + \cdots + \theta_n X_n,$$

where $\theta_1, \dots, \theta_n$ are positive weights (constants). Bock et al. [64] showed that for $n = 2$, if

$$t \leq \frac{a(\theta_1 + \theta_2)}{\lambda},$$

then $P\{W \leq t\}$ is Schur convex in (θ_1, θ_2), and if

$$t \geq \frac{(a+1/2)(\theta_1+\theta_2)}{\lambda},$$

then $P\{W \geq t\}$ is Schur convex in (θ_1, θ_2). For general $n>2$, $\mathrm{P}\{W \leq t\}$ is Schur convex in the region

$$\left\{\boldsymbol{\theta} : \min_{1\leq i\leq n} \theta_i \geq \frac{t\lambda}{na+1}\right\},$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$, and $\mathrm{P}\{W \geq t\}$ is Schur convex in $\boldsymbol{\theta}$ for

$$t \geq \frac{(na+1)(\theta_1+\theta_2+\cdots+\theta_n)}{\lambda}.$$

Diaconis and Perlman [123] further studied the tail probabilities of convolution of gamma random variables. They pointed out that if

$$(\theta_1, \ldots, \theta_n) \overset{m}{\succeq} (\theta'_1, \ldots, \theta'_n) \tag{10.1.1}$$

then

$$\mathrm{Var}\left[\sum_{i=1}^{n} \theta_i X_i\right] \geq \mathrm{Var}\left[\sum_{i=1}^{n} \theta'_i X_i\right],$$

where $\overset{m}{\succeq}$ means the majorization order (see Definition 10.3.1).

This property states that the convolutions are more dispersed about their means as measured by their variances if the weights are more dispersed in the sense of majorization. Diaconis and Perlman [123] also wondered whether $\sum_{i=1}^{n} \theta_i X_i$ is more dispersed than $\sum_{i=1}^{n} \theta'_i X_i$ as measured by the stronger criterion of their tail probabilities. They tried to answer this question by proving that under the condition (10.1.1), the distribution functions of $\sum_{i=1}^{n} \theta_i X_i$ and $\sum_{i=1}^{n} \theta'_i X_i$ have only one crossing. However, they only proved this result for $n = 2$. For $n \geq 3$, they required further restrictions. Hence, this problem has been open for a long time, which is also known as *Unique Crossing Conjecture*.

The rest of this paper is organized as follows: In Sect. 10.2, we first review some stochastic orders and majorization orders. In Sect. 10.3, we investigate the crossing properties of two convolutions of gamma random variables under various conditions on the parameters for n=2. In Sect. 10.4, we establish the right spread ordering between two convolutions of independent gamma random variables for arbitrary n. We conclude our discussion with some remarks in the last section.

10.2 Preliminaries

In this section, we will review some notions of stochastic orders and majorization orders.

Assume that the positive random variables X and Y have distribution functions F and G, survival functions $\overline{F} = 1 - F$ and $\overline{G} = 1 - G$, density functions f and g, and failure rate functions $r_X = f/\overline{F}$ and $r_Y = g/\overline{G}$, respectively. The following orders are usually used to compare the magnitude of random variables:

Definition 10.2.1. X is said to be smaller than Y in the:

(i) Likelihood ratio order (denoted by $X \leq_{\mathrm{lr}} Y$) if $g(x)/f(x)$ is increasing in x

(ii) Hazard rate order (denoted by $X \leq_{\mathrm{hr}} Y$) if $\overline{G}(x)/\overline{F}(x)$ is increasing in x

(iii) Stochastic ordering (denoted by $X \leq_{\mathrm{st}} Y$) if $\overline{F}(x) \leq \overline{G}(x)$ for every x

(iv) Mean residual life order, denoted by $X \leq_{\mathrm{mrl}} Y$, if

$$\frac{\int_t^\infty \overline{F}(x)\mathrm{d}x}{\overline{F}(t)} \leq \frac{\int_t^\infty \overline{G}(x)\mathrm{d}x}{\bar{G}(t)}$$

It is known that (cf. Shaked and Shanthikumar [427])

$$X \leq_{\mathrm{lr}} Y \Longrightarrow X \leq_{\mathrm{hr}} Y \Longrightarrow X \leq_{\mathrm{mrl}} Y \Longrightarrow \mathrm{E}[X] \leq \mathrm{E}[Y],$$

and

$$X \leq_{\mathrm{lr}} Y \Longrightarrow X \leq_{\mathrm{hr}} Y \Longrightarrow X \leq_{\mathrm{st}} Y \Longrightarrow \mathrm{E}[X] \leq \mathrm{E}[Y].$$

The following order, called the dispersive order, is used to compare the variabilities of two random variables.

Definition 10.2.2. X is said to be less dispersed than Y (denoted by $X \leq_{\mathrm{disp}} Y$) if

$$F^{-1}(\beta) - F^{-1}(\alpha) \leq G^{-1}(\beta) - G^{-1}(\alpha)$$

for all $0 < \alpha \leq \beta < 1$.

A weaker order called the right spread order has also been proposed to compare the variabilities of two distributions (cf. Fernández-Ponce et al. [160]).

Definition 10.2.3. X is said to be less right spread than Y (denoted by $X \leq_{\mathrm{RS}} Y$) if

$$\int_{F^{-1}(p)}^{\infty} \overline{F}(x)\,\mathrm{d}x \leq \int_{G^{-1}(p)}^{\infty} \overline{G}(x)\mathrm{d}x, \qquad \text{for all } 0 \leq p \leq 1.$$

It is known that

$$X \leq_{\mathrm{disp}} Y \Longrightarrow X \leq_{\mathrm{RS}} Y \Longrightarrow \mathrm{Var}(X) \leq \mathrm{Var}(Y).$$

Bagai and Kochar [24] proved the following result:

Theorem 10.2.4. *If $X \leq_{\mathrm{disp}} Y$ and F or G is IFR (increasing failure rate), then $X \leq_{\mathrm{hr}} Y$.*

Definition 10.2.5. X is said to be smaller than Y in the star order, denoted by $X \leq_* Y$ (or $F \leq_* G$) if $G^{-1}F(x)/x$ is increasing in x on the support of X, where G^{-1} is the right continuous inverse of G.

It is known that if $X \leq_* Y$, then $\overline{F}(x)$ crosses $\overline{G}(\theta x)$ at most once and from above as x increases from 0 to ∞, for each $\theta > 0$. If $X \leq_* Y$, then Y is more skewed than X as explained in Marshall and Olkin [311]. The star order is also called more IFRA (increasing failure rate in average) order in reliability theory for reason explained below. The average failure of F at x is

$$\tilde{r}_X(x) = \frac{1}{x}\int_0^x r_X(u)\mathrm{d}u = \frac{-\ln \overline{F}(x)}{x}.$$

Thus $F \leq_* G$ can be interpreted in terms of average failure rates as

$$\frac{\tilde{r}_X(F^{-1}(u))}{\tilde{r}_Y(G^{-1}(u))} = \frac{G^{-1}(u)}{F^{-1}(u)}$$

being increasing in $u \in (0,1)$. A random variable X is said to have an IFRA distribution if its average failure rate $\tilde{r}_X(x)$ is increasing. Note that X has an IFRA distribution if and only if F is star ordered with respect to exponential distribution.

Definition 10.2.6. X is said to be more NBUE (new better than used in expectation) than Y or X is smaller than Y in the NBUE order (written as $X \leq_{\text{NBUE}} Y$) if

$$\frac{1}{\mu_F}\int_{F^{-1}(u)}^{\infty} \overline{F}(x)\mathrm{d}x \leq \frac{1}{\mu_G}\int_{G^{-1}(u)}^{\infty} \overline{G}(x)\mathrm{d}x, \quad \text{for all } u \in (0,1],$$

where μ_F (μ_G) denotes the expectation of X (Y).

It has been shown in Kochar [246] that

$$X \leq_* Y \Longrightarrow X \leq_{\text{NBUE}} Y \Longrightarrow X \leq_{\text{Lorenz}} Y,$$

where $\leq_{\text{Lorenz}}$ means the Lorenz order, a well-known order in economics. It is also known that (Marshall and Olkin [311], p. 69),

$$X \leq_{\text{Lorenz}} Y \Longrightarrow \gamma_X \leq \gamma_Y,$$

where $\gamma_X = \sqrt{\text{Var}[X]}/\text{E}[X]$ denotes the coefficient of variation of X. A good discussion of those orders can be found in Barlow and Proschan [39], Marshall and Olkin [311], and Shaked and Shanthikumar [427].

When $\text{E}[X] = \text{E}[Y]$, the order $\leq_{\text{RS}}$ is equivalent to the order $\leq_{\text{NBUE}}$. However, they are distinct when $\text{E}[X] \neq \text{E}[Y]$. For more details, please refer to Kochar et al. [250].

We will also use majorization in the following discussion. Let $\{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$ denote the increasing arrangement of the components of the vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

Definition 10.2.7. The vector $\boldsymbol{x}$ in $\mathbb{R}_+^n$ is said to majorize the vector $\boldsymbol{y}$ in $\mathbb{R}_+^n$ (denoted by $\boldsymbol{x} \overset{m}{\succeq} \boldsymbol{y}$) if

$$\sum_{i=1}^{j} x_{(i)} \leq \sum_{i=1}^{j} y_{(i)}$$

for $j = 1, \ldots, n-1$ and $\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$.

Relaxing the equality condition gives the following weak majorization order:

Definition 10.2.8. The vector $\boldsymbol{x}$ in $\mathbb{R}_+^n$ is said to weakly submajorize the vector $\boldsymbol{y}$ in $\mathbb{R}_+^n$ (denoted by $\boldsymbol{x} \overset{w}{\succeq} \boldsymbol{y}$) if, for $j = 1, \ldots, n$,

$$\sum_{i=1}^{j} x_{[i]} \geq \sum_{i=1}^{j} y_{[i]}$$

where $\{x_{[1]}, x_{[2]}, \ldots, x_{[n]}\}$ denotes the decreasing arrangement of the components of the vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$.

For extensive and comprehensive details on the theory of the majorization order and its applications, please refer to Marshall and Olkin [308].

Another interesting weaker order related to the majorization order introduced by Bon and Păltănea [75] is the p-larger order.

Definition 10.2.9. A vector $\boldsymbol{x}$ in $\mathbb{R}_+^n$ is said to be p-larger than another vector $\boldsymbol{y}$ in $\mathbb{R}_+^n$ (denoted by $\boldsymbol{x} \overset{p}{\succeq} \boldsymbol{y}$) if

$$\prod_{i=1}^{j} x_{(i)} \leq \prod_{i=1}^{j} y_{(i)}, \quad j = 1, \ldots, n.$$

Zhao and Balakrishnan [488] introduced the following order of reciprocal majorization.

Definition 10.2.10. A vector $\boldsymbol{x}$ in $\mathbb{R}_+^n$ is said to reciprocally majorize the another vector $\boldsymbol{y}$ in $\mathbb{R}_+^n$ (denoted by $\boldsymbol{x} \overset{rm}{\succeq} \boldsymbol{y}$) if

$$\sum_{i=1}^{j} \frac{1}{x_{(i)}} \geq \sum_{i=1}^{j} \frac{1}{y_{(i)}}, \quad j = 1, \ldots, n.$$

It has been pointed out in Kochar and Xu [259] that

$$\boldsymbol{x} \overset{m}{\succeq} \boldsymbol{y} \Longrightarrow \boldsymbol{x} \overset{p}{\succeq} \boldsymbol{y} \Longrightarrow \boldsymbol{x} \overset{rm}{\succeq} \boldsymbol{y}.$$

10.3 Magnitude and Dispersive Orderings

Let $X_1, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$, and $Y_1, \ldots, Y_n$ be another set of independent exponential random variables with Y_i having hazard rate λ_i', $i = 1, \ldots, n$. Boland, EI-Newehi and Proschan [71] showed that under the condition of the majorization order,

$$(\lambda_1, \ldots, \lambda_n) \overset{m}{\succeq} (\lambda_1', \ldots, \lambda_n') \Longrightarrow \sum_{i=1}^{n} X_i \geq_{\mathrm{lr}} \sum_{i=1}^{n} Y_i.$$

Under the same condition, Kochar and Ma [252] proved that

$$(\lambda_1, \ldots, \lambda_n) \overset{m}{\succeq} (\lambda'_1, \ldots, \lambda'_n) \Longrightarrow \sum_{i=1}^{n} X_i \geq_{\text{disp}} \sum_{i=1}^{n} Y_i. \tag{10.3.1}$$

This topic has been extensively investigated by Bon and Păltănea [75]. They pointed out that, under the p-larger order, which is a weaker order than the majorization order,

$$(\lambda_1, \ldots, \lambda_n) \overset{p}{\succeq} (\lambda'_1, \ldots, \lambda'_n) \Longrightarrow \sum_{i=1}^{n} X_i \geq_{\text{hr}} \sum_{i=1}^{n} Y_i.$$

This result has been strengthened by Khaledi [226] as

$$(\lambda_1, \ldots, \lambda_n) \overset{p}{\succeq} (\lambda'_1, \ldots, \lambda'_n) \Longrightarrow \sum_{i=1}^{n} X_i \geq_{\text{disp}} \sum_{i=1}^{n} Y_i. \tag{10.3.2}$$

More recently, Zhao and Balakrishnan [488] proved that, under the condition of reciprocal order,

$$(\lambda_1, \ldots, \lambda_n) \overset{rm}{\succeq} (\lambda'_1, \ldots, \lambda'_n) \Longrightarrow \sum_{i=1}^{n} X_i \geq_{\text{mrl}} \sum_{i=1}^{n} Y_i. \tag{10.3.3}$$

The result (10.3.1) of Kochar and Ma [252] can be immediately extended to convolutions of Erlang random variables as follows:

Theorem 10.3.1. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that for $i = 1, \ldots, n$, X_{λ_i} has gamma distribution with scale parameter λ_i and a common shape parameter a which is an integer such that $a \geq 1$. Then*

$$\boldsymbol{\lambda} \overset{m}{\succeq} \boldsymbol{\lambda}^* \Longrightarrow \sum_{i=1}^{n} X_{\lambda_i} \geq_{\text{disp}} \sum_{i=1}^{n} X_{\lambda_i^*}. \tag{10.3.4}$$

Korwar [266] generalized Theorem (10.3.1) to the case of $a \geq 1$. Khaledi and Kochar [233] strengthened this result with majorization replaced by p-larger ordering in (10.3.1).

Theorem 10.3.2. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_i, for $i = 1, \ldots, n$. Then*

$$\boldsymbol{\lambda} \overset{p}{\succeq} \boldsymbol{\lambda}^* \Longrightarrow \sum_{i=1}^{n} X_{\lambda_i} \geq_{\text{disp}} \sum_{i=1}^{n} X_{\lambda_i^*}. \tag{10.3.5}$$

The following result is an immediate consequence of Theorem 10.3.2, Theorem 10.2.4, and the fact that convolutions of IFR distributions are IFR.

Corollary 10.3.3. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_i, for $i = 1, \ldots, n$. Then,*

$$\boldsymbol{\lambda} \overset{p}{\succeq} \boldsymbol{\lambda}^* \Longrightarrow \sum_{i=1}^{n} X_{\lambda_i} \geq_{\mathrm{hr}} \sum_{i=1}^{n} X_{\lambda_i^*}. \tag{10.3.6}$$

Since $(\lambda_1, \lambda_2, \ldots, \lambda_n) \overset{p}{\succeq} (\tilde{\lambda}, \tilde{\lambda}, \ldots, \tilde{\lambda})$, where $\tilde{\lambda}$ is the geometric mean of the λ_i's, the following lower bounds on various quantities of interest associated with convolutions of gamma random variables are given next.

Corollary 10.3.4. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_i, for $i = 1, \ldots, n$. Then,*

(a) $\sum_{i=1}^{n} X_{\lambda_i} \geq_{\mathrm{disp}} \sum_{i=1}^{n} Y_i$.

(b) $\sum_{i=1}^{n} X_{\lambda_i} \geq_{\mathrm{hr}} \sum_{i=1}^{n} Y_i$ *which implies*

(c) $\sum_{i=1}^{n} X_{\lambda_i} \geq_{\mathrm{st}} \sum_{i=1}^{n} Y_i$,

where $Y_1, \ldots, Y_n$ is a random sample from a gamma distribution with shape parameter $a \geq 1$ and scale parameter $\tilde{\lambda}$, the geometric mean of λ_i's.

This result leads to better bounds for measures of variability for $\sum_{i=1}^{n} X_{\lambda_i}$ by replacing λ's by their geometric mean. On the other hand the bounds given by Korwar [266] uses arithmetic mean $\overline{\lambda} = \sum_{i=1}^{n} \lambda_i$ instead of the geometric mean on the right-hand sides of the above inequalities.

In Figs. 10.1 and 10.2 we plot the distribution functions of convolutions of two independent gamma random variables along with the bounds given by Corollary 10.3.4(c) and by Korwar [266].

In Figs. 10.3 and 10.4 we plot the hazard functions of convolutions of two independent gamma random variables along with the bounds given by Corollary 10.3.4(b) and by Korwar [266]. The vector of parameters in Figs. 10.1 and 10.3 is $\boldsymbol{\lambda}_1 = (1, 2)$ and that in Figs. 10.2

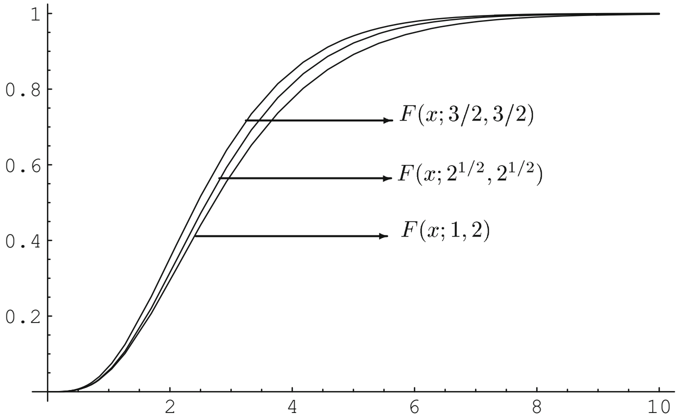


Figure 10.1: Graphs of distribution functions of $S(\lambda_1, \lambda_2)$

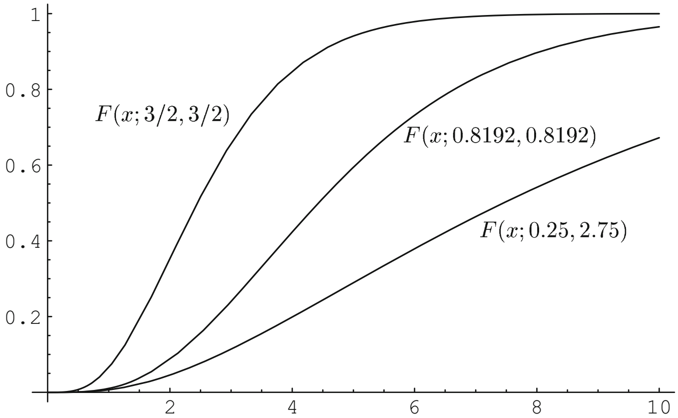


Figure 10.2: Graphs of distribution functions of $S(\lambda_1, \lambda_2)$

and 10.4 is $\boldsymbol{\lambda}_2 = (0.25, 2.75)$. Note that $\boldsymbol{\lambda}_2 \overset{m}{\succeq} \boldsymbol{\lambda}_1$. It appears from these figures that the improvements on the bounds are relatively more if λ_i's are more dispersed in the sense of majorization. The fact that this is true follows because the geometric mean is Schur concave, whereas the arithmetic mean is Schur constant and the distribution (hazard rate) of convolutions of i.i.d. gamma random variables with common parameter $\tilde{\lambda}$ is decreasing (increasing) in $\tilde{\lambda}$.

The following result due to Amiri, Khaledi and Samaniego [11] is a generalization of Theorem 4.1 of Zhao and Balakrishnan [488] and Corollary 3.8 in Kochar and Xu [259] from convolutions of independent

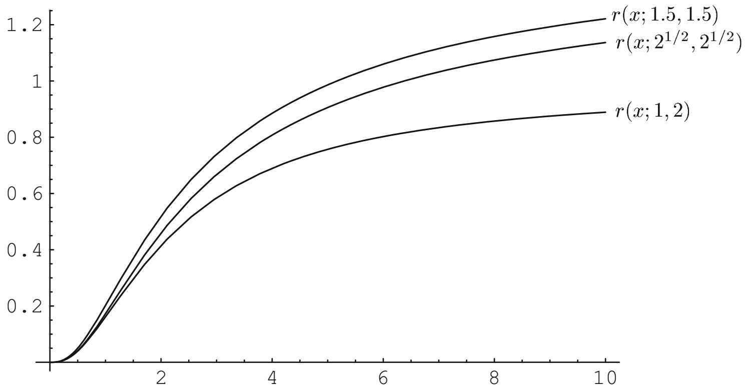


Figure 10.3: Graphs of hazard rates of $S(\lambda_1, \lambda_2)$

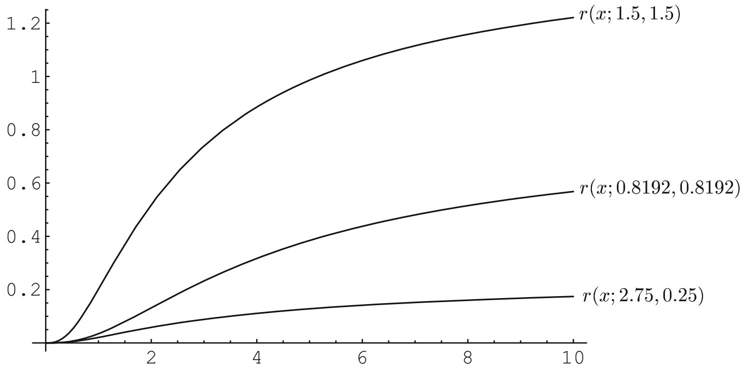


Figure 10.4: Graphs of hazard rates of $S(\lambda_1, \lambda_2)$

exponential distributions to convolutions of gamma distributions with common shape parameters $a \geq 1$.

Theorem 10.3.5. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_i, for $i = 1, \ldots, n$. Then,*

$$(\lambda_1, \ldots, \lambda_n) \overset{rm}{\succeq} (\lambda_1^*, \ldots, \lambda_n^*) \Longrightarrow \sum_{i=1}^{n} X_{\lambda_i} \geq_{\mathrm{mrl}} \sum_{i=1}^{n} X_{\lambda_i^*}.$$

Corollary 10.3.6. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_i, for $i = 1, \ldots, n$ and $Y_1, \ldots, Y_n$ be a random* *sample from a gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_H, where λ_H is harmonic mean of λ_i's. Then,*

$$(\lambda_1, \ldots, \lambda_n) \overset{rm}{\succeq} (\lambda_1^*, \ldots, \lambda_n^*) \Longrightarrow \sum_{i=1}^{n} X_{\lambda_i} \geq_{\text{mrl}} \sum_{i=1}^{n} Y_i.$$

This corollary provides a computable lower bound on mrl function of convolutions of gamma random variables which is sharper than those that can be obtained from Theorem 3.4 of Korwar [266] in terms of arithmetic mean and from Corollary 2.2 of Khaledi and Kochar [233] in terms of geometric mean of λ_i's. To justify these observations, in Figs. 10.5 and 10.6, we plot the mean residual life functions of convolutions of two independent gamma random variables with bound given in terms of arithmetic mean, geometric mean, and harmonic mean of λ_i's. In Fig. 10.5, we plot the mean residual functions for $\lambda_1 = 3.6$ and $\lambda_2 = 0.4$.

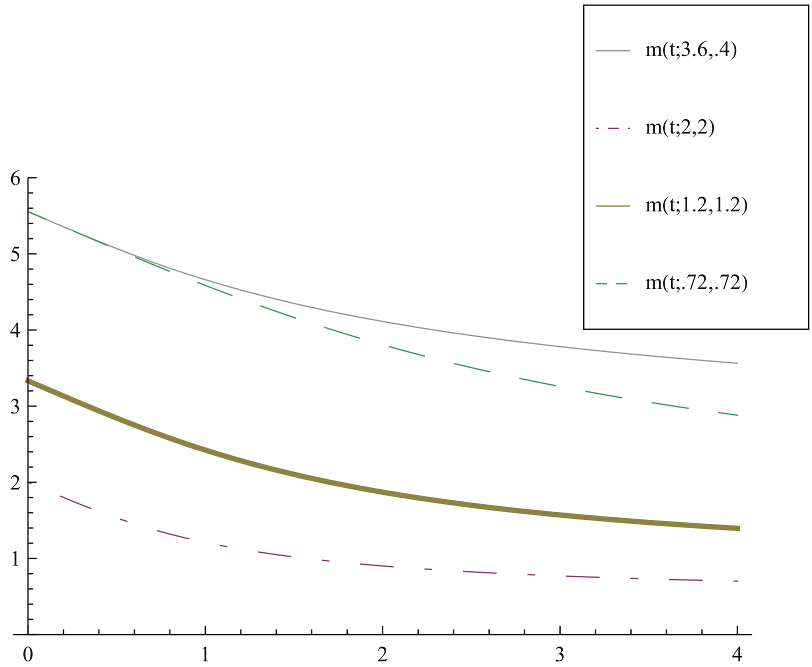


Figure 10.5: Mean residual function of convolutions of gamma random variables

We also plot the mean residual life functions of convolutions of independent gamma random variables for different sets of λ_i's

$$(2,6) \overset{rm}{\succeq} (5.2,2.4) \overset{rm}{\succeq} (3,6) \overset{rm}{\succeq} (4,4)$$

that show how rm ordering between λ_i's will affect the mean residual life function of convolutions of gamma random variables.

Mi, Shi and Zhou [320] studied linear combinations of independent gamma random variables with different integer shape parameters (i.e., Erlang random variables). They established the likelihood ratio ordering between two linear combinations of Erlang random variables under some restrictions on the coefficients and shape parameters. It is interesting to note that Yu [484] proved

$$\sum_{i=1}^{n} \beta_i X_i \geq_{\mathrm{st}} \sum_{i=1}^{n} \beta X_i \Longleftrightarrow \prod_{i=1}^{n} \beta_i^{a_i} \geq \prod_{i=1}^{n} \beta^{a_i}, \tag{10.3.7}$$

where $\beta, \beta_i \in \mathbb{R}_+$, and X_i's are gamma random variables $\Gamma(a_i, \lambda)$ for $i = 1, \ldots, n$, respectively.

Kochar and Xu [264] presented the following equivalent characterization of stochastic ordering between two linear combinations of independent gamma random variables.

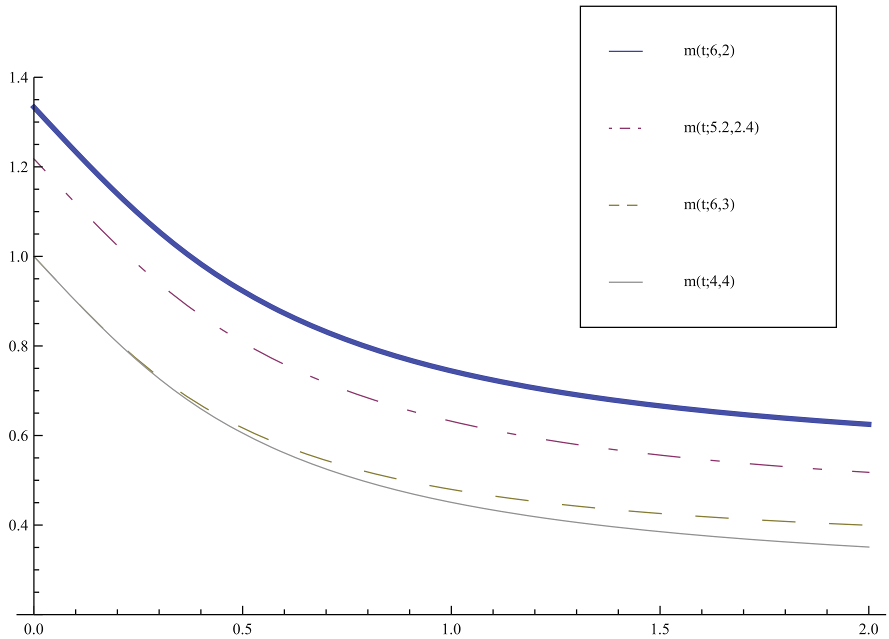


Figure 10.6: Mean residual function of convolutions of gamma random variables

Lemma 10.3.7. *Let X_1 and X_2 be independent gamma random variables $\Gamma(a_1, \lambda)$ and $\Gamma(a_2, \lambda)$, respectively. If $\beta_{(2)}/\beta_{(1)} \geq \beta'_{(2)}/\beta'_{(1)}$, then the following statements are equivalent:*

1. $\beta_{(1)}^{a_1}\beta_{(2)}^{a_2} \geq \left(\beta'_{(1)}\right)^{a_1}\left(\beta'_{(2)}\right)^{a_2}$
2. $\beta_{(1)}X_1 + \beta_{(2)}X_2 \geq_{\text{st}} \beta'_{(1)}X_1 + \beta'_{(2)}X_2$

They also proved the following result, which recovers Theorem 3.3 in Zhao [487]:

Theorem 10.3.8. *Let $X_1, \ldots, X_n$ be independent gamma random variables $\Gamma(a_1, \lambda), \ldots, \Gamma(a_n, \lambda)$, respectively. If $1 \leq a_1 \leq a_2 \leq \ldots \leq a_n$, then*

$$(\log(\beta_1), \ldots, \log(\beta_n)) \overset{w}{\succeq} \left(\log(\beta'_1), \ldots, \log(\beta'_n)\right) \Longrightarrow \sum_{i=1}^{n} \beta_{(i)}X_i \geq_{\text{disp}} \sum_{i=1}^{n} \beta'_{(i)}X_i.$$

10.4 Star Ordering

Kochar and Xu [259] proved the following result on convolutions of exponential random variables:

$$\left(\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}\right) \overset{m}{\succeq} \left(\frac{1}{\lambda_1^*}, \ldots, \frac{1}{\lambda_n^*}\right) \Longrightarrow \sum_{i=1}^{n} E_{\lambda_i} \geq_{\triangle} \sum_{i=1}^{n} E_{\lambda_i^*}, \qquad (10.4.1)$$

where E_{λ_i}, $i = 1, \ldots, n$ is exponential random variable with hazard rate λ_i and $\triangle$ order stands for NBUE and Lorenz order. For more details of Lorenz order the reader is referred to Sect. 3.A. in Shaked and Shanthikumar [426].

Let X_{θ_1}, X_{θ_2}, $X_{\theta'_1}$, and $X_{\theta'_2}$ be independent gamma random variables with a common shape parameter a and scale parameters $\theta_1 = 1/\lambda_1$, $\theta_2 = 1/\lambda_2$, $\theta'_1 = 1/\lambda'_1$, and $\theta'_2 = 1/\lambda'_2$, respectively. Proposition 2.1 of Diaconis and Perlman [123] shows that if

$$(\lambda_1, \lambda_2) \overset{m}{\succeq} (\lambda'_1, \lambda'_2),$$

then the distribution function of $X_{\theta_1} + X_{\theta_2}$ crosses that of $X_{\theta'_1} + X_{\theta'_2}$ exactly once.

In this section, it will be shown that under various conditions on the scale parameters, one can establish star ordering between convolutions of gamma random variables .

Recently, Kochar and Xu [263] studied the problem of comparing the skewness of linear combinations of independent gamma random variables. Let X_1 and X_2 be independent and identically distributed gamma random variables $\Gamma(a, \lambda)$. They proved that for $(\beta_i, \beta_i') \in \mathbb{R}_+^2$, $i = 1, 2$, if either

$$(\beta_1, \beta_2) \overset{m}{\succeq} (\beta_1', \beta_2') \tag{10.4.2}$$

or

$$\left(\frac{1}{\beta_1}, \frac{1}{\beta_2}\right) \overset{m}{\succeq} \left(\frac{1}{\beta_1'}, \frac{1}{\beta_2'}\right) \tag{10.4.3}$$

then

$$\beta_1 X_1 + \beta_2 X_2 \geq_* \beta_1' X_1 + \beta_2' X_2, \tag{10.4.4}$$

where $\geq_*$ denotes the star order and $\overset{m}{\succeq}$ denotes the majorization order. Amiri et al. [11] also independently proved the above results when $a \geq 1$. These results are closely related to a result of Yu [484], who proved that, for $\beta_i \in \mathbb{R}_+$,

$$\sum_{i=1}^{n} \beta_i X_i \geq_* \sum_{i=1}^{n} X_i, \tag{10.4.5}$$

where X_i's are gamma random variables $\Gamma(a_i, \lambda)$ for $i = 1, \dots, n$, respectively. These results reveal that if the coefficients are more dispersed, then the linear combinations are more skewed as compared by star ordering.

This topic is further pursued by Zhao [487] who extended the results of Eqs. (10.4.2)–(10.4.4) to the case of independent gamma random variables with different shape parameters. More precisely, let X_1 and X_2 be independent gamma random variables $\Gamma(a_1, \lambda)$ and $\Gamma(a_2, \lambda)$. Zhao proved that for $\beta_1 \leq \beta_2$ and $\beta_1' \leq \beta_2'$,

$$(\beta_1, \beta_2) \overset{m}{\succeq} (\beta_1', \beta_2') \Longrightarrow \beta_1 X_1 + \beta_2 X_2 \geq_* \beta_1' X_1 + \beta_2' X_2, \tag{10.4.6}$$

and if $\beta_1 \leq \beta_2$, $\beta_1' \leq \beta_2'$, and $a_1 \leq a_2$, then

$$\left(\frac{1}{\beta_1}, \frac{1}{\beta_2}\right) \overset{m}{\succeq} \left(\frac{1}{\beta_1'}, \frac{1}{\beta_2'}\right) \Longrightarrow \beta_1 X_1 + \beta_2 X_2 \geq_* \beta_1' X_1 + \beta_2' X_2. \tag{10.4.7}$$

Kochar and Xu [262] give a different sufficient condition on the scale parameters of the convoluting gamma random variables for star ordering to hold.

Theorem 10.4.1. *Let X_{θ_1}, X_{θ_2}, $X_{\theta_1'}$, $X_{\theta_2'}$ be independent gamma random variables with a common shape parameter a and scale parameters $\theta_1 = 1/\lambda_1$, $\theta_2 = 1/\lambda_2$, $\theta_1' = 1/\lambda_1'$, and $\theta_2' = 1/\lambda_2'$. Then,*

$$(\lambda_1, \lambda_2) \overset{m}{\succeq} (\lambda_1', \lambda_2') \Longrightarrow X_{\theta_1} + X_{\theta_2} \geq_* X_{\theta_1'} + X_{\theta_2'}.$$

Remark 10.4.2. Theorem 10.4.1 implies that the distribution function of $X_{\theta_1}+X_{\theta_2}$ crosses that of $X_{\theta_1'}+X_{\theta_2'}$ at most once, no matter how $X_{\theta_1} + X_{\theta_2}$ is scaled. As a special case, they have exactly one crossing when both sides have the same mean which strengthens Proposition 2.1 in Diaconis and Perlman [123].

Recently Kochar and Xu [264] have given a new sufficient condition for ordering the skewness of linear combinations of two independent gamma random variables with arbitrary shape parameters and this result unifies the previous results on this topic as given above.

Theorem 10.4.3. *Let X_1 and X_2 be independent gamma random variables $\Gamma(a_1, \lambda)$ and $\Gamma(a_2, \lambda)$, respectively. Then,*

$$\frac{\beta_{(2)}}{\beta_{(1)}} \geq \frac{\beta_{(2)}'}{\beta_{(1)}'} \Longrightarrow \beta_{(1)}X_1 + \beta_{(2)}X_2 \geq_* \beta_{(1)}'X_1 + \beta_{(2)}'X_2,$$

where $\{\beta_{(1)}, \beta_{(2)}\}$ denotes the increasing arrangement of the components of the vector $(\beta_1, \beta_2) \in \mathbb{R}_+^2$.

Remark 10.4.4. The condition given in the Theorem 10.4.3 is very general. It is weaker than any of the following conditions, which are commonly used in the literature:

1. $(\beta_1, \beta_2) \overset{m}{\succeq} (\beta_1', \beta_2')$
2. $(\log(\beta_1), \log(\beta_2)) \overset{m}{\succeq} (\log(\beta_1'), \log(\beta_2'))$
3. $(1/\beta_1, 1/\beta_2) \overset{m}{\succeq} (1/\beta_1', 1/\beta_2')$

Remark 10.4.5. Conditions (a) and (c) have been used to prove Theorems 4.2 and 4.3 in Zhao [487] [see also Eqs. (10.4.6) and (10.4.7)]. The proof of Theorem 4.2 of Zhao [487] is quite involved. However, it follows immediately from Remark 1.

10.5 Right Spread Order of Linear Combinations

Amiri et al. [11] proved the following result on RS ordering between convolutions of gamma random variables with a common shape parameter.

Theorem 10.5.1. *Let $X_{\lambda_1}, \ldots, X_{\lambda_n}$ be independent random variables such that X_{λ_i} has gamma distribution with shape parameter $a \geq 1$ and scale parameter λ_i, for $i = 1, \ldots, n$. Then,*

$$(\lambda_1, \ldots, \lambda_n) \overset{rm}{\succeq} (\lambda_1^*, \ldots, \lambda_n^*) \Longrightarrow \sum_{i=1}^{n} X_{\lambda_i} \geq_{\mathrm{RS}} \sum_{i=1}^{n} X_{\lambda_i^*}.$$

Theorem 10.5.1 generalizes Corollary 3.9 of Kochar and Xu [259] from convolutions of independent Erlang distributions to convolutions of gamma distributions with common shape parameters $a \geq 1$.

Now we consider the case when the shape parameters of the gamma random variables are not necessarily equal. The first result (cf. Kochar and Xu [264]) gives the following characterization of right spread order for linear combinations of two gamma random variables:

Lemma 10.5.2. *Let X_1 and X_2 be independent gamma random variables $\Gamma(a_1, \lambda)$ and $\Gamma(a_2, \lambda)$, respectively. If $\beta_{(2)}/\beta_{(1)} \geq \beta'_{(2)}/\beta'_{(1)}$, then the following statements are equivalent:*

1. $\beta_{(1)} a_1 + \beta_{(2)} a_2 \geq \beta'_{(1)} a_1 + \beta'_{(2)} a_2$
2. $\beta_{(1)} X_1 + \beta_{(2)} X_2 \geq_{\mathrm{RS}} \beta'_{(1)} X_1 + \beta'_{(2)} X_2$

Proof: It follows from Theorem 4.3 in Fernández-Ponce et al. [160] that for two nonnegative random variables X and Y, if $X \leq_* Y$, then

$$\mathrm{E}[X] \leq \mathrm{E}[Y] \Longleftrightarrow X \leq_{\mathrm{RS}} Y.$$

It follows from Theorem 10.4.3 that under the given assumption,

$$\beta_{(1)} X_1 + \beta_{(2)} X_2 \geq_* \beta'_{(1)} X_1 + \beta'_{(2)} X_2.$$

Hence,

$$\beta_{(1)} X_1 + \beta_{(2)} X_2 \geq_{\mathrm{RS}} \beta'_{(1)} X_1 + \beta'_{(2)} X_2$$

is equivalent to

$$\mathrm{E}\left[\beta_{(1)} X_1 + \beta_{(2)} X_2\right] \left[\geq \mathrm{E}\left[\beta'_{(1)} X_1 + \beta'_{(2)} X_2\right] .\right.$$

So, the desired result follows. ■

Remark 10.5.3. Theorem 4.5 in Zhao [487] states that if $1 \le a_1 \le a_2$, then

$$(\beta_1, \beta_2) \overset{w}{\succeq} (\beta_1', \beta_2') \Longrightarrow \beta_{(1)} X_1 + \beta_{(2)} X_2 \ge_{\mathrm{RS}} \beta_{(1)}' X_1 + \beta_{(2)}' X_2.$$

Lemma 10.5.2 removes the restriction on the shape parameters.

As a direct consequence, we have the following result:

Corollary 10.5.4. *Let X_1 and X_2 be independent gamma random variables $\Gamma(a_1, \lambda)$ and $\Gamma(a_2, \lambda)$, respectively. Then,*

$$(\beta_1, \beta_2) \overset{m}{\succeq} (\beta_1', \beta_2') \Longrightarrow \beta_{(1)} X_1 + \beta_{(2)} X_2 \ge_{\mathrm{RS}} \beta_{(1)}' X_1 + \beta_{(2)}' X_2.$$

The following result of Zhao [487] immediately follows from Corollary 10.5.4, Theorem 3.C.7 of Shaked and Shanthikumar [427], and similar argument to Theorem 10.3.8.

Corollary 10.5.5. *Let mutually independent random variables $X_1, \ldots, X_n$ have gamma distributions $\Gamma(a_1, \lambda), \ldots, \Gamma(a_n, \lambda)$, respectively. If $1 \le a_1 \le a_2 \le \cdots \le a_n$, then*

$$(\beta_1, \ldots, \beta_n) \overset{w}{\succeq} (\beta_1', \ldots, \beta_n') \Longrightarrow \sum_{i=1}^{n} \beta_{(i)} X_i \ge_{\mathrm{RS}} \sum_{i=1}^{n} \beta_{(i)}' X_i.$$

Yu [484] gave necessary and sufficient conditions for stochastically comparing linear combinations of heterogeneous and homogeneous gamma random variables. The following result gives necessary and sufficient conditions for comparing linear combinations of gamma random variables according to right spread order:

Proposition 10.5.6. *Let mutually independent random variables $X_1, \ldots, X_n$ have gamma distributions $\Gamma(a_1, \lambda), \ldots, \Gamma(a_n, \lambda)$, respectively. Then,*

$$\sum_{i=1}^{n} \beta_i X_i \ge_{\mathrm{RS}} \beta \sum_{i=1}^{n} X_i \Longleftrightarrow \beta \le \frac{\sum_{i=1}^{n} \beta_i a_i}{\sum_{i=1}^{n} a_i}.$$

Proof: It follows from Yu [484] [see also (10.4.5)] that

$$\sum_{i=1}^{n} \beta_i X_i \ge_{*} \beta \sum_{i=1}^{n} X_i.$$

Using Theorem 4.3 in Fernández-Ponce et al. [160] again, we have

$$\sum_{i=1}^{n} \beta_i X_i \geq_{\mathrm{RS}} \beta \sum_{i=1}^{n} X_i \Longleftrightarrow \mathrm{E}\left(\sum_{i=1}^{n} \beta_i X_i\right) \leq \mathrm{E}\left[\sum_{i=1}^{n} \beta X_i\right].$$

Hence, the required result follows. ■

Compared to Corollary 10.5.5, Proposition 10.5.6 poses no restriction on the shape parameters.

Acknowledgments

Baha-Eldin Khaledi and Subhash Kochar are partially supported by Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad.

Chapter 11

On Used Systems and Systems with Used Components

Xiaohu Li, Franco Pellerey, and Yinping You

Abstract: Consider an n–component coherent system having random lifetime $T_{\boldsymbol{X}}$, where $\boldsymbol{X} = (X_1, \ldots, X_n)$ is the vector of the non-independent components' lifetimes. Stochastic comparisons of the residual life of $T_{\boldsymbol{X}}$ at a fixed time $t \geq 0$, conditioned on $\{T_{\boldsymbol{X}} > t\}$ or on $\{X_i > t, \forall i = 1, \ldots, n\}$, are investigated. Sufficient conditions on the vector $\boldsymbol{X}$ that imply this comparison in the usual stochastic order are provided, together with sufficient conditions under which the lifetime $T_{\boldsymbol{X}}$ satisfies the NBU aging property.

X. Li • Y. You
School of Mathematical Sciences, Xiamen University, Xiamen 361005, China
e-mail: mathxhli@hotmail.com; youyinping19881203@163.com

F. Pellerey (✉)
Dipartimento di Scienze Matematiche, Politecnico di Torino,
C.so Duca degli Abruzzi, 24, I-10129 Torino, Italy
e-mail: franco.pellerey@polito.it

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_11,

11.1 Introduction

Coherent systems are often considered in reliability theory to describe the structure and the performance of complex systems. Consider an item formed by a number n of components, i.e., an *n-component system*. Its *structure function* $\phi : \{0,1\}^n \to \{0,1\}$ is a function that maps the state vector $\widehat{\boldsymbol{x}} = (\widehat{x}_1, \ldots, \widehat{x}_n)$ of its components (where $\widehat{x}_i = 1$ if component i is working and $\widehat{x}_i = 0$ if it is failed) to the state $\widehat{y} \in \{0,1\}$ of the system itself. The system is said to be *coherent* whenever every component is relevant (i.e., it affects the working or failure of the system) and the structure function is monotone in every component (i.e., replacing a failed component by a working component cannot cause a working system to fail). For example, k-out-of-n systems, and series and parallel systems in particular, are coherent systems. See Esary and Marshall [152] or Barlow and Proschan [39] for a detailed introduction to coherent systems and related properties and applications.

Several problems and results dealing with aging properties for lifetimes of coherent systems, or with stochastic comparisons among coherent systems, have been considered in reliability literature. In particular, the closure property of some aging notions with respect to construction of coherent systems has been investigated, in most of the cases assuming independence among the lifetimes of the system's components (see, e.g, Barlow and Proschan [39], Samaniego [407], Deshpande et al. [121], Franco et al. [169], Li and Chen [286]).

Among others, a natural question dealing with coherent systems is on the comparison between the reliability of a used coherent system and the reliability of a systems with used components. Precisely, denoted with $\boldsymbol{X}$ the vector of the component's lifetimes and with $T_{\boldsymbol{X}}$ the lifetime of the system, one can consider stochastic comparisons between the residual lifetimes

$$[T_{\boldsymbol{X}} - t \mid T_{\boldsymbol{X}} > t]$$

and

$$[T_{\boldsymbol{X}} - t \mid X_i > t, \forall i = 1, \ldots, n],$$

for $t \geq 0$. In fact, it is commonly assumed that the former is smaller, in some stochastic sense, than the latter. The intuitive explanation of this fact is that the reliability of a system with all components being in working state is higher with respect to the case with some of them being in failure state, even if the system is not in failure state. This assertion, which is actually true under assumption of independence among components (see, e.g., Pellerey and Petakos [371], or Li and Lu [291]), is not always verified for non-independent components, as shown, for example, in Sect. 11.2.

This problem, and similar problems, has been recently investigated, for example, in Khaledi and Shaked [235], Navarro et al. [339], or Samaniego et al. [409], under the assumption of independence among components' lifetimes, or in Zhang [486], under assumption of exchangeability of components' lifetimes. The purpose of this paper is to generalize some of the results appearing in the above-mentioned references, in particular providing conditions on the vector $\boldsymbol{X}$ such that

$$[T_{\boldsymbol{X}} - t \mid T_{\boldsymbol{X}} > t] \leq_{\text{st}} [T_{\boldsymbol{X}} - t \mid X_i > t, \, \forall i = 1, \ldots, n], \tag{11.1.1}$$

even under the case of components having non-independent or exchangeable lifetimes, where $\leq_{\text{st}}$ denotes the usual stochastic order (whose definition is recalled below). These conditions are described in Sect. 11.2. As a corollary of the main result, a few statements that describe conditions on $\boldsymbol{X}$ such that the system's lifetime $T_{\boldsymbol{X}}$ satisfies some of the most well-known aging properties are presented in Sect. 11.3.

For ease of reference, some notations are introduced, and the definitions of several stochastic orders and dependence concepts which will be used in sequel are recalled.

Throughout this note, the terms *increasing* and *decreasing* stand for nondecreasing and nonincreasing, respectively. A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be *increasing* when $\phi(\boldsymbol{x}) \leq \phi(\boldsymbol{y})$ for $\boldsymbol{x} \leq \boldsymbol{y}$, which denotes $x_i \leq y_i$ for all $i = 1, \ldots, n$. All random variables under investigation are nonnegative, and expectations are implicitly assumed to be finite once they appear. The notation $[X \mid A]$ stands for the random object whose distribution is the conditional distribution of X given the event A. The dimension of a random vector is clear from the context and unless otherwise stated it is assumed to be n. We will denote with $I = \{1, \ldots, n\}$ the set of component's indices and with $I_i = \{1, \ldots, i\}$, for $i = 1, \ldots, n$, their subsets. For any nonempty $A \subset I$, $\boldsymbol{X}_A$ and $\boldsymbol{x}_A$ denote the random vector of those X_i's with $i \in A$ and the corresponding constant vector, respectively. Besides, for any $s \geq 0$, notation $\boldsymbol{s}$ denotes the constant vector $(s, \ldots, s)$ with the dimension conforming to its circumstance. Finally, the following notation is adopted:

$\boldsymbol{x} \wedge \boldsymbol{y} = (x_1 \wedge y_1, \ldots, x_n \wedge y_n)$, $\boldsymbol{x} \vee \boldsymbol{y} = (x_1 \vee y_1, \ldots, x_n \vee y_n)$, and $u \wedge v = \min\{u, v\}$, $u \vee v = \max\{u, v\}$.

Some well-known stochastic orders are recalled in the following definition. Further details, properties, and applications of these orders may be found in Shaked and Shanthikumar [426].

Definition 11.1.1. Given two random vectors (or variables) $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{Y}$ in the following:

(i) *Likelihood ratio order* (denoted by $\boldsymbol{X} \leq_{\mathrm{lr}} \boldsymbol{Y}$) if their joint densities f and g satisfy $f(\boldsymbol{x})g(\boldsymbol{y}) \leq f(\boldsymbol{x} \wedge \boldsymbol{y})g(\boldsymbol{x} \vee \boldsymbol{y})$ for any $\boldsymbol{x}$ and $\boldsymbol{y}$

(ii) *Stochastic order* (denoted by $\boldsymbol{X} \leq_{\mathrm{st}} \boldsymbol{Y}$) if $\mathrm{E}[\phi(\boldsymbol{X})] \leq \mathrm{E}[\phi(\boldsymbol{Y})]$ for any increasing function ϕ with finite expectations;

(iii) *Increasing convex order* (denoted by $\boldsymbol{X} \leq_{\mathrm{icx}} \boldsymbol{Y}$) if $\mathrm{E}[\phi(\boldsymbol{X})] \leq \mathrm{E}[\phi(\boldsymbol{Y})]$ for any increasing and convex function ϕ with finite expectations

(iv) *Increasing concave order* (denoted by $\boldsymbol{X} \leq_{\mathrm{icv}} \boldsymbol{Y}$) if $\mathrm{E}[\phi(\boldsymbol{X})] \leq \mathrm{E}[\phi(\boldsymbol{Y})]$ for any increasing and concave function ϕ with finite expectation

(v) *Upper orthant order* (denoted by $\boldsymbol{X} \leq_{\mathrm{uo}} \boldsymbol{Y}$) if $\mathrm{E}[\prod_{i=1}^n \phi_i(X_i)] \leq \mathrm{E}[\prod_{i=1}^n \phi_i(Y_i)]$ for any set of nonnegative increasing functions $\phi_i, i = 1 \ldots, n$ such that expectations exist

Recall that, in the univariate case, $X \leq_{\mathrm{st}} Y$ if, and only if, $\mathrm{P}\{X > t\} \leq \mathrm{P}\{Y > t\}$ for all $t \in \mathbb{R}$. The following two positive dependence notions also are well known (see, e.g., Joe [211], or Shaked and Shanthikumar [426]):

Definition 11.1.2. A random vector $\boldsymbol{X}$ is said to be *multivariate total positive of order 2* (MTP2) if its joint density f satisfies $f(\boldsymbol{x})f(\boldsymbol{y}) \leq f(\boldsymbol{x} \vee \boldsymbol{y})f(\boldsymbol{x} \wedge \boldsymbol{y})$ for any $\boldsymbol{x}$, $\boldsymbol{y}$.

Definition 11.1.3. For a bivariate vector $\boldsymbol{X} = (X_1, X_2)$, X_2 is said to be *right tail increasing* (RTI) in X_1 if $[X_2 \mid X_1 > x_1]$ is stochastically increasing in x_1 (and similarly X_1 is said to be RTI in X_2 if $[X_1 \mid X_2 > x_2]$ is stochastically increasing in x_2).

It should be mentioned that MTP2 property implies RTI property in both directions, while the reverse may not be true (see, e.g., Joe [211], or Müller and Scarsini [333], and references therein).

Finally, we recall that for a coherent system having structure function ϕ, the relationship between the vector $\boldsymbol{X}$ of component's lifetimes and system's lifetime $T_{\boldsymbol{X}}$ is described by the relation $T_{\boldsymbol{X}} = \tau(\boldsymbol{X})$, where the *coherent life function* $\tau : \mathbb{R}^n \to \mathbb{R}$ is defined as

$$\tau(x_1, \dots, x_n) = \sup\{t \geq 0 : \phi(\widehat{x}_{1,t}, \dots, \widehat{x}_{n,t}) = 1\},$$

where $\widehat{x}_{i,t} = 1$ if $x_i > t$, or $\widehat{x}_{i,t} = 0$ if $x_i \leq t$, for $i \in I$. It should recall that coherent life functions are increasing and such that

$$\tau(t_1 - s, \dots, t_n - s) = \tau(t_1, \dots, t_n) - s, \tag{11.1.2}$$

for every $s \geq 0$ and $t_i \geq s$, $i \in I$ (see Esary and Marshall [152]). Also, a subset $J = \{i_1, \dots, i_J\} \subseteq \{1, \dots, n\}$ of the components indices is said to be a *path set* if the system is working whenever the components indexed in J are working.

11.2 Main Results

First, we show that stochastic inequality (11.1.1) does not necessarily hold. In fact, let $\boldsymbol{X} = (X_1, X_2)$ be such that

$$\begin{aligned} \mathrm{P}\{(X_1, X_2) = (2, 1)\} &= 1/4 \\ \mathrm{P}\{(X_1, X_2) = (2, 2)\} &= 3/8 \\ \mathrm{P}\{(X_1, X_2) = (3, 1)\} &= 1/4 \\ \mathrm{P}\{(X_1, X_2) = (3, 2)\} &= 1/8 \end{aligned}$$

and let $T_{\boldsymbol{X}} = \max\{X_1, X_2\}$. Letting $t = 1.5$ and $s = 1$, it holds that

$$\mathrm{P}\{T_{\boldsymbol{X}} - t > s | T_{\boldsymbol{X}} > t\} = \frac{\mathrm{P}\{\max\{X_1, X_2\} > 2.5\}}{\mathrm{P}\{\max\{X_1, X_2\} > 1.5\}} = 3/8,$$

while

$$\mathrm{P}\{T_{\boldsymbol{X}} - t > s | X_i > t,\ \forall i\} = \frac{\mathrm{P}\{\max\{X_1, X_2\} > 2.5, X_1 > 1.5, X_2 > 1.5\}}{\mathrm{P}\{X_1 > 1.5, X_2 > 1.5\}} = 1/4,$$

so that (11.1.1) cannot be satisfied.

The following statement provides the first sufficient condition under which the stochastic comparison between $[T_{\boldsymbol{X}} - t \mid T_{\boldsymbol{X}} > t]$ and $[T_{\boldsymbol{X}} - t \mid X_i > t, \forall i = 1, \dots, n]$ does hold.

Theorem 11.2.1. *Let $\boldsymbol{X}$ be a vector of component's lifetimes such that, for any nonempty $A \subset I$ and $\boldsymbol{s} = (s, \ldots, s)$ with $s \geq 0$,*

$$[\boldsymbol{X}_{\bar{A}} - \boldsymbol{s} \mid \boldsymbol{X} > \boldsymbol{s}] \geq_{\text{st}} [\boldsymbol{X}_{\bar{A}} - \boldsymbol{s} \mid \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}]. \tag{11.2.1}$$

Then, (11.1.1) holds for any coherent system with lifetime $T_{\boldsymbol{X}} = \tau(\boldsymbol{X})$, i.e.,

$$[T_{\boldsymbol{X}} - s \mid T_{\boldsymbol{X}} > s] \leq_{\text{st}} [T_{\boldsymbol{X}} - s \mid \boldsymbol{X} > \boldsymbol{s}], \quad s \geq 0.$$

Proof: Denote with $J_1, J_2, \ldots, J_\ell = I$ all possible path sets of the coherent system which has lifetime $T_{\boldsymbol{X}}$. Then it holds that, for any $s \geq 0$,

$$\{T_{\boldsymbol{X}} > s\} = \bigcup_{i=1}^{\ell} \left\{\boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \leq \boldsymbol{s}\right\}.$$

For any s, $t \geq 0$, let

$$\begin{aligned} a_i &= \mathrm{P}\{\boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \leq \boldsymbol{s}\}, \quad i = 1, \ldots, \ell, \\ b_i &= \mathrm{P}\{T_{\boldsymbol{X}} > s + t,\, \boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \leq \boldsymbol{s}\}, \quad i = 1, \ldots, \ell. \end{aligned}$$

We have

$$\begin{aligned} & \mathrm{P}\{T_{\boldsymbol{X}} > s + t \mid T_{\boldsymbol{X}} > s\} \\ = \; & \frac{\mathrm{P}\left\{T_{\boldsymbol{X}} > s + t,\; T_{\boldsymbol{X}} > s\right\}}{\mathrm{P}\{T_{\boldsymbol{X}} > s\}} \\ = \; & \frac{\sum_{i=1}^{\ell} \mathrm{P}\left\{T_{\boldsymbol{X}} > s + t,\, \boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \leq \boldsymbol{s}\right\}}{\sum_{i=1}^{\ell} \mathrm{P}\left\{\boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \leq \boldsymbol{s}\right\}} \\ = \; & \frac{\sum_{i=1}^{\ell} b_i}{\sum_{i=1}^{\ell} a_i}. \end{aligned}$$

Now, for any path set J_i, denoted with n_i its cardinality, consider the system corresponding to the structure function $\phi_{J_i} : \{0, 1\}^{n_i} \to \{0, 1\}$ defined as $\phi_{J_i}(\widehat{\boldsymbol{x}}_{J_i}) = \phi(\widehat{\boldsymbol{x}}_{J_i}, 0_{\bar{J}_i})$, i.e., letting in failed state all the components outside the path set. Let $T^i_{\boldsymbol{X}_{J_i}} = \tau_i(\boldsymbol{X}_{J_i})$ denote the lifetime of the subsystem whose structure function is ϕ_{J_i}. Clearly, for any $\widehat{\boldsymbol{x}} \in \{0, 1\}^n$ we have $\phi_{J_i}(\widehat{\boldsymbol{x}}_{J_i}) = \phi(\widehat{\boldsymbol{x}}_{J_i}, \mathbf{0}_{\bar{J}_i}) \leq \phi(\widehat{\boldsymbol{x}}_{J_i}, \widehat{\boldsymbol{x}}_{\bar{J}_i}) = \phi(\widehat{\boldsymbol{x}})$, so that $\{T^i_{\boldsymbol{X}_{J_i}} > t\} \subseteq \{T_{\boldsymbol{X}} > t\}$. Moreover, since coherent life functions are increasing, by (11.1.2) and (11.2.1) it holds that

$$\begin{aligned}
\frac{b_i}{a_i} &= \mathrm{P}\{T_{\boldsymbol{X}} > s+t \mid \boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \le \boldsymbol{s}\} \\
&= \mathrm{P}\{\tau(\boldsymbol{X}) > s+t \mid \boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \le \boldsymbol{s}\} \\
&= \mathrm{P}\{\tau(\boldsymbol{X}-\boldsymbol{s}) > t \mid \boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \le \boldsymbol{s}\} \\
&= \mathrm{P}\{\tau_i(\boldsymbol{X}_{J_i}-\boldsymbol{s}) > t \mid \boldsymbol{X}_{J_i} > \boldsymbol{s},\, \boldsymbol{X}_{\bar{J}_i} \le \boldsymbol{s}\} \\
&\le \mathrm{P}\{\tau_i(\boldsymbol{X}_{J_i}-\boldsymbol{s}) > t \mid \boldsymbol{X}_{J_\ell} > \boldsymbol{s}\} \\
&\le \mathrm{P}\{\tau(\boldsymbol{X}-\boldsymbol{s}) > t \mid \boldsymbol{X}_{J_\ell} > \boldsymbol{s}\} \\
&= \mathrm{P}\{\tau(\boldsymbol{X}) > s+t \mid \boldsymbol{X}_{J_\ell} > \boldsymbol{s}\} \\
&= \mathrm{P}\{T_{\boldsymbol{X}} > s+t \mid \boldsymbol{X}_{J_\ell} > \boldsymbol{s}\} \\
&= \frac{b_\ell}{a_\ell}, \qquad \text{for any } i = 1, \ldots, \ell.
\end{aligned}$$

Thus, $b_i a_\ell \le a_i b_\ell$ for $i = 1, \ldots, \ell$. This invokes

$$a_\ell b_1 + \cdots + a_\ell b_\ell \le a_1 b_\ell + \cdots + a_\ell b_\ell$$

and hence

$$\frac{\sum_{i=1}^{\ell} b_i}{\sum_{i=1}^{\ell} a_i} \le \frac{b_\ell}{a_\ell},$$

which is just

$$\mathrm{P}\{T_{\boldsymbol{X}} - s > t \mid T_{\boldsymbol{X}} > s\} \le \mathrm{P}\{T_{\boldsymbol{X}} - s > t \mid \boldsymbol{X} > \boldsymbol{s}\},$$

i.e., the assertion. ■

Theorem 11.2.1 has a very nice physical implication and describes conditions under which a coherent system of used components is better than a used coherent system, in the sense of having stochastically larger life length. This essentially claims that the positive dependence, or the independence, among the components of the coherent system is a sufficient condition for this property. Herewith, we address some other sufficient conditions for the assumption (11.2.1) to hold.

Theorem 11.2.2. *If the joint density of $\boldsymbol{X} = (X_1, \ldots, X_n)$ is MTP2, then (11.2.1) holds for any nonempty $A \subseteq I$ and $s \ge 0$.*

Proof: Recall that the MTP2 property of $(X_1, \ldots, X_n)$ is equivalent to $\boldsymbol{X} \le_{\mathrm{lr}} \boldsymbol{X}$. Taking A and B as $\{\boldsymbol{X}_{\bar{A}} > \boldsymbol{s}, \boldsymbol{X}_A \le \boldsymbol{s}\}$ and $\{X_i > s,\, i = 1, \ldots, n\}$, respectively, in Theorem 6.E.2 of Shaked and Shanthikumar [426], we immediately obtain

$$[\boldsymbol{X} \mid \boldsymbol{X} > \boldsymbol{s}] \ge_{\mathrm{lr}} [\boldsymbol{X} \mid \boldsymbol{X}_A \le \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}].$$

Now, by Theorem 6.E.4(b) of Shaked and Shanthikumar [426], it follows that

$$[\boldsymbol{X}_{\bar{A}} \mid \boldsymbol{X} > \boldsymbol{s}] \geq_{\mathrm{lr}} [\boldsymbol{X}_{\bar{A}} \mid \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}],$$

and, by Theorem 6.E.8 in the same reference, we have

$$[\boldsymbol{X}_{\bar{A}} \mid \boldsymbol{X} > \boldsymbol{s}] \geq_{\mathrm{st}} [\boldsymbol{X}_{\bar{A}} \mid \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}],$$

for any $s \geq 0$. ■

A long list of multivariate distributions are MTP2. For example, a large number of vectors of lifetimes having an archimedean survival copula, or described by means of multivariate frailty models, satisfy this property (see, on this aim, Bassan and Spizzichino [44], or Durante et al. [137], and references therein). Other examples may be found in Marshall and Olkin [308] or Joe [211]. However, there are also many cases where this property is not satisfied, for example, when $\boldsymbol{X}$ does not admit a density. In this case, property (11.2.1) may be verified under alternative conditions, described in the following two statements.

Before giving the next statements, observe that inequality (11.2.1) is verified by all joint distributions that satisfy the dynamic multivariate positive aging notions defined in Shaked and Shanthikumar [420] and references therein. Among them, the weaker one is the property introduced in Norros [361], called *weakened by failures* (WBF): a vector $\boldsymbol{X}$ is said to be WBF if

$$[\boldsymbol{X}_{\bar{A}} - \boldsymbol{s} \mid \boldsymbol{X}_A = \boldsymbol{x}_A, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}] \geq_{\mathrm{st}} [\boldsymbol{X}_{\bar{A}} - \boldsymbol{s} \mid \boldsymbol{X}_A = \boldsymbol{x}_A, X_i = x_i, \boldsymbol{X}_{\bar{A}-\{i\}} > \boldsymbol{s}]$$

for all $A \subseteq I, i \in I$, $\boldsymbol{x}_A \leq \boldsymbol{s}$, and $x_i \leq s$. Clearly, the assumptions of Theorem 11.2.1 are satisfied whenever $\boldsymbol{X}$ is WBF. The next result shows that inequality (11.2.1) is satisfied even under weaker assumptions.

Theorem 11.2.3. *If, for any* $B \subset \bar{A} \subseteq I$, *any* $\boldsymbol{x}_B \geq \boldsymbol{0}$ *and any* $\boldsymbol{y}_{\bar{B}} \geq \boldsymbol{x}_{\bar{B}}$,

$$[\boldsymbol{X}_B - \boldsymbol{x}_B \mid \boldsymbol{X}_B > \boldsymbol{x}_B, \boldsymbol{X}_{\bar{B}} = \boldsymbol{y}_{\bar{B}}] \geq_{\mathrm{uo}} [\boldsymbol{X}_B - \boldsymbol{x}_B \mid \boldsymbol{X}_B > \boldsymbol{x}_B, \boldsymbol{X}_{\bar{B}} = \boldsymbol{x}_{\bar{B}}], \tag{11.2.2}$$

then the inequality (11.2.1) *holds.*

Proof: Without loss of generality, let $\bar{A} = \{1, \dots, k\}$, and fix $\boldsymbol{s} = (s, \dots, s), s \geq 0$. For $i = 2, \dots, k$, set $B = I_{i-1} = \{1, \dots, i-1\}$ in (11.2.2). Let us denote $\bar{I}_i = \{i+1, \dots, n\}$ and $\bar{I}_{i-1} = \{i, \dots, n\}$. Thus, for any $\boldsymbol{y}_{I_{i-1}} \geq \boldsymbol{x}_{I_{i-1}} \geq \boldsymbol{s}$,

$$\begin{aligned}
&\mathrm{P}\{X_i > s+t, \boldsymbol{X}_{\bar{I}_i} > \boldsymbol{s} \mid \boldsymbol{X}_{\bar{I}_{i-1}} > \boldsymbol{s}, \boldsymbol{X}_{I_{i-1}} = \boldsymbol{y}_{I_{i-1}}\} \\
&\quad \geq \ \mathrm{P}\{X_i > s+t, \boldsymbol{X}_{\bar{I}_i} > \boldsymbol{s} \mid \boldsymbol{X}_{\bar{I}_{i-1}} > \boldsymbol{s}, \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}\},
\end{aligned}$$

which implies

$$\begin{aligned}
& \lim_{\boldsymbol{\Delta}\to\boldsymbol{0}+} \frac{\mathrm{P}\{X_i > s+t, \boldsymbol{X} > \boldsymbol{s},\ \boldsymbol{y}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{y}_{I_{i-1}} + \boldsymbol{\Delta}\}}{\mathrm{P}\{\boldsymbol{X} > \boldsymbol{s},\ \boldsymbol{y}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{y}_{I_{i-1}} + \boldsymbol{\Delta}\}} \\
= & \lim_{\boldsymbol{\Delta}\to\boldsymbol{0}+} \frac{\mathrm{P}\{X_i > s+t, \boldsymbol{X}_{\bar{I}_i} > \boldsymbol{s},\ \boldsymbol{y}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{y}_{I_{i-1}} + \boldsymbol{\Delta}\}}{\mathrm{P}\{X_i > s, \boldsymbol{X}_{\bar{I}_i} > \boldsymbol{s},\ \boldsymbol{y}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{y}_{I_{i-1}} + \boldsymbol{\Delta}\}} \\
\geq & \lim_{\boldsymbol{\Delta}\to\boldsymbol{0}+} \frac{\mathrm{P}\{X_i > s+t, \boldsymbol{X}_{\bar{I}_i} > \boldsymbol{s},\ \boldsymbol{x}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{x}_{I_{i-1}} + \boldsymbol{\Delta}\}}{\mathrm{P}\{X_i > s, \boldsymbol{X}_{\bar{I}_i} > \boldsymbol{s},\ \boldsymbol{x}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{x}_{I_{i-1}} + \boldsymbol{\Delta}\}} \\
= & \lim_{\boldsymbol{\Delta}\to\boldsymbol{0}+} \frac{\mathrm{P}\{X_i > s+t, \boldsymbol{X} > \boldsymbol{s},\ \boldsymbol{x}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{x}_{I_{i-1}} + \boldsymbol{\Delta}\}}{\mathrm{P}\{\boldsymbol{X} > \boldsymbol{s},\ \boldsymbol{x}_{I_{i-1}} \leq \boldsymbol{X}_{I_{i-1}} < \boldsymbol{x}_{I_{i-1}} + \boldsymbol{\Delta}\}}.
\end{aligned}$$

This yields, for any $i = 2, \dots, k$ and $\boldsymbol{y}_{\bar{B}} \geq \boldsymbol{x}_{\bar{B}} \geq \boldsymbol{s}$,

$$\mathrm{P}\{X_i > s+t \mid \boldsymbol{X} > \boldsymbol{s}, \boldsymbol{X}_{I_{i-1}} = \boldsymbol{y}_{I_{i-1}}\} \geq \mathrm{P}\{X_i > s+t \mid \boldsymbol{X} > \boldsymbol{s}, \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}\}. \tag{11.2.3}$$

Moreover, the inequality (11.2.2) implies, for $\boldsymbol{y}_{\bar{B}} \geq \boldsymbol{x}_{\bar{B}}$ and $\boldsymbol{y}_B \geq \boldsymbol{x}_B$,

$$\frac{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_{\bar{B}} = \boldsymbol{y}_{\bar{B}}\}}{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_{\bar{B}} = \boldsymbol{y}_{\bar{B}}\}} \geq \frac{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_{\bar{B}} = \boldsymbol{x}_{\bar{B}}\}}{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_{\bar{B}} = \boldsymbol{x}_{\bar{B}}\}}.$$

Denote C the complimentary set of B with respect to $\bar{A}$, i.e., $B \cup C = \bar{A}$ and $B \cap C = \emptyset$. Then, $\bar{B} = A \cup C$. Setting $\boldsymbol{y}_C = \boldsymbol{x}_C$, it follows that

$$\begin{aligned}
&\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{t}_A\} \cdot \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{v}_A\} \\
&\quad \geq \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{t}_A\} \cdot \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{v}_A\},
\end{aligned}$$

for every $\boldsymbol{t}_A \geq \boldsymbol{v}_A$.

Fix any $\boldsymbol{x}_A$, and denote $D_1 = \{\boldsymbol{v}_A : \boldsymbol{0} \leq \boldsymbol{v}_A \leq \boldsymbol{x}_A\}$, $D_2 = \{\boldsymbol{t}_A : \boldsymbol{t}_A \geq \boldsymbol{x}_A\}$. By the previous inequality we have

$$\begin{aligned}
&\int_{D_2} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{t}_A\} \mathrm{d}F_{\boldsymbol{X}_A \mid \boldsymbol{X}_C}(\boldsymbol{t}_A \mid \boldsymbol{x}_C) \\
&\quad \cdot \int_{D_1} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{v}_A\} \mathrm{d}F_{\boldsymbol{X}_A \mid \boldsymbol{X}_C}(\boldsymbol{v}_A \mid \boldsymbol{x}_C)
\end{aligned}$$

$$\geq \int_{D_2} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{t}_A\} \mathrm{d}F_{\boldsymbol{X}_A|\boldsymbol{X}_C}(\boldsymbol{t}_A \mid \boldsymbol{x}_C)$$
$$\cdot \int_{D_1} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{v}_A\} \mathrm{d}F_{\boldsymbol{X}_A|\boldsymbol{X}_C}(\boldsymbol{v}_A \mid \boldsymbol{x}_C),$$

and hence

$$\frac{\int_{D_2} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{t}_A\} \mathrm{d}F_{\boldsymbol{X}_A|\boldsymbol{X}_C}(\boldsymbol{t}_A \mid \boldsymbol{x}_C)}{\int_{D_2} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{t}_A\} \mathrm{d}F_{\boldsymbol{X}_A|\boldsymbol{X}_C}(\boldsymbol{t}_A \mid \boldsymbol{x}_C)}$$
$$\geq \frac{\int_{D_1} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{v}_A\} \mathrm{d}F_{\boldsymbol{X}_A|\boldsymbol{X}_C}(\boldsymbol{v}_A \mid \boldsymbol{x}_C)}{\int_{D_1} \mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A = \boldsymbol{v}_A\} \mathrm{d}F_{\boldsymbol{X}_A|\boldsymbol{X}_C}(\boldsymbol{v}_A \mid \boldsymbol{x}_C)},$$

i.e.,

$$\frac{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B, \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A > \boldsymbol{x}_A\}}{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B, \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A > \boldsymbol{x}_A\}} \geq \frac{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B, \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A \leq \boldsymbol{x}_A\}}{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B, \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A \leq \boldsymbol{x}_A\}}.$$

The last inequality is equivalent to

$$\frac{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A > \boldsymbol{x}_A\}}{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A > \boldsymbol{x}_A\}} \geq \frac{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{y}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A \leq \boldsymbol{x}_A\}}{\mathrm{P}\{\boldsymbol{X}_B > \boldsymbol{x}_B \mid \boldsymbol{X}_C = \boldsymbol{x}_C, \boldsymbol{X}_A \leq \boldsymbol{x}_A\}}, \tag{11.2.4}$$

whenever $\boldsymbol{y}_B \geq \boldsymbol{x}_B$.

Now, setting $B = \bar{A}$, $C = \emptyset$, $\boldsymbol{x}_B = \boldsymbol{s} = (s, \ldots, s)$, and $\boldsymbol{y}_B = (s + t, s, \ldots, s)$ in (11.2.4) yields

$$\begin{aligned}
& \mathrm{P}\{X_1 > t + s \mid \boldsymbol{X} > \boldsymbol{s}\} \\
= & \frac{\mathrm{P}\{X_1 > t + s, \boldsymbol{X}_A > \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}\}}{\mathrm{P}\{\boldsymbol{X}_A > \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}\}} \\
= & \frac{\mathrm{P}\{X_1 > t + s, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s} \mid \boldsymbol{X}_A > \boldsymbol{s}\}}{\mathrm{P}\{\boldsymbol{X}_{\bar{A}} > \boldsymbol{s} \mid \boldsymbol{X}_A > \boldsymbol{s}\}} \\
\geq & \frac{\mathrm{P}\{X_1 > t + s, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s} \mid \boldsymbol{X}_A \leq \boldsymbol{s}\}}{\mathrm{P}\{\boldsymbol{X}_{\bar{A}} > \boldsymbol{s} \mid \boldsymbol{X}_A \leq \boldsymbol{s}\}} \\
= & \mathrm{P}\{X_1 > t + s \mid \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}, \boldsymbol{X}_A \leq \boldsymbol{s}\}, \quad \text{for any } s, t \geq 0.
\end{aligned}$$

That is,

$$[X_1 - s \mid \boldsymbol{X} > \boldsymbol{s}] \geq_{\mathrm{st}} [X_1 - s \mid \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}], \quad \text{for any } s \geq 0. \tag{11.2.5}$$

By (11.2.4) again, letting $i = 2, \dots, k$ and $C = I_{i-1}$, it holds that, for s, $t \geq 0$ and $\boldsymbol{x}_{I_{i-1}} \geq \boldsymbol{s}$,

$$\begin{aligned} & \mathrm{P}\{X_i > t+s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X} > \boldsymbol{s}\} \\ = \; & \frac{\mathrm{P}\{X_i > t+s, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s} \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A > \boldsymbol{s}\}}{\mathrm{P}\{\boldsymbol{X}_{\bar{A}_{i-1}} > \boldsymbol{s} \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A > \boldsymbol{s}\}} \\ \geq \; & \frac{\mathrm{P}\{X_i > t+s, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s} \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A \leq \boldsymbol{s}\}}{\mathrm{P}\{\boldsymbol{X}_{\bar{A}_{i-1}} > \boldsymbol{s} \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A \leq \boldsymbol{s}\}} \\ = \; & \mathrm{P}\{X_i > t+s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}\}. \end{aligned}$$

That is, for $i = 2, \dots, k$,

$$[X_i - s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X} > \boldsymbol{s}] \geq_{\mathrm{st}} [X_i - s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}].$$

On the other hand, by (11.2.3), we have, for $\boldsymbol{y}_{I_{i-1}} \geq \boldsymbol{x}_{I_{i-1}} \geq \boldsymbol{s}$,

$$[X_i - s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{y}_{I_{i-1}}, \boldsymbol{X} > \boldsymbol{s}] \geq_{\mathrm{st}} [X_i - s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X} > \boldsymbol{s}],$$

and thus

$$[X_i - s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{y}_{I_{i-1}}, \boldsymbol{X} > \boldsymbol{s}] \geq_{\mathrm{st}} [X_i - s \mid \boldsymbol{X}_{I_{i-1}} = \boldsymbol{x}_{I_{i-1}}, \boldsymbol{X}_A \leq \boldsymbol{s}, \boldsymbol{X}_{\bar{A}} > \boldsymbol{s}]. \tag{11.2.6}$$

Finally, by applying Theorem 6.B.3 of Shaked and Shanthikumar [426] to (11.2.5) and (11.2.6), we reach the desired result (11.2.1). ■

The following statement provides alternative conditions for (11.2.1) in the bivariate case.

Theorem 11.2.4. *If X_2 is RTI in X_1 and X_1 is RTI in X_2, then, for any $s \geq 0$,*

$$[X_1 - s \mid X_1 > s, X_2 > s] \geq_{\mathrm{st}} [X_1 - s \mid X_1 > s, X_2 \leq s]$$

and

$$[X_2 - s \mid X_2 > s, X_1 > s] \geq_{\mathrm{st}} [X_2 - s \mid X_2 > s, X_1 \leq s].$$

That is, the inequality (11.2.1) holds.

Proof: Let $s, t \geq 0$ and denote

$$\begin{aligned} A &= \{X_1 > s+t, X_2 > s\}, \\ B &= \{s+t \geq X_1 > s, X_2 > s\}, \end{aligned}$$

$$\begin{aligned} C &= \{X_1 > s+t, X_2 \le s\}\ , \\ D &= \{s+t \ge X_1 > s, X_2 \le s\}\ . \end{aligned}$$

Since X_2 is RTI in X_1, it holds that

$$\frac{\mathrm{P}\{A\}}{\mathrm{P}\{A \cup C\}} = \frac{\mathrm{P}\{X_1 > s+t, X_2 > s\}}{\mathrm{P}\{X_1 > s+t\}} \ge \frac{\mathrm{P}\{X_1 > s, X_2 > s\}}{\mathrm{P}\{X_1 > s\}} = \frac{\mathrm{P}\{A \cup B\}}{\mathrm{P}\{A \cup B \cup C \cup D\}}.$$

Note that A, B, C, and D are mutually exclusive, the above inequality may be rephrased as

$$\frac{\mathrm{P}\{A\}}{\mathrm{P}\{A\} + \mathrm{P}\{C\}} \ge \frac{\mathrm{P}\{A\} + \mathrm{P}\{B\}}{\mathrm{P}\{A\} + \mathrm{P}\{C\} + \mathrm{P}\{B\} + \mathrm{P}\{D\}}.$$

Equivalently,

$$\frac{\mathrm{P}\{A\}}{\mathrm{P}\{A\} + \mathrm{P}\{C\}} \ge \frac{\mathrm{P}\{B\}}{\mathrm{P}\{B\} + \mathrm{P}\{D\}},$$

which implies

$$\mathrm{P}\{A\} \cdot \mathrm{P}\{D\} \ge \mathrm{P}\{B\} \cdot \mathrm{P}\{C\},$$

and hence

$$\mathrm{P}\{A\} \cdot \mathrm{P}\{D\} + \mathrm{P}\{A\} \cdot \mathrm{P}\{C\} \ge \mathrm{P}\{B\} \cdot \mathrm{P}\{C\} + \mathrm{P}\{A\} \cdot \mathrm{P}\{C\}.$$

This is just

$$\frac{\mathrm{P}\{A\}}{\mathrm{P}\{A \cup B\}} \ge \frac{\mathrm{P}\{C\}}{\mathrm{P}\{C \cup D\}}.$$

Consequently, we have, for any $s, t \ge 0$,

$$\begin{aligned} & \mathrm{P}\{X_1 > s+t \mid X_1 > s, X_2 > s\} \\ = \ & \frac{\mathrm{P}\{X_1 > s+t, X_2 > s\}}{\mathrm{P}\{X_1 > s, X_2 > s\}} \\ = \ & \frac{\mathrm{P}\{A\}}{\mathrm{P}\{A \cup B\}} \\ \ge \ & \frac{\mathrm{P}\{C\}}{\mathrm{P}\{C \cup D\}} \\ = \ & \frac{\mathrm{P}\{X_1 > s+t, X_2 \le s\}}{\mathrm{P}\{X_1 > s, X_2 \le s\}} \\ = \ & \mathrm{P}\{X_1 > s+t \mid X_1 > s, X_2 \le s\}. \end{aligned}$$

That is, $[X_1 - s \mid X_1 > s, X_2 > s] \ge_{\mathrm{st}} [X_1 - s \mid X_1 > s, X_2 \le s]$.

In a completely similar manner, we also have, for any $s \ge 0$

$$[X_2 - s \mid X_2 > s, X_1 > s] \ge_{\mathrm{st}} [X_2 - s \mid X_2 > s, X_1 \le s].$$

Thus, (11.2.1) is validated. ∎

11.3 Sufficient Conditions for Positive Aging Properties

Conditions under which lifetimes of coherent systems satisfy aging properties have been studied extensively in the literature (see, e.g., Barlow and Proschan [39], or Lai and Xie [269]), in most of the cases under the assumption of independence among component's lifetimes. Some interesting results dealing with the case of dependent components have been recently shown, for example, in Hu and Li [199] and Navarro and Shaked [351], where conditions on the joint density of the vector of component's lifetimes such that parallel and series systems have monotonic hazard and reverse hazard rates are described. Some results in the same spirit, but for more general coherent systems and weaker aging notions, are provided in this section.

Denote with $X_t = (X - t \mid X > t)$ the residual life of a random lifetime X at time $t \geq 0$. The following are among the most important univariate aging concepts.

Definition 11.3.1. A nonnegative random variable X is said to be

(i) *New better than used* (NBU) if $X \geq_{\mathrm{st}} X_t$ for all $t \geq 0$

(ii) *New better than used in the 2nd stochastic dominance* (NBU(2)) if $X \geq_{\mathrm{icv}} X_t$ for all $t \geq 0$

(iii) *New better than used in the increasing convex order* (NBUC) if $X \geq_{\mathrm{icx}} X_t$ for all $t \geq 0$

The aging notions defined above can be generalized to the multivariate setting as follows. Denote with

$$\boldsymbol{X}_t = [(X_1 - t, \ldots, X_n - t) \mid X_1 > t, \ldots, X_n > t]$$

the residual life vector of $\boldsymbol{X}$ at time $t \geq 0$.

Definition 11.3.2. A nonnegative random vector $\boldsymbol{X}$ is said to be

(i) *Multivariate new better than used* (M-NBU) if $\boldsymbol{X} \geq_{\mathrm{st}} \boldsymbol{X}_t$ for all $t \geq 0$

(ii) *Multivariate new better than used in the 2nd stochastic dominance* (M-NBU(2)) if $\boldsymbol{X} \geq_{\mathrm{icv}} \boldsymbol{X}_t$ for all $t \geq 0$

(iii) *Multivariate new better than used in the increasing convex order* (M-NBUC) if $\boldsymbol{X} \geq_{\text{icx}} \boldsymbol{X}_t$ for all $t \geq 0$

Readers may refer to Pellerey [370] or Li and Pellerey [292] for examples of bivariate distributions with the M-NBU property.

According to Theorem 5.1 of Barlow and Proschan [39], a coherent system may inherit the NBU property of its independent components. Theorem 11.3.3 below builds this preservation property for coherent systems of dependent components. Note that the assumption in (11.2.1) holds when all concerned components are mutually independent; thus, Theorem 11.3.3 forms an interesting extension for Theorem 5.1 of Barlow and Proschan [39].

Theorem 11.3.3. *Under the assumption of* (11.2.1)*, any coherent system is NBU whenever the components' lifetimes vector* $\boldsymbol{X}$ *is M-NBU.*

Proof: By Theorem 11.2.1 and inequality (11.1.2), we have

$$[T_{\boldsymbol{X}} - s \mid T_{\boldsymbol{X}} > s] \leq_{\text{st}} [T_{\boldsymbol{X}} - s \mid \boldsymbol{X} > \boldsymbol{s}] \stackrel{st}{=} T_{\boldsymbol{X}_s}, \quad \text{for any } s \geq 0.$$

The M-NBU property of $\boldsymbol{X}$ implies $\boldsymbol{X}_s \leq_{\text{st}} \boldsymbol{X}$ for any $s \geq 0$. Due to the monotonicity of the coherent life functions, we have

$$T_{\boldsymbol{X}_s} \leq_{\text{st}} T_{\boldsymbol{X}}, \quad \text{for any } s \geq 0.$$

Thus, it holds that

$$[T_{\boldsymbol{X}} - s \mid T_{\boldsymbol{X}} > s] \leq_{\text{st}} T_{\boldsymbol{X}}, \quad \text{for any } s \geq 0.$$

This completes the proof. ■

Example 11.3.4. Consider a random vector $\boldsymbol{X}$ having the joint survival function

$$\bar{F}(x_1, \ldots, x_n) = \left(\frac{e^{bx_1} + e^{bx_2} + \cdots + e^{bx_n}}{n} \right)^{-\theta}, \qquad \theta, b > 0.$$

One may easily verify that the series system of these components has the reliability function $e^{-b\theta x}$ of an exponential distribution and thus is NBU. In fact, it can be verified that $\boldsymbol{X}$ has MTP2 density and satisfies the M-NBU property (Pellerey [370]). According to Theorem 11.3.3, any coherent system with components having lifetimes vector $\boldsymbol{X}$ is also of NBU property. ■

Example 11.3.5. Consider the random vector $\boldsymbol{X}$ having a Marshall-Olkin bivariate exponential distribution, i.e., having joint survival function

$$\overline{F}(x_1, x_2) = \mathrm{P}\{X_1 > x_1,\, X_2 > x_2\} = \exp\big\{-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_3 (x_1 \vee x_2)\big\},$$

with $x_1, x_2 \geq 0$ and $\lambda_i \geq 0$, $i = 1, 2, 3$. As shown in Corollary 4.2 in Li and Pellerey [292], such a vector $\boldsymbol{X}$ satisfies the M-NBU property. Moreover, even if it does not satisfy the MTP2 property because of the singularity due to $\mathrm{P}\{X_1 = X_2\} > 0$, it satisfies the RTI property, as can be easily verified. Thus, according to Theorem 11.3.3 and Theorem 11.2.4, the lifetime $T_{\boldsymbol{X}}$ of any coherent system whose components' lifetimes are described by $\boldsymbol{X}$ is NBU. ■

In a similar fashion, we may build the following result, which serves as a generalization of Theorem 1 in Pellerey and Petakos [371].

Theorem 11.3.6. *Under the assumption* (11.2.1)*, any coherent system with convex [concave] coherent life function has a lifetime* $T_{\boldsymbol{X}}$ *which is NBUC [NBU(2)] whenever the components vector* $\boldsymbol{X}$ *is M-NBUC [M-NBU(2)].*

As an immediate consequence, we get Corollary 11.3.7 below, which generalizes the preservation properties of NBUC and NBU(2) aging notions under parallel (series) systems with independent components due to Li et al. [290] and Li and Kochar [289].

Corollary 11.3.7. *Under the assumption* (11.2.1)*, the lifetime of a parallel [series] system is NBUC [NBU(2)] whenever the vector of components' lifetimes* $\boldsymbol{X}$ *is M-NBUC [(M-NBU(2)].*

Acknowledgement

Authors would like to thank Professor Jorge Navarro for illuminating discussions on the properties of coherent systems, which invoked our interest in the subject of this note. Xiaohu Li and Yinping You are supported by National Natural Science Foundation of China (10771090).

Chapter 12

On Allocation of Active Redundancies to Systems: A Brief Review

Xiaohu Li and Weiyong Ding

Abstract: In reliability engineering and system security, it is of great practical interest to allocate active redundancies at either component or system level so as to enhance the system's lifetime or improve some other performances of the system. This topic has been paid much attention ever since reliability theory took its form in the early 1970s. In this paper, we review these results in the reliability literature on allocating active redundancies to a coherent system. Using these theories, engineers are guided to design a better system or to optimize the performance of the original system in the sense of some stochastic orders. Recent theoretical results as well as some applications are also presented.

X. Li (✉)
School of Mathematical Sciences, Xiamen University, Xiamen 361005, China
e-mail: mathxhli@hotmail.com

W. Ding
College of Science, Hebei United University, Tangshan 063009, China
e-mail: mathdwy@hotmail.com

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_12,

12.1 Introduction and Preliminaries

A coherent system is one for which the structure function is monotone in each component and in which every component is relevant, i.e., the behavior of a component does affect the performance of the system (Barlow and Proschan [39]). In particular, the k-out-of-n system, which functions if and only if at least k-out-of-n its components work, has been paid much special attention. In the context of reliability, the lifetime of the k-out-of-n system corresponds to the $(n-k+1)$-th order statistic $X_{n-k+1:n}$ from random lifetimes $X_1, X_2, \ldots, X_n$. Specifically, the lifetimes of series and parallel systems are just the smallest order statistic $X_{1:n}$ and the largest one $X_{n:n}$, respectively.

In reliability engineering and system security, it is of great practical interest to allocate active redundancies at either component or system level so as to enhance the system's lifetime or improve some other indices of the system's performance. In general, the following two types of allocations are commonly used in practice, namely, (i) *active redundancy* (hot standby), which is put in parallel to component/system and starts functioning at the same time as the component/system is initiated, and (ii) *standby redundancy* (cold standby), which is put in standby and starts functioning once the working component/system fails. Recently, Cha et al. [82] considered the so-called *general standby*, which operates in a milder environment in the standby state (and hence the hazard rate is nonzero and smaller than that in the usual environment) and bear the normal stress in a working state. Obviously, this is just an intermediate state between the cold standby and the hot one. We will only focus on the active redundancy. For more on cold standby and general standby, please refer to Boland et al. [70], Cha et al. [82], EI-Newehi and Sethuraman [147], Li et al. [295], Li et al. [293], Misra et al. [325], and Shaked and Shanthikumar [421]. It is worth mentioning that this review is by no means an exhaustive summary of all related works on active redundancy allocation. The papers we selected for inclusion only reflect our focused theme.

As far as we know, it was pointed out in Barlow and Proschan [39] that for a coherent system, allocating active redundancies at component level is more effective than allocating them at system level. EI-Newehi et al. [146] were among the first to employ the powerful tools of majorization and Schur-convex/Schur-concave function to pursue the optimal allocation of components to parallel-series and series-parallel systems in the sense of maximizing the system's reliability function, and hence majorization plays a rather important role in comparing policies to allocate active redundancies. Till now several models on allocation of active redundancy to coherent system have been developed and they can be reasonably compared in terms of various stochastic orders.

The rest of this paper is organized as follows: Sect. 12.2 includes results on stochastic comparison between the lifetime of a coherent system with active redundancies at component level and that at system level. In Sect. 12.3, we discuss the allocation of more than one active redundancies to a k-out-of-n system; the optimal allocation policy is presented in some interesting scenarios such as the stochastically ordered working components and mutually independent working components and redundancies. Finally, Sect. 12.4 concludes with considering other models of allocating active redundancies to a coherent system.

In order to be self-contained, let us recall definitions of the following stochastic orders. For further detail, see Müller and Stoyan [335] and Shaked and Shanthikumar [426].

Definition 12.1.1. Let X and Y be two random variables with absolutely continuous cumulative distribution functions F and G, probability density functions f and g, and survival functions $\bar{F} = 1 - F$ and $\bar{G} = 1 - G$, respectively. X is said to be smaller than Y in the following:

(i) *Usual stochastic order*, denoted by $X \leq_{\mathrm{st}} Y$, if $\bar{F}(x) \leq \bar{G}(x)$ for all x

(ii) *Hazard rate order*, denoted by $X \leq_{\mathrm{hr}} Y$, if $\bar{G}(x)/\bar{F}(x)$ is increasing in x

(iii) *Reversed hazard rate order*, denoted by $X \leq_{\mathrm{rh}} Y$, if $G(x)/F(x)$ is increasing in x

(iv) *Likelihood ratio order*, denoted by $X \leq_{\mathrm{lr}} Y$, if $g(x)/f(x)$ is increasing in x

(v) *Increasing convex order*, denoted by $X \leq_{\mathrm{icx}} Y$, if $\int_x^{\infty} \bar{F}(t)\mathrm{d}t \leq \int_x^{\infty} \bar{G}(t)\mathrm{d}t$ for all x

(vi) *Increasing concave order*, denoted by $X \leq_{\mathrm{icv}} Y$, if $\int_{-\infty}^{x} F(t)\mathrm{d}t \geq \int_{-\infty}^{x} G(t)\mathrm{d}t$ for all x

For ease of reference, implications among these orders are presented in the following diagram:

$$\begin{array}{ccccccc} & & X \le_{\mathrm{hr}} Y & & & & X \le_{\mathrm{icx}} Y \\ & \nearrow & & \searrow & & \nearrow & \\ X \le_{\mathrm{lr}} Y & & & & X \le_{\mathrm{st}} Y & & \\ & \searrow & & \nearrow & & \searrow & \\ & & X \le_{\mathrm{rh}} Y & & & & X \le_{\mathrm{icv}} Y \end{array}$$

The other interesting notion is the majorization order, which is quite useful in establishing various inequalities. For extensive and comprehensive discussions on the theory of majorization order and its application, readers may refer to Marshall et al. [312].

Definition 12.1.2. Let $x_{(1)} \le \cdots \le x_{(n)}$ be the increasing arrangement of the components of the vector $\boldsymbol{x} = (x_1, \ldots, x_n)$.

(i) $\boldsymbol{x}$ is said to *majorize* $\boldsymbol{y} = (y_1, \ldots, y_n)$ (denoted by $\boldsymbol{x} \overset{m}{\succeq} \boldsymbol{y}$) if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$ and

$$\sum_{i=1}^{j} x_{(i)} \le \sum_{i=1}^{j} y_{(i)}, \qquad j = 1, \ldots, n-1.$$

(ii) A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is said to be *Schur-convex* (*Schur-concave*) if $\boldsymbol{x} \overset{m}{\succeq} \boldsymbol{y}$ implies

$$\phi(\boldsymbol{x}) \ge (\le) \phi(\boldsymbol{y}), \qquad \text{for } \boldsymbol{x},\, \boldsymbol{y} \in \mathbb{R}^n.$$

12.2 Redundancy at Component Level Versus that at System Level

In order to improve the reliability of the system, a measure to take in practice is to allocate one redundancy to each component. Since those redundant components may form another system according to the same structure, we have two choices: (i) at component level, that is, put one spare in parallel to one component, and (ii) at system level, that is, put the other system composed of spares in parallel to the system of working components. Now, one question arises naturally in this situation: should the redundancies be allocated to the system at component level or at system level?

$\boldsymbol{X} = (X_1, \ldots, X_n)$	Lifetimes of n independent components
$\boldsymbol{Y} = (Y_1, \ldots, Y_n)$	Lifetimes of n independent active spares
F_i, G_i	Distribution functions of X_i and Y_i, $i = 1, \ldots, n$
$\boldsymbol{X} \vee \boldsymbol{Y}$	$(X_1 \vee Y_1, X_2 \vee Y_2, \ldots, X_n \vee Y_n)$, here $x \vee y = \max\{x, y\}$
$T(\boldsymbol{X})$	Lifetime of a coherent system with components' lifetimes $\boldsymbol{X}$
$T_{k\|n}(\boldsymbol{X})$	Lifetime of a k-out-of-n system with components' lifetimes $\boldsymbol{X}$
$h(p)$	Reliability of a coherent system with i.i.d. components

Here is the milestone conclusion: *it is better to allocate the redundancies at component level than at the system level in the sense of attaining stochastically larger lifetime of the redundant system.* This principle is well known among design engineers.

Theorem 12.2.1 (Esary, Marshall and Proschan [153])**.** *For a coherent system with the structural function T, it holds that*

$$T(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\text{st}} T(\boldsymbol{X}) \vee T(\boldsymbol{Y}), \tag{12.2.1}$$

for any n-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$.

Naturally, one may wonder whether allocating redundancies at component level is also better than allocating redundancy at system level in some other stronger orders such as the (reversed) hazard rate order and the likelihood ratio order. Boland and EI-Newehi [66] provided a generally negative answer by displaying a simple counterexample, in which a series system consists of two components having a common standard exponential distribution and two spares have exponential distributions with common rate 2. That is,

$$X_1 \overset{\text{st}}{=} X_2 \sim \exp\{-t\} \quad \text{and} \quad Y_1 \overset{\text{st}}{=} Y_2 \sim \exp\{-2t\}.$$

However, as will be seen, for several specific coherent structures, the stronger stochastic order may be possible. Now, let us present these results in the literature in the following two scenarios: coherent systems with matching spares and nonmatching spares, respectively.

12.2.1 Redundancies with Matching Spares

For a series system which is the simplest popular structure in coherent systems, allocating redundancies at component level is more effective than at system level in the sense of the hazard rate order.

Theorem 12.2.2 (Boland and EI-Newehi [66]). *Suppose* $X_i \overset{st}{=} Y_i$ *for* $i = 1, 2, \ldots, n$. *Then,*

$$T_{n|n}(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{hr}} T_{n|n}(\boldsymbol{X}) \vee T_{n|n}(\mathbf{Y}).$$

However, Example 12.2.3 below tells that the hazard rate order in Theorem 12.2.2 cannot be replaced by the reversed hazard rate order and hence nor by the likelihood ratio order.

Example 12.2.3. Suppose the lifetimes X_1 and X_2 of the two components in a series system have the standard exponential distribution and Fréchet distribution, respectively. Denote Y_1 and Y_2 the lifetimes of their respective matching spares. That is,

$$X_1 \overset{\mathrm{st}}{=} Y_1 \sim \exp\{-t\} \quad \text{and} \quad X_2 \overset{\mathrm{st}}{=} Y_2 \sim 1 - \exp\{-t^{-1}\}.$$

As can be seen in Fig. 12.1, the reversed hazard rates of the system with redundancies at component level and that with redundancy at system level cross with each other. That is to say, the reversed hazard rate order does not exist at all. ■

When the working components of a coherent system and their matching spares are independent and identically distributed, allocating redundancies at component level is more effective than allocating redundancy at system level in the sense of stronger stochastic orders for specific coherent structures.

Theorem 12.2.4 (Boland and EI-Newehi [66]). *Suppose lifetimes of the components and their spares are i.i.d.*

(i) If $ph'(p)/h(p)$ *decreases in* p *and* $h(p) < p$ *for all* $p \in (0, 1)$, *then*

$$T(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{hr}} T(\boldsymbol{X}) \vee T(\boldsymbol{Y}).$$

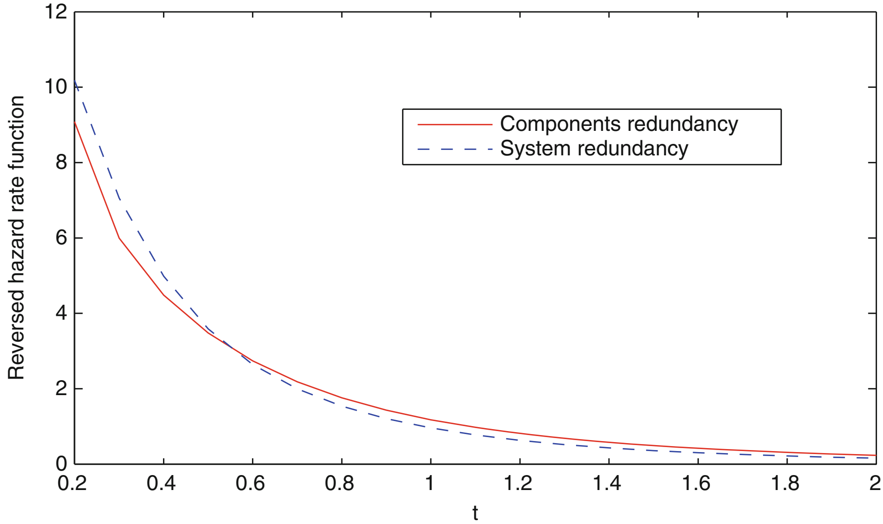


Figure 12.1: A series system with redundancy at component and system levels

(ii) It holds that $T_{2|n}(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{rh}} T_{2|n}(\boldsymbol{X}) \vee T_{2|n}(\boldsymbol{Y})$

Since the structure of an $(n-1)$-out-of-n does not satisfy the condition in Theorem 12.2.4(i), Theorem 12.2.4(ii) is imperative. Afterward, this theorem was improved and generalized as follows:

Theorem 12.2.5 (Singh and Singh[435]). *Suppose lifetimes of the components and their spares are i.i.d. Then,*

$$T_{k|n}(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{lr}} T_{k|n}(\boldsymbol{X}) \vee T_{k|n}(\boldsymbol{Y}).$$

for $k = 1, \ldots, n$*.*

Recently, Theorem 12.2.5 was further extended to the general coherent systems as below.

Theorem 12.2.6 (Misra et al. [322]). *Suppose that a coherent system has the reliability function* $h(p)$*, the lifetimes of components and the spares are i.i.d. If*

$$\frac{1-h(p)}{1-p}\frac{h'(p)}{h'(p(2-p))}$$

increases in $p \in (0, 1)$*, then,*

$$T(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{lr}} T(\boldsymbol{X}) \vee T(\boldsymbol{Y}).$$

Remark 12.2.7. One can verify that the reliability function of a k-out-of-n system satisfies the monotone condition in Theorem 12.2.6. Thus, Theorem 12.2.5 is just a specific case of Theorem 12.2.6. In addition, it should be mentioned here that, in order to establish the likelihood ratio order, both Singh and Singh [435] and Misra et al. [322] made the tacit assumption that the common distribution of the lifetimes of both components and spares is absolutely continuous. By virtue of *signature*, Kochar et al. [253] made up this drawback through providing an alternative proof under the assumption of the continuity of the common lifetime distribution. For comprehensive studies on signature, we refer readers to Samaniego [408].

12.2.2 Redundancies with Nonmatching Spares

Here, we assume that lifetimes of the working components in a coherent systems are i.i.d. components and lifetimes of all its spares are also i.i.d., but the two distributions are not necessary identical. In this context, Misra et al. [322] proved for some particular coherent structures that allocating redundancies at component level is more effective than allocating a redundant system at system level in the sense of attaining larger reversed hazard rate.

Theorem 12.2.8 (Misra et al. [322])**.** *Suppose that a coherent system composed of i.i.d. components has reliability function $h(p)$ and the spares are i.i.d. also. If, for any fixed $q \in (0,1)$,*

$$\frac{h'(1-(1-p)q)}{h'(p)} \quad \textit{decreases in } p \in (0,1),$$

then,

$$T(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{rh}} T(\boldsymbol{X}) \vee T(\boldsymbol{Y}).$$

Corollary 12.2.9 follows from Theorem 12.2.6 immediately.

Corollary 12.2.9. *For a k-out-of-n system with i.i.d. components and i.i.d. spares, it holds that*

$$T_{k|n}(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\mathrm{rh}} T_{k|n}(\boldsymbol{X}) \vee T_{k|n}(\boldsymbol{Y}),$$

for $k = 1, \ldots, n$.

As illustrated by a counterexample due to Boland and EI-Newehi [66], the hazard rate order generally does not hold for k-out-of-n systems, and hence neither does the likelihood rate order.

From several our numerical experiments, it seems that Theorem 12.2.2 may be generalized to k-out-of-n systems. That is, for a k-out-of-n system, if $X_i \stackrel{\text{st}}{=} Y_i$, $i = 1, 2, \ldots, n$, then,

$$T_{k|n}(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\text{hr}} T_{k|n}(\boldsymbol{X}) \vee T_{k|n}(\boldsymbol{Y}), \qquad k = 1, \ldots, n.$$

However, to the best of our knowledge, this is not proved/disproved yet and hence is still an open problem.

12.3 Allocation of Active Redundancies to a k-out-of-n System

Boland et al. [68] and Shaked and Shanthikumar [421] were among the first to study the optimal allocation of active redundancies to coherent systems by means of the majorization order. Other than the technique based on the *redundancy importance* due to Boland et al. [68], Shaked and Shanthikumar [421] exploited the stochastic order to deal with the optimality. Consequently, researchers focused their attention on comparing allocation policies of active redundancies to coherent system with respect to a variety of stochastic orders.

This section deals with the optimal allocation of m i.i.d. active redundancies to a k-out-of-n system. The main work in the literature involves the following three situations: (i) components and redundancies are i.i.d., (ii) component and redundancies are independent but not necessarily identical, and (iii) components are stochastically ordered. All results summarized here justify the following intuitively reasonable fact: *by balancing the allocation policy of active redundancies as much as possible, one can stochastically maximize the lifetime of the redundant k-out-of-n system.*

$\boldsymbol{X} = (X_1, \ldots, X_n)$	Lifetimes of n independent components
$\boldsymbol{Y} = (Y_1, \ldots, Y_m)$	Lifetimes of m independent and identically distributed spares
F_i, G	Distribution functions of components X_i and spare Y_i's
$\boldsymbol{r} = (r_1, \ldots, r_n)$	Allocation policy with r_i redundancies allocated to X_i
$\boldsymbol{s} = (s_1, \ldots, s_n)$	Allocation policy with s_i redundancies allocated to X_i
$T_{k\|n}(\boldsymbol{r})$	Lifetime of the k-out-of-n redundant system with policy $\boldsymbol{r}$

12.3.1 The Case with i.i.d. Components and Redundancies

Shaked and Shanthikumar [421] were among the first to study the problem of allocating m active redundancies to a parallel system with n components in the situation that lifetimes of components and redundancies are independent and identically distributed. They developed the following influential benchmark in this line of research: many authors devoted themselves to pursuing more general results over the past two decades.

Theorem 12.3.1 (Shaked and Shanthikumar [421])**.** *Suppose that the working components and redundancies are independent and identically distributed. Then, for two allocation policies* $\boldsymbol{r}$ *and* $\boldsymbol{s}$*,*

$$T_{n|n}(\boldsymbol{s}) \le_{\rm st} T_{n|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s}.$$

Since the hazard rate order is strictly stronger than the usual stochastic order, an allocation policy that stochastically optimizes the lifetime of a redundant series system may not also optimize the hazard rate of the system. However, it was found subsequently that the allocation policy stochastically optimizing the lifetime of a series system does optimizes the hazard rate of the system.

Theorem 12.3.2 (Singh and Singh [436])**.** *Suppose that components and redundancies are independent and identically distributed. For two allocation policies* $\boldsymbol{s}$ *and* $\boldsymbol{r}$*,*

$$T_{n|n}(\boldsymbol{s}) \le_{\rm hr} T_{n|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s}.$$

Afterward, Hu and Wang [201] extended the result on the series system in Theorem 12.3.1 to the k-out-of-n systems.

Theorem 12.3.3 (Hu and Wang [201])**.** *Suppose that components and redundancies are independent and identically distributed. For two allocation policies* $\boldsymbol{s}$ *and* $\boldsymbol{r}$*,*

$$T_{k|n}(\boldsymbol{s}) \le_{\rm st} T_{k|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s},$$

for $k = 1, \ldots, n$*.*

It should be remarked here that the above result due to Hu and Wang [201] can also be derived from Theorem 2.13 of Boland et al. [68]. Recently, we further strengthen the stochastic order in Theorem 12.3.3 to the hazard rate order. Thus, the allocation policy stochastically optimizing the lifetime of a k-out-of-n system also optimizes the hazard rate of the system.

Theorem 12.3.4 (Ding and Li [133])**.** *Suppose that components and redundancies are independent and identically distributed. For two allocation policies* $\boldsymbol{s}$ *and* $\boldsymbol{r}$*,*

$$T_{k|n}(\boldsymbol{s}) \leq_{\rm hr} T_{k|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s},$$

for $k = 1, \ldots, n$*.*

Hu and Wang [201] and Misra et al. [322] independently demonstrated through counterexamples that the hazard rate order in Theorem 12.3.4 cannot be replaced by the reversed hazard rate for coherent systems with more than two components although they simultaneously pointed out the following fact:

Theorem 12.3.5 (Hu and Wang [201]; Misra et al. [322])**.** *Suppose that components and redundancies are independent and identically distributed. For two allocation policies* $\boldsymbol{s}$ *and* $\boldsymbol{r}$*,*

$$T_{2|2}(\boldsymbol{s}) \leq_{\rm rh} T_{2|2}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s}.$$

Finally, Hu and Wang [201] conjectured that the above reversed hazard rate order on the series systems with two components may be strengthened to the likelihood ratio order. However, this is still an open problem.

12.3.2 The Case with i.i.d Components and i.i.d. Redundancies

In the situation that components and redundancies have nonidentical life distributions, Boland et al. [68] by means of *redundancy importance* showed that the survival function of the lifetime of a redundant k-out-of-n system is Schur concave, and this result may be precisely stated as follows:

Theorem 12.3.6 (Boland et al. [68]). *Suppose that components and redundancies are independent and identically distributed, respectively, and the two common distributions are not necessarily identical. For allocation policies* $\boldsymbol{s}$ *and* $\boldsymbol{r}$*,*

$$T_{k|n}(\boldsymbol{s}) \leq_{\text{st}} T_{k|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s},$$

for $k = 1, \ldots, n$*.*

In contrast to Theorem 12.3.4, which extended the usual stochastic order to the hazard rate order for k-out-of-n systems with i.i.d. components and redundancies, Misra et al. [322] illustrated through a series system with two components in Theorem 12.3.6 that the usual stochastic order cannot be upgraded to the hazard rate order when components and redundancies have nonidentical distributions. However, they obtained the hazard rate order for the series system.

Theorem 12.3.7 (Misra et al. [322]). *Suppose that components and redundancies are independent and identically distributed, respectively, and the two common distributions are not necessarily identical. For allocation policies* $\boldsymbol{r}$ *and* $\boldsymbol{s}$*, if* $\ln G(x)/\ln F(x)$ *increases in* x*, then,*

$$T_{n|n}(\boldsymbol{s}) \leq_{\text{hr}} T_{n|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s}.$$

We conjecture that Theorem 12.3.7 above may be extended to k-out-of-n systems. Precisely, if $\ln G(x)/\ln F(x)$ increases in x, then,

$$T_{k|n}(\boldsymbol{s}) \leq_{\text{hr}} T_{k|n}(\boldsymbol{r}) \qquad \text{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s},$$

for $k = 1, \ldots, n$.

12.3.3 The Case with Stochastically Ordered Components

So far, all the developments on this topic are based on the critical assumption that lifetimes of components in the system are i.i.d. However, it may be more realistic to consider a system consisting of independent and heterogeneous components. Misra et al. [322] firstly considered the system with stochastically ordered lifetimes of components and i.i.d. redundancies.

Denote the class of allocation policies

$$\mathcal{R}_m = \{(r_1, \ldots, r_n) : r_1 \geq \cdots \geq r_n,\, r_1 + \cdots + r_n = m\}.$$

Misra et al. [322] showed the following theorem, asserting that the lifetime of a redundant series system may be stochastically increased through balancing the allocation policy in $\mathcal{R}_m$.

Theorem 12.3.8 (Misra et al. [322])**.** *Suppose $X_1 \geq_{\rm st} \cdots \geq_{\rm st} X_n$ and $Y_1, \ldots, Y_m$ are i.i.d.. For allocation policies $\boldsymbol{r}, \boldsymbol{s} \in \mathcal{R}_m$, if $Y_1 \geq_{\rm st} X_1$, then,*

$$T_{n|n}(\boldsymbol{s}) \leq_{\rm st} T_{n|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s}.$$

However, intuitively, one needs to allocate more redundancies to less reliable components. In fact, we achieved the nice result for redundant k-out-of-n systems.

Theorem 12.3.9 (Li and Ding [287])**.** *Suppose $X_1 \geq_{\rm st} \cdots \geq_{\rm st} X_n$, $Y_1, \ldots, Y_m$ are i.i.d. and $Y_1 \geq_{\rm st} X_1$. Let two policies $\boldsymbol{r}$ and $\boldsymbol{s}$ such that $r_l = s_l$ for $l \notin \{i, j\}$ $(i < j)$. Then, for any $k = 1, \ldots, n$,*

$$T_{k|n}(\boldsymbol{s}) \leq_{\rm st} T_{k|n}(\boldsymbol{r}) \qquad \textit{if and only if} \qquad r_j = s_i > s_j = r_i$$

and

$$T_{k|n}(\boldsymbol{s}) \leq_{\rm st} T_{k|n}(\boldsymbol{r}) \quad \textit{whenever } s_i - s_j \geq 2,\ r_i = s_i + 1 \textit{ and } r_j = s_j - 1.$$

This motivates us to further consider the so-called class of *admissible* allocation policies

$$\bar{\mathcal{R}}_m = \{(r_1, \ldots, r_n) : r_1 \leq \cdots \leq r_n,\, r_1 + \cdots + r_n = m\}.$$

instead. And it was found once again that the lifetime of a redundant k-out-of-n system may be stochastically increased through balancing the allocation policy in $\bar{\mathcal{R}}_m$.

Theorem 12.3.10 (Li and Ding [287])**.** *Suppose $X_1 \geq_{\rm st} \cdots \geq_{\rm st} X_n$ and $Y_1, \ldots, Y_m$ are i.i.d. For allocation policies $\boldsymbol{r}, \boldsymbol{s} \in \bar{\mathcal{R}}_m$, if $Y_1 \geq_{\rm st} X_1$, then,*

$$T_{k|n}(\boldsymbol{s}) \leq_{\rm st} T_{k|n}(\boldsymbol{r}) \qquad \textit{whenever } \boldsymbol{r} \overset{m}{\preceq} \boldsymbol{s},$$

for $k = 1, \ldots, n$.

Now, taking the above three theorems into account, the *optimal allocation policy* $\boldsymbol{r}^* = (r_1^*, \ldots, r_n^*)$ should be achieved in $\bar{\mathcal{R}}_m$ as follows.

Theorem 12.3.11 (Li and Ding [287]). *Suppose $X_1 \geq_{\rm st} \cdots \geq_{\rm st} X_n$ and $Y_1, \ldots, Y_m$ are i.i.d.. For any allocation policy $\boldsymbol{r} \in \bar{\mathcal{R}}_m$, if $Y_1 \geq_{\rm st} X_1$, then, for $k = 1, \ldots, n$,*

$$T_{k|n}(\boldsymbol{r}) \leq_{\rm st} T_{k|n}(\boldsymbol{r}^*),$$

here $\boldsymbol{r}^* \in \bar{\mathcal{R}}_m$ and $|r_i^* - r_j^*| \leq 1$ for each pair of location (i, j).

$Y_1 \geq_{\rm st} X_1$, which claims that redundancies are not worse than any active component, is a bit restrictive and not always the case in practice. However, it may not be dropped off.

Example 12.3.12. Consider three components in series with

$$\bar{F}_1(t) = e^{-0.2t}, \quad \bar{F}_2(t) = e^{-0.5t}, \quad \bar{F}_3(t) = e^{-2t}$$

and three redundancies with common survival function $\bar{G}(t) = e^{-t}$, it holds that

$$X_1 \geq_{\rm st} X_2 \geq_{\rm st} Y_1 \geq_{\rm st} X_3.$$

In this setup, there are only admissible allocation policies

$$\boldsymbol{r}_1 = (0, 0, 3), \quad \boldsymbol{r}_2 = (0, 1, 2), \quad \boldsymbol{r}_3 = (1, 1, 1).$$

As is seen in Fig. 12.2, the corresponding survival curves cross with each other, and none of them is superior to the other. Due to the violation of $Y_1 \geq_{\rm st} X_1 \geq_{\rm st} \cdots \geq_{\rm st} X_3$, the optimal allocation policy does not exist. ■

As is seen, the stochastically optimal allocation policy embodies the stochastic order among the components by increasing one redundancy at some location and balance the two resulted subsets. It is still an open problem to derive the optimal allocation policy in the sense of attaining the smallest hazard rate.

12.4 Other Allocations of Active Redundancies

For coherent systems with heterogenous components, the research work in the literature mainly focuses on the following two scenarios: allocating a single redundancy to a system with heterogenous components and allocating more than one redundancy to a system with heterogenous components.

12.4.1 One Single Redundancy

Boland et al. [70] studied for the first time the allocation of one active redundancy to a k-out-of-n system with lifetimes of the components being stochastically ordered. Denote $T_{k|n}^{(i)}(\boldsymbol{X};Y)$ the lifetime of the k-out-of-n system with the redundancy being put in parallel to the ith component, $i = 1, 2, \ldots, n$. That is, $T_{k|n}^{(i)}(\boldsymbol{X};Y)$ is the $(n-k+1)$th order statistics based upon the independent random lifetimes $X_1, \ldots, X_{i-1}, X_i \vee Y, X_{i+1}, \ldots, X_n$.

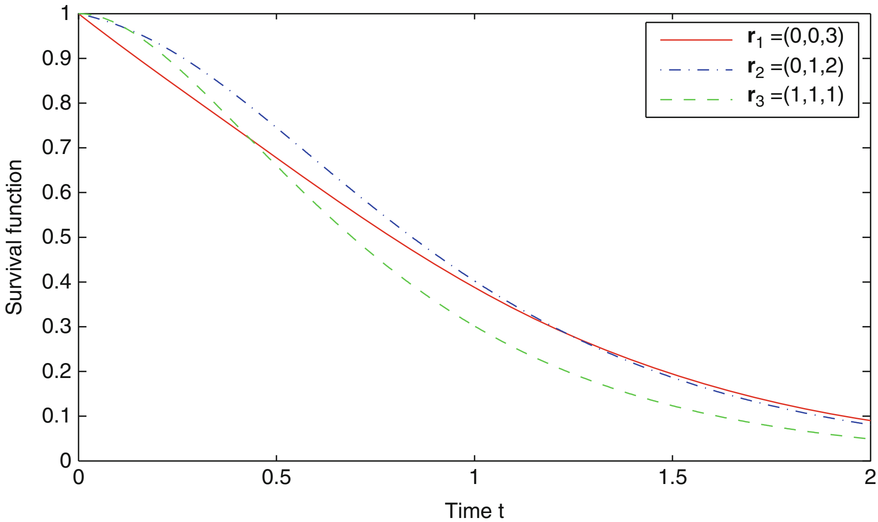


Figure 12.2: Survival curves corresponding to policies $\boldsymbol{r}_i$, $i = 1, 2, 3$

Theorem 12.4.1 (Boland et al. [70]). *Suppose* $X_1 \geq_{\rm st} X_2 \geq_{\rm st} \cdots \geq_{\rm st} X_n$. *Then,*

$$T_{k|n}^{(1)}(\boldsymbol{X};Y) \leq_{\rm st} T_{k|n}^{(2)}(\boldsymbol{X};Y) \leq_{\rm st} \cdots \leq_{\rm st} T_{k|n}^{(n)}(\boldsymbol{X};Y)$$

for $k = 1, \ldots, n$.

Subsequently, Li and Hu [288] derived the similar result with respect to the increasing concave order for the series system with two components. Here, we state the general version for series systems with more than two components without proof.

Theorem 12.4.2 (Li and Hu [288]). *Suppose* $X_1 \geq_{\rm icv} X_2 \geq_{\rm icv} \cdots \geq_{\rm icv} X_n$. *Then,*

$$T^{(1)}_{n|n}(\boldsymbol{X};Y) \leq_{\rm st} T^{(2)}_{n|n}(\boldsymbol{X};Y) \leq_{\rm st} \cdots \leq_{\rm st} T^{(n)}_{n|n}(\boldsymbol{X};Y).$$

In Li and Hu [288], it was also pointed out that the increasing concave order cannot be replaced by the increasing convex order in Theorem 12.4.2 through a counterexample of one series system of two components. On the other hand, except for the usual stochastic order and the increasing concave order, allocation policies may also be compared in the sense of other stochastic orders.

Theorem 12.4.3 (Singh and Misra [434]). *Let* X_1, X_2 *and* Y *have exponential distributions with hazard rates* λ_1, λ_2, *and* λ, *respectively. If* $\lambda_2 \geq (\lambda_1 \vee \lambda)$, *then,*

$$T^{(1)}_{2|2}(X_1, X_2; Y) \leq_{\rm hr} T^{(2)}_{2|2}(X_1, X_2; Y).$$

According to Theorems 12.4.1, 12.4.2, and 12.4.3, in order to stochastically maximize the lifetime of a k-out-of-n system, one needs to allocate the redundancy in parallel with the weakest component.

12.4.2 Several Redundancies

For a k-out-of-n system with n heterogenous components, the following results provide the optimal allocation of n heterogenous active redundancies in the sense of the usual stochastic order and the increasing concave order, respectively. The usual stochastic order is derived by Boland et al. [68] by means of redundancy importance and arrangement decreasing function.

Theorem 12.4.4 (Boland et al. [68]). *Suppose* $X_1 \geq_{\rm st} \cdots \geq_{\rm st} X_n$ *and* $Y_1 \leq_{\rm st} \cdots \leq_{\rm st} Y_n$. *Then,*

$$T_{k|n}(\boldsymbol{X} \vee \boldsymbol{Y}) \geq_{\rm st} T_{k|n}(X_1 \vee Y_{i_1}, \ldots, X_n \vee Y_{i_n}),$$

for any permutation $(i_1, \ldots, i_n)$ *of* $(1, \ldots, n)$ *and* $k = 1, \ldots, n$.

Theorem 12.4.5 (Valdès et al. [457]). *Suppose* $X_1 \geq_{\rm icv} \cdots \geq_{\rm icv} X_n$ *and* $Y_1 \leq_{\rm st} \cdots \leq_{\rm st} Y_m$ *with* $m < n$. *Then,*

$$\begin{aligned} &T_{n|n}(X_1, \ldots, X_{n-m}, X_{n-m+1} \vee Y_1, \ldots, X_n \vee Y_m) \\ &\geq_{\rm icv} T_{n|n}(X_1, \ldots, X_{n-m}, X_{n-m+1} \vee Y_{i_1}, \ldots, X_n \vee Y_{i_m}), \end{aligned}$$

for any permutation $(i_1, \ldots, i_m)$ *of* $(1, \ldots, m)$.

For series systems with two components, Valdès and Zequeira [459] derived the following hazard rate order.

Theorem 12.4.6 (Valdès and Zequeira [459]). *Assume $X_i \overset{st}{=} Y_i$ with hazard rate $\lambda_i(t)$ for $i = 1, 2$. If either of (i)-(iv) holds,*

(i) $X_1 \leq_{\mathrm{hr}} X_2$ *and* $\frac{\lambda_2(t)}{\lambda_1(t)}$ *is decreasing in* $t > 0$

(ii) $X_1 \leq_{\mathrm{hr}} X_2$ *and* $\frac{2\lambda_2(t)}{\lambda_1(t)} - 1 \leq \inf_{t>0} \frac{\lambda_2(t)}{\lambda_1(t)}$ *for all* $t > 0$

(iii) $X_1 \leq_{\mathrm{st}} X_2$ *and* $\frac{\lambda_2(t)}{\lambda_1(t)} \leq \min_{t \leq 0} \frac{1+F_2(t)}{1+F_1(t)}$ *for all* $t > 0$

(iv) $\lambda_2(t) = c\lambda_1(t)$ *for some* $c > 0$ *and all* $t > 0$

then,

$$T_{2|2}(X_1 \vee Y_1, X_2 \vee Y_2) \leq_{\mathrm{hr}} T_{2|2}(X_1 \vee Y_2, X_2 \vee Y_1).$$

In comparison with Theorem 12.4.6(i), Misra et al. [325] derived a more general sufficient condition. A similar condition was also given for the reversed hazard rate order:

Theorem 12.4.7 (Misra et al. [325]). *Assume $X_i \overset{st}{=} Y_i$ with hazard rate $\lambda_i(t)$ and reversed hazard rate $r_i(t)$, for $i = 1, 2$.*

(i) If $X_1 \leq_{\mathrm{st}} X_2$ *and* $\lambda_2(t)F_1(t) \leq \lambda_1(t)F_2(t)$ *for* $t > 0$*, then,*

$$T_{2|2}(X_1 \vee Y_1, X_2 \vee Y_2) \leq_{\mathrm{hr}} T_{n|n}(X_1 \vee Y_2, X_2 \vee Y_1)$$

(ii) If $X_1 \leq_{\mathrm{st}} X_2$ *and* $r_2(t)\bar{F}_1(t) \leq r_1(t)\bar{F}_2(t)$ *for* $t > 0$*, then,*

$$T_{2|2}(X_1 \vee Y_1, X_2 \vee Y_2) \leq_{\mathrm{rh}} T_{2|2}(X_1 \vee Y_2, X_2 \vee Y_1)$$

For a series system with two components and two active redundancies, Valdés et al. [458] considered two allocation policies: either allocate Y_1 to component X_1 or allocate Y_2 to X_2. Formally, the lifetimes of the two resulting redundant systems are

$$U_1 = (X_1 \vee Y_1) \wedge X_2 \qquad \text{and} \qquad U_2 = X_1 \wedge (X_2 \vee Y_2).$$

The two policies of allocation were compared with respect to the usual stochastic order, the hazard rate order, and the increasing concave order, respectively.

Theorem 12.4.8 (Valdés et al. [458]). *(i) If $X_1 \leq_{\rm st} X_2$ and $Y_1 \geq_{\rm st} Y_2$, then $U_1 \geq_{\rm st} U_2$*

(ii) If $X_1 \leq_{\rm st} (X_2 \wedge Y_1)$ and $X_2 \geq_{\rm st} Y_2$, then $U_1 \geq_{\rm st} U_2$

(iii) If $X_1 \leq_{\rm hr} X_2$, $X_1 \leq_{\rm hr} Y_1 \leq_{\rm hr} Y_2$, and $\frac{\lambda_2(t)}{\lambda_1(t)}$ is non-increasing, then $U_1 \geq_{\rm hr} U_2$

Theorem 12.4.9 (Valdés et al. [457]). *(i) If $X_1 \leq_{\rm icv} X_2$ and $Y_1 \geq_{\rm st} Y_2$, then $U_1 \geq_{\rm icv} U_2$*

(ii) If $X_1 \leq_{\rm icv} X_2$, $X_1 \leq_{\rm st} Y_1$ and $X_2 \geq_{\rm st} Y_2$, then $U_1 \geq_{\rm icv} U_2$

Recently, Misra et al. [324] generalized/supplemented Valdés et al. [457, 458] and presented a similar sufficient condition for the reversed hazard rate order.

Theorem 12.4.10 (Misra et al. [324]). *(i) If $X_1 \leq_{\rm rh} X_2$, $Y_1 \geq_{\rm rh} Y_2$, $X_1 \leq_{\rm rh} Y_1$ and for all $t \geq 0$, $F_1(t)G_1(t) \geq F_2(t)G_2(t)$, then $U_1 \geq_{\rm rh} U_2$*

(ii) If $X_1 \leq_{\rm st} X_2$, $Y_1 \geq_{\rm hr} Y_2$, $X_1 \leq_{\rm hr} Y_1$, and $\lambda_1(t)F_2(t) \geq \lambda_2(t)F_1(t)$ for all $t \geq 0$, then $U_1 \geq_{\rm hr} U_2$

Theorem 12.4.11 (Misra et al. [324]). *If either of (i), (ii), and (iii) holds, then $U_1 \geq_{\rm icv} U_2$:*

(i) $X_1 \leq_{\rm icv} X_2$ and $\bar{F}_2(t)\bar{G}_1(t) \geq \bar{F}_1(t)\bar{G}_2(t)$ for all $t \geq 0$

(ii) $Y_1 \leq_{\rm icv} Y_2$ and $\bar{F}_2(t)\bar{G}_1(t) \geq \bar{F}_1(t)\bar{G}_2(t)$ for all $t \geq 0$

(iii) $X_1 \leq_{\rm icv} X_2$ and $F_2(t)G_2(t) \geq F_1(t)G_1(t)$ for all $t \geq 0$

Another sufficient condition for the reversed hazard rate order was presented by Li et al. [294].

Theorem 12.4.12 (Li et al. [294]). *Suppose $X_1 \leq_{\rm rh} X_2$ and $Y_1 \geq_{\rm rh} Y_2$. Then, $U_1 \geq_{\rm rh} U_2$ if*

$$F_1(t)G_1(t)\bar{G}_2(t) \leq F_2(t)G_2(t)\bar{G}_1(t), \qquad F_1(t)G_1(t) \geq F_2(t)G_2(t),$$

for all $t \geq 0$.

Let us move to the allocation of active redundancies put forward by Mi [319]. Given $n+r$ stochastically ordered components, one wants to form the most reliable k-out-of-n system with $r \leq n$ active redundancies in the sense of stochastically maximizing the lifetime of the resulting system. The optimal structure was presented as below.

Theorem 12.4.13 (Mi [319]). *Suppose $X_1 \leq_{\mathrm{st}} \cdots \leq_{\mathrm{st}} X_{n+r}$. Then,*

$$T_{k|n}(X_1,\ldots,X_n;X_{n+1}\ldots,X_{n+r})\geq_{\mathrm{st}} T_{k|n}(X_{i_1},\ldots,X_{i_n};X_{i_{n+1}}\ldots,X_{i_{n+r}}),$$

for any permutation $(i_1,\ldots,i_{n+r})$ of $(1,\ldots,n+r)$ and $k=1,\ldots,n$.

According to Theorems 12.4.4–12.4.13, in order to make the redundant system more reliable, one may improve the weakest component by allocating an active redundancy.

12.4.3 System with Two Dependent Components

So far, all the research work is developed in the context of mutually independent components. Belzunce et al. [53] seem to be the first to take the dependence between components into consideration.

Let, for $1 \leq i < j \leq n$, the transform

$$\tau_{ij}(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = (x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n).$$

Recall that a function $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ is said to be *arrangement increasing* (AI) if

$$(x_i - x_j)[g(\boldsymbol{x}) - g(\tau_{ij}(\boldsymbol{x}))] \leq 0, \quad \text{for all } i \text{ and } j \text{ such that } 1 \leq i < j \leq n.$$

It is plain that $X_1 \leq_{\mathrm{lr}} \cdots \leq_{\mathrm{lr}} X_n$ along with mutual independence implies AI joint density. Actually, multivariate F-distribution, Marshall–Olkin multivariate exponential distribution, and multivariate Pareto distribution of type I all have an AI joint density. One may refer to Marshall et al. [312] for more details on AI functions and their applications.

Suppose X_1, X_2 are lifetimes of the two dependent active components and Y is the lifetime of the redundant one. For the series structure, the following theorem once again confirms that it is better to allocate the redundant to the weaker component.

Theorem 12.4.14 (Belzunce et al. [53], Theorem 3.2(b)). *Suppose the redundant lifetime Y is independent of the (X_1, X_2) with an AI joint density. Then,*

$$X_1 \wedge (X_2 \vee Y) \leq_{\mathrm{st}} (X_1 \vee Y) \wedge X_2.$$

The above excellent pioneering work forms one generalization of Theorem 12.4.1 with two independent components in series. Suppose the active component lifetimes $\boldsymbol{X} = (X_1, \ldots, X_n)$ has an AI joint density and is independent of the redundant lifetime Y, we conjecture that

$$(\boldsymbol{X} \vee \boldsymbol{Y})_k \leq_{\mathrm{st}} (\boldsymbol{X} \vee \tau_{ir}(\boldsymbol{Y}))_k,$$

for any $i < r$ and $k = 1, \ldots, n$.

At last, we believe, these interesting results will attract the attention from other researchers and will propel them toward more exciting research work along this line.

Acknowledgments

Xiaohu Li and Weiyong Ding are supported by National Natural Science Foundation of China (10771090)

Part IV

Stochastic Orders in Risk Analysis

Chapter 13

Dynamic Risk Measures within Discrete-Time Risk Models

Hélène Cossette and Etienne Marceau

Abstract: In this paper, we examine the capital assessment for an insurance portfolio within the classical discrete-time risk model and within two of its extensions: the classical discrete-time risk model with dependent lines of business and the classical discrete-time risk model with random income. We use finite-time ruin probabilities VaR and TVaR dynamic risk measures over a finite-time horizon. We apply results on stochastic orders to examine the riskiness of the portfolio via the dynamic TVaR. Numerical examples are provided to illustrate the topics discussed in this paper.

H. Cossette • E. Marceau (✉)
École d'Actuariat, Université Laval, Québec city, QC Canada
e-mail: helene.cossette@act.ulaval.ca; Etienne.Marceau@act.ulaval.ca

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_13,

13.1 Introduction

An appropriate assessment of the global risk of a portfolio is crucial for an insurance company. A portfolio can be a set of policies or different lines of business. Ruin theory in actuarial science has for main objective to evaluate this global risk with long-term dynamic risk models by examining the behavior of the portfolio over several periods, either on a discrete time or a continuous time basis. These models describe the evolution through time of the surplus associated to the portfolio by modeling the claims that will be paid and the premiums that will be received in the future. The global risk is generally quantified in terms of ruin measures, which are based on the event that the portfolio becomes insolvent. One important objective of ruin theory is to analyze and to compute ruin measures, especially the ruin probabilities over a finite-time and infinite-time horizon. One may use the ruin probabilities to determine the initial capital that needs to be allocated to the portfolio such that the ruin probability does not exceed a predetermined value, say, 1 %, 0.5 %, or 0.1 %. This procedure corresponds to the computation of a dynamic VaR. A general introduction to ruin theory can be found, e.g., in Gerber [178], Rolski et al. [392], or in Asmussen and Albrecher [22].

In the present paper, we examine the capital assessment for an insurance portfolio within the classical discrete time due to De Finetti [110] (see also, e.g., Bühlmann [81], Gerber [178], and Dickson [124]) and within two extensions: the classical discrete-time risk model with dependent lines of business and the classical discrete-time risk model with random income. We use the finite-time ruin probabilities defined within the classical discrete-time risk model and within both extensions to determine dynamic risk measures VaR and TVaR over a finite-time horizon. Traditionally, in discrete-time risk models, the investigation of the riskiness of an insurance portfolio is carried through the adjustment coefficient. In this paper, we apply results on stochastic orders to examine the riskiness of the portfolio via the dynamic TVaR within the classical discrete-time risk model and the two extensions previously mentioned

This paper is structured as follows. In Sect. 13.2, we briefly recall the characteristics of the classical discrete-time risk model, we define dynamic risk measures VaR and TVaR, and we use the increasing convex order to analyze the impact of the riskiness of the aggregate claim amount on the dynamic TVaR. In Sect. 13.3, we present the additional specifications of the classical discrete-time risk model with dependent lines of business and we use the supermodular order to investigate the impact of the dependence relation between the lines of business on the dynamic TVaR. In Sect. 13.4, we provide the additional definitions for the classical discrete-time risk model with random income and we use the concordance order to analyze the impact of the dependence relation between the premium income and the aggregate claim amount of the portfolio. For the three sections, numerical examples are provided to illustrate the topics discussed in this paper.

13.2 Discrete-Time Risk Model

13.2.1 Definitions

We consider a portfolio of an insurance company in the context of the classical discrete-time risk model. Let $\boldsymbol{W} = \{W_k, k \in \mathbb{N}_+\}$ be a sequence of independent and identically distributed (i.i.d.) random variables (random variables), where rv W_k is the aggregate claim amount in period $k \in \mathbb{N}_+$. In this section, we assume that W_k follows a compound distribution, meaning that it can be written as the random sum

$$W_k = \begin{cases} \sum_{j=1}^{M_k} B_{k,j}, & M_k > 0, \\ 0, & M_k = 0, \end{cases}$$

where the rv M_k is the number of claims in period k and $\{B_{k,j}, j \in \mathbb{N}_+\}$ are the individual claim amounts in period k. As it is often the case, it is assumed that $\{B_{k,j}, j \in \mathbb{N}_+\}$ is a sequence of i.i.d. positive random variables and also independent of N_k. The premium income per period is $\pi = (1 + \eta) E[W]$, with a strictly positive security margin $\eta > 0$. Furthermore, it is assumed that $\boldsymbol{W} = \{W_k, k \in \mathbb{N}_+\}$ forms a sequence of i.i.d. random variables. The classical discrete-time risk model is due to De Finetti [110] and it is a standard model in risk theory (see, e.g., Bühlmann [81], Gerber [178], and Dickson [124] for details).

Let $\boldsymbol{U} = \{U_k, k \in \mathbb{N}\}$ be the surplus process of the insurance portfolio where U_k corresponds to the surplus level at time $k \in \mathbb{N}$. For $k = 0$, $U_0 = u$ corresponds to the initial surplus and, for $k \in \mathbb{N}_+$, the expression for U_k is

$$U_k = U_{k-1} + \pi - W_k = u - \sum_{j=1}^{k} (W_j - \pi).$$

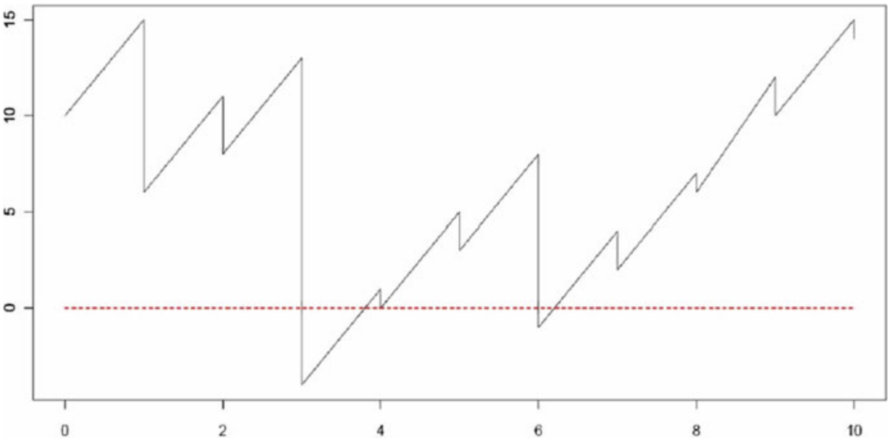


Figure 13.1: A typical sample path of the surplus process $\boldsymbol{U}$

We denote by the rv τ the time of ruin where

$$\tau = \begin{cases} \inf\limits_{k\in\mathbb{N}_+} \{k, U_k < 0\}, \text{ if } \boldsymbol{U} \text{ goes below 0 at least once} \\ \infty, \text{ if } \boldsymbol{U} \text{ never goes below 0} \end{cases}.$$

The finite-time ruin probability over n periods is given by $\psi(u,n) = P\{\tau \leq n | U_0 = u\}$.

In the sequel, we need to examine the behavior of the maximum of the random walk associated to the surplus process $\boldsymbol{U}$. Let $L_k = (W_k - \pi)$ be the net loss in period $k \in \mathbb{N}_+$. Then, L_1, L_2, ... form a sequence of i.i.d. random variables with $\mathrm{E}[L_k] = \mathrm{E}[W_k] - \pi < 0$ (since $\eta > 0$), for $k \in \mathbb{N}_+$. We define the random walk with negative drift $\boldsymbol{Y} = \{Y_k, k \in \mathbb{N}\}$ where $Y_0 = 0$ and

$$Y_k = \sum_{j=1}^{k} L_j, \qquad k \in \mathbb{N}_+.$$

The maximum net cumulative loss process associated to $\boldsymbol{Y}$ is defined by $\boldsymbol{Z} = \{Z_k, k \in \mathbb{N}\}$, where

$$Z_k = \max_{j=0,1,2,..,k} \{Y_j\}.$$

An alternative definition for the finite-time ruin probability over n periods is provided by

$$\psi(u,n) = \mathrm{P}\{\tau \leq n | U_0 = u\} = \mathrm{P}\{Z_n > u\} = 1 - F_{Z_n}(u),$$

where F_{Z_n} corresponds to the cumulative distribution function (cdf) of Z_n. The finite-time ruin probability can hence be studied through the behavior of Z_n for a fixed n. Note that, given the definition of Z_n, its distribution has a mass at 0, which is equal to $\overline{\psi}\,(0,n) = 1 - \psi\,(0,n)$.

In Figs. 13.1 and 13.2, we provide an example of a sample path of $\boldsymbol{U}$ and the corresponding sample path for $\boldsymbol{Y}$.

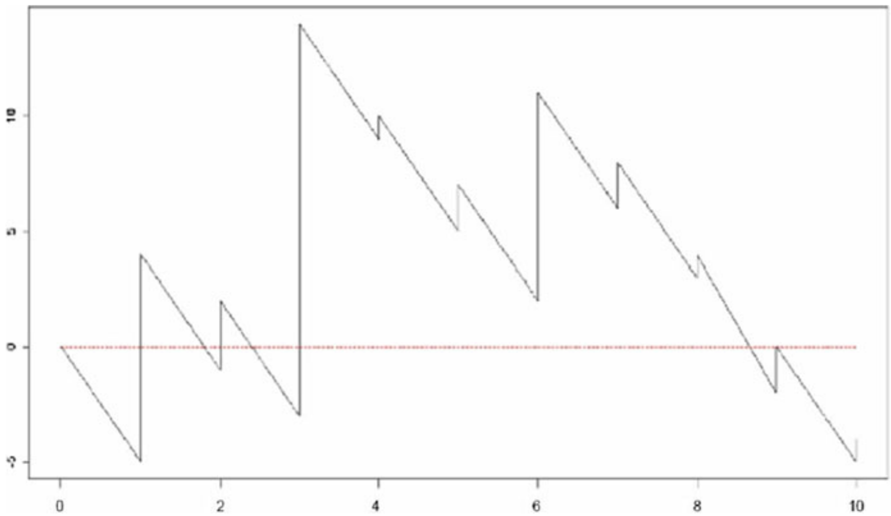


Figure 13.2: A typical sample path of the surplus process $\boldsymbol{U}$

13.2.2 Dynamic VaR and TVaR

In actuarial science and quantitative risk management, very popular risk measures are the Value-at-Risk (VaR) and the Tail-Value-at-Risk (TVaR), which are used to determine the initial capital for a fixed period of time. See, e.g., Acerbi [3], Acerbi and Tasche [4], and McNeil et al. [314] for details on the VaR and TVaR. Given this usage of such risk measures, it would be also relevant to determine the initial capital for an insurance portfolio by assessing its stochastic behavior over a finite-time horizon. Therefore, under a risk management perspective, the knowledge of the cdf F_{Z_n} of the maximum Z_n of the random walk $\boldsymbol{Y}$ can be used to determine the initial capital required. Inspired by, e.g., Trufin et al. [455], we consider two risk measures based on Z_n, namely, the dynamic VaR and the dynamic TVaR.

Definition 13.2.1. The dynamic VaR is defined by

$$VaR_\kappa(Z_n) = F_{Z_n}^{-1}(\kappa),$$

where $F_{Z_n}^{-1}$ is the inverse of F_{Z_n} defined by

$$F_X^{-1}(u) = \inf\{x \in \mathbb{R} : F_X(x) \geq u\},$$

for $u \in]0,1[$.

One may interpret $VaR_\kappa(Z_n)$ as the capital required such that the probability that the maximum net cumulative aggregate loss Z_n of the random walk $\boldsymbol{Y}$ over the first n periods exceeds that capital is equal to $1-\kappa$.

Definition 13.2.2. The dynamic TVaR is defined by

$$TVaR_\kappa(Z_n) = \frac{1}{1-\kappa}\int_{1-\kappa}^{1} VaR_v(Z_n)\mathrm{d}v$$

which is equal to

$$\frac{\mathrm{E}\left[Z_n \times 1_{\{Z_n > VaR_\kappa(Z_n)\}}\right] + VaR_\kappa(Z_n)\left(F_{Z_n}\left(VaR_\kappa(Z_n)\right) - \kappa\right)}{1-\kappa}. \tag{13.2.1}$$

Compared to the dynamic VaR, the dynamic TVaR has the advantage of being more sensitive to the stochastic behavior of Z_n in the tail of its distribution.

13.2.3 Numerical Computation of the Dynamic VaR and TVaR

To compute the values of the dynamic VaR and TVaR and to simplify the presentation, we assume that individual claim amount $B_j \in \{0, 1h, 2h, \ldots\}$ and annual premium $\pi \in \{1h, 2h, \ldots\}$. It implies that $W_k \in \{0, 1h, 2h, \ldots\}$, $k \in \mathbb{N}_+$, $L_k \in \{-\pi, -\pi + h, \ldots, -1h, 0, 1h, 2h, \ldots\}$, $k \in \mathbb{N}_+$, and $Z_n \in \{0, 1h, 2h, \ldots\}$, $n \in \mathbb{N}^+$. The cdf of Z_n is then given by

$$F_{Z_n}(kh) = 1 - \psi(kh, n) = \overline{\psi}(kh, n), \qquad k \in \mathbb{N}.$$

Let us denote the probability mass function (pmf) of W by $f_W(kh) = \mathrm{P}\{W = kh\}$, $k \in \mathbb{N}$. The values of $\overline{\psi}(kh, n)$ are computed with the following recursive relation:

$$\overline{\psi}(kh,n) = \sum_{j=0}^{k+\frac{\pi}{h}} f_W(jh)\,\overline{\psi}(kh+\pi-jh,n-1), \qquad k \in \mathbb{N}^+,$$

with $\overline{\psi}(kh,0) = 1$, $k \in \mathbb{N}$. The pmf of Z_n is given by

$$f_{Z_n}(kh) = \begin{cases} \overline{\psi}(0,n), & k = 0 \\ \overline{\psi}(kh,n) - \overline{\psi}((k-1)h,n), & k \in \mathbb{N}_+ \end{cases}. \tag{13.2.2}$$

Let $VaR_\kappa(Z_n) = k_0 h$, for $k_0 \in \mathbb{N}$. Then, the expression (13.2.1) for $TVaR_\kappa(Z_n)$ becomes

$$TVaR_\kappa(Z_n) = \frac{\mathrm{E}\left[Z_n \times 1_{\{Z_n > k_0 h\}}\right] + k_0 h\left(F_{Z_n}(k_0 h) - \kappa\right)}{1-\kappa}, \tag{13.2.3}$$

where

$$\mathrm{E}\left[Z_n \times 1_{\{Z_n > k_0 h\}}\right] = \sum_{k=k_0+1}^{\infty} kh f_{Z_n}(kh) = \mathrm{E}[Z_n] - \sum_{k=0}^{k_0} kh f_{Z_n}(kh).$$

Remark 13.2.3. If the insurance company has allocated an initial surplus of $VaR_\kappa(Z_n) = k_0 h$ ($k_0 \in \mathbb{N}$) and for some reason $Z_n > k_0 h$, then only an amount of $k_0 h$ is available for the payment of losses. On the other hand, if $Z_n > k_0 h$ but the initial capital allocated is $TVaR_\kappa(Z_n)$, then the company is in a more suitable position since it has at its disposal an amount corresponding to the average of Z_n exceeding $VaR_\kappa(Z_n)$.

The computation of both the $VaR_\kappa(Z_n)$ and $TVaR_\kappa(Z_n)$ is illustrated in the following example.

Example 13.2.4. Let W follow a compound negative binomial distribution meaning that the number of claims M has a negative binomial distribution with parameters r and q such that $\mathrm{E}[M] = r\frac{1-q}{q}$ and $f_M(k) = \binom{r+k-1}{k}(q)^r(1-q)^k$, $k \in \mathbb{N}$. Let the claim amount rv B have a geometric distribution with parameter θ such that $\mathrm{E}[B] = \frac{1}{\theta}$ and $f_B(k) = \theta(1-\theta)^{k-1}$, $k \in \mathbb{N}_+$. For $r = 2$ and 50 with $q = \frac{r}{r+2}$ such that $\mathrm{E}[M] = 2$ and for $\theta = \frac{1}{10}$, we have obtained the following values for $VaR_{0.995}(Z_{12})$ and $TVaR_{0.995}(Z_{12})$ when the net relative security loading is fixed at $\eta = 25\,\%$ (i.e. $\pi = 25$):

r	$\mathrm{E}[Z_{12}]$	$\sqrt{\mathrm{Var}(Z_{12})}$	$VaR_{0.995}(Z_{12})$	$TVaR_{0.995}(Z_{12})$
2	31.7731	44.7324	220	257.5559
50	23.0041	33.0174	163	190.5177

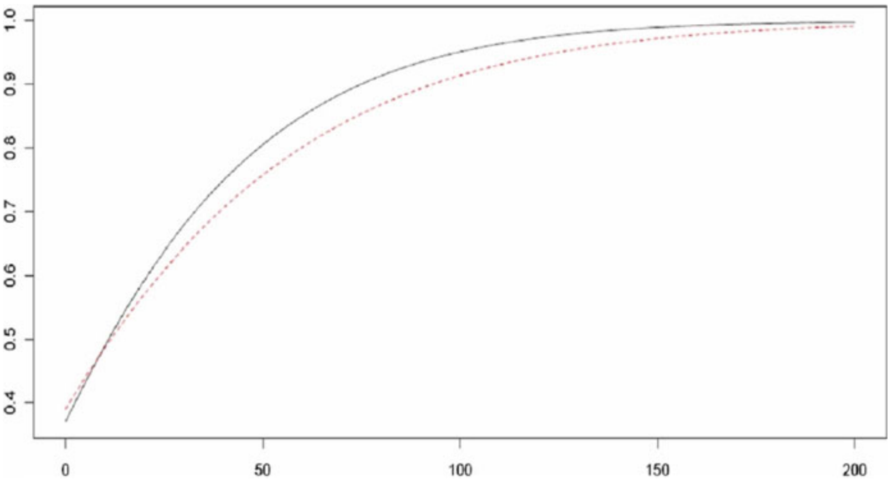


Figure 13.3: CDF of Z_{12} for $r = 2$ and $r = 50$

In Fig. 13.3, we depict the values of $F_{Z_{12}}(k)$, $k \in \mathbb{N}$, for $r = 2$ and $r = 50$. ■

13.2.4 Dynamic TVaR and Increasing Convex Order

In order to examine the impact of the definition of the aggregate claim amount W on the value of $TVaR_\kappa(Z_n)$, we briefly recall some basic definitions on stochastic orders (see, e.g., Shaked and Shanthikumar [427], Müller and Stoyan [335], or Denuit et al. [117] for further details). Univariate random variables V and V' are ordered in increasing convex order (denoted $V \leq_{\text{icx}} V'$) if $\mathrm{E}[\phi(V)] \leq \mathrm{E}[\phi(V')]$ holds for all increasing convex functions ϕ, such that the expectations exist. For example, $\phi(z) = \max(z - x; 0)$, called the stop-loss function in actuarial science, is an increasing convex function. From, e.g., Müller & Stoyan [335] and Shaked & Shanthikumar [427], $M \leq_{\text{icx}} M'$ implies that $W \leq_{\text{icx}} W'$ and $B \leq_{\text{icx}} B'$ implies that $W \leq_{\text{icx}} W'$.

In actuarial science and quantitative risk management, the increasing convex order is used to assess the riskiness of insurance risk portfolio. Within the classical discrete-time risk model, the increasing convex order is used to compare the riskiness of two portfolios via their respective adjustment coefficients. A comparison of the finite-time ruin probabilities of two portfolios and therefore their dynamic VaR cannot be made based on the increasing convex order. However, as we shall see in the following proposition, it is possible to compare the dynamic TVaR of two portfolios using the increasing convex order.

Proposition 13.2.5. *If* $W \leq_{\text{icx}} W'$*, then* $TVaR_\kappa(Z_n) \leq TVaR_\kappa(Z'_n)$ *for* $n \in \mathbb{N}$ *and for* $\kappa \in (0,1)$.

Proof: Clearly, $W \leq_{\text{icx}} W'$ implies that $L \leq_{\text{icx}} L'$. We need to define the process $\widetilde{\boldsymbol{Z}} = \left\{ \widetilde{Z}_k, k \in \mathbb{N} \right\}$ where

$$\widetilde{Z}_k = Y_k - \min_{j=0,1,2,..,k} \{Y_j\} = \max_{j=0,1,2,..,k} \{Y_k - Y_j\}.$$

Using the Lindley's recursive equation (see, e.g., Asmussen [21]), we have

$$\widetilde{Z}_k = \max\left(\widetilde{Z}_{k-1} + L_k, 0\right),$$

for $k \in \mathbb{N}_+$ and $\widetilde{Z}_0 = 0$. By duality, since $\{L_j, j \in \mathbb{N}^+\}$ is a sequence of i.i.d. random variables, it implies that $\widetilde{Z}_k \stackrel{d}{=} Z_k$, $k \in \mathbb{N}^+$. Using the results from the proofs of Theorems 4 and 5 of Müller [332], $L \leq_{\text{icx}} L'$ implies $\widetilde{Z}_n \leq_{\text{icx}} \widetilde{Z}'_n$, for $n \in \mathbb{N}$. By duality, if $\widetilde{Z}_n \leq_{\text{icx}} \widetilde{Z}'_n$ then $Z_n \leq_{\text{icx}} Z'_n$, for $n \in \mathbb{N}$. Then, from Denuit et al. [117], it follows that $Z_n \leq_{\text{icx}} Z'_n$ implies $TVaR_\kappa(Z_n) \leq TVaR_\kappa(Z'_n)$, for $n \in \mathbb{N}$. ■

In Example 13.2.4, the parameters r and q of the negative binomial distribution for the rv M are fixed such that $M \leq_{\text{icx}} M'$, which implies that $W \leq_{\text{icx}} W'$. Then, by Proposition 13.2.5, the value of $TVaR_\kappa(Z_n)$ will decrease as the value of the parameter r increases such that $\mathrm{E}[M]$ remains unchanged. If the initial capital is computed with that measure, its value will thereby decrease.

13.3 Discrete-Time Risk Model with Dependent Lines of Business

13.3.1 Additional Definitions

In this section, we consider a portfolio which has m lines of business. Several authors have examined the problem of correlated aggregate claim amounts for a portfolio of dependent lines of business. Among them, in the context of discrete-time risk models, Cossette and Marceau [95] considered the Poisson model with common shock and the negative binomial model with common component and studied the impact of the dependence relation in each model on the ruin probability. Wu & Yuen [477] have also investigated a discrete-time risk model with dependent lines of business. They investigate their model for a family of claim-number distributions, they carry numerical studies to compare finite-time ruin probabilities, and, for the infinite-time ruin probabilities, they analyze the model in terms of the adjustment coefficient.

To study this model, further definitions and assumptions must be added. Let the rv $X_{i,k}$ be the aggregate claim amount in period k for the line i ($i = 1, 2, \ldots, m$, $k \in \mathbb{N}_+$) and let the rv $W_k = \sum_{i=1}^{m} X_{i,k}$ be the aggregate claim amount in period $k \in \mathbb{N}_+$. We assume that $\{(X_{1,k}, \ldots, X_{m,k}), k \in \mathbb{N}_+\}$ forms a sequence of i.i.d. random vectors. The rv $X_{i,k}$ follows a compound distribution and

$$X_{i,k} = \begin{cases} \sum_{j=1}^{M_{i,k}} B_{i,k,j}, & M_{i,k} > 0, \\ 0, & M_{i,k} = 0, \end{cases}$$

where the rv $M_{i,k}$ is the number of claims in period k for line i and $\{B_{i,k,j}, j \in \mathbb{N}_+\}$ is a sequence of individual claim amounts in period k (positive random variables) for line i which are assumed to be i.i.d. random variables. Also, we assume that $\{B_{1,k,j}, j \in \mathbb{N}_+\}, \ldots, \{B_{n,k,j}, j \in \mathbb{N}^+\}$, and $(M_{1,k}, \ldots, M_{n,k})$ are independent. The premium income per period for the line i is $\pi_i = (1 + \eta_i) \operatorname{E}[W_i]$, $i = 1, 2, \ldots, m$ and the total premium income per period is $\pi = \sum_{i=1}^{m} \pi_i$.

We define the maximum net cumulative loss process for line i ($i = 1, 2, \ldots, m$) by $\boldsymbol{Z}^{(i)} = \left\{Z_k^{(i)}, k \in \mathbb{N}\right\}$ with $Z_0^{(i)} = 0$ and

$$Z_k^{(i)} = \max\left\{0, \sum_{l=1}^{1} (X_{i,l} - \pi_i), \ldots, \sum_{l=1}^{k} (X_{i,l} - \pi_i)\right\}, \qquad \text{for } k \in \mathbb{N}_+.$$

The risk measures VaR and TVaR associated to $Z_n^{(i)}$ are denoted $VaR_\kappa\left(Z_n^{(i)}\right)$ and $TVaR_\kappa\left(Z_n^{(i)}\right)$, $i = 1, 2, \ldots, m$. The mutualization benefit resulting from the aggregation of m lines of business is given by

$$MB_{\kappa,n}^{VaR} = \sum_{i=1}^{n} VaR_\kappa\left(Z_n^{(i)}\right) - VaR_\kappa(Z_n)$$

or

$$MB_{\kappa,n}^{TVaR} = \sum_{i=1}^{n} TVaR_\kappa\left(Z_n^{(i)}\right) - TVaR_\kappa(Z_n).$$

Example 13.3.1. The portfolio is composed of $m = 10$ lines of business. We assume that $(M_1, \ldots, M_m)$ follows a multivariate Poisson distribution with parameters $(\lambda_1, \ldots, \lambda_m, \alpha_0)$ which is defined as follows: Let K_0, K_1,..., K_n be independent random variables with $K_0 \sim Pois\,(\alpha_0)$, $0 \leq \alpha_0 \leq \min\,(\lambda_1, \ldots, \lambda_m)$, and $K_i \sim Pois\,(\alpha_i = \lambda_i - \alpha_0)$, $i = 1, 2, \ldots, m$. Let the random variables $M_1, \ldots, M_m$ be defined as

$$M_1 = K_1 + K_0, \cdots\cdots, M_m = K_m + K_0.$$

We assume here that $\lambda_1 = \ldots = \lambda_m = 0.1$. For each line $i = 1, \ldots, m$, the claim amount follows a geometric distribution as defined in Example 13.2.4 with $\theta = \frac{1}{10}$. This means that $(M_1, \ldots, M_m)$ follows a multivariate Poisson distribution. Then, by Cossette et al. [94], $W = X_1 + \ldots + X_n$ follows a compound Poisson distribution with Poisson parameter

$$\gamma_{m,\gamma_0} = \sum_{i=0}^{n} \alpha_i = m\,(0.1 - \alpha_0) + \alpha_0$$

and

$$f_{C^{(m,\gamma_0)}}\,(k) = \frac{m\,(0.1 - \alpha_0)}{\gamma_{m,\gamma_0}} f_B\,(k) + \frac{\alpha_0}{\gamma_{m,\gamma_0}} f_B^{*m}(k), \qquad k \in \mathbb{N}_+.$$

For each line $i = 1, \ldots, m$, the net relative security margin is assumed to be $\eta = 20\,\%$ which implies that the premium income per period for the whole portfolio is $\pi = 10 \times 1.2 \times 0.1 \times \frac{1}{10} = 12$. Over a finite-time horizon of $n = 4$ periods and for $\kappa = 70\,\%$, we obtain the following results:

$10\mathrm{E}\left[Z_4^{(1)}\right]$	$\sqrt{10\mathrm{Var}\left(Z_4^{(1)}\right)}$	$10VaR_\kappa\left(Z_4^{(1)}\right)$	$10TVaR_\kappa\left(Z_4^{(1)}\right)$
32.0261	25.00814	0	106.7535

α_0	$\mathrm{E}\,[Z_4]$	$\sqrt{\mathrm{Var}\,(Z_4)}$	$VaR_\kappa\,(Z_4)$	$TVaR_\kappa\,(Z_4)$
0	11.8199	18.4778	14	34.4667
0.05	20.1890	41.4363	8	66.1472
0.1	29.6237	53.3903	31	96.7878

Clearly, the values of $VaR_\kappa\,(Z_4)$ do not increase with the dependence parameter α_0 and the values of $TVaR_\kappa\,(Z_4)$ increase with the dependence parameter α_0. This result has important consequences on capital assessment. Moreover, we observe the incoherence of the risk measure VaR with the following values of $MB_{\kappa,n}^{VaR}$:

	$\alpha_0 = 0$	$\alpha_0 = 0.05$	$\alpha_0 = 0.1$
$MB_{\kappa,n}^{VaR}$	-14	-8	-31

As expected, if one uses the risk measure TVaR rather than the risk measure VaR, the coherence of this risk measure is observed (regarding the mutualization of m risks) with the following values $MB_{\kappa,n}^{TVaR}$:

	$\alpha_0 = 0$	$\alpha_0 = 0.05$	$\alpha_0 = 0.1$
$MB_{\kappa,n}^{VaR}$	72.2868	40.6064	9.9661

Also, the mutualization benefit decreases as the positive dependence relation among the lines of business becomes stronger. ■

13.3.2 Dynamic TVaR and Supermodular Order

In this subsection, we analyze the impact of the dependence between the components of $(M_1, \ldots, M_m)$ on $TVaR_\kappa(Z_n)$ with the supermodular order. Let $\boldsymbol{V} = (V_1, \ldots, V_m)$ and $\boldsymbol{V}' = (V_1', \ldots, V_m')$ be two random vectors where, for each i, V_i and V_i' have the same marginal distributions (i.e., $V_i \sim V_i'$ for $i = 1, 2, \ldots, m$). Then, $\boldsymbol{V}$ is less than $\boldsymbol{V}'$ under supermodular order, denoted $\boldsymbol{V} \leq_{\mathrm{sm}} \boldsymbol{V}'$, if $\mathrm{E}[g(\boldsymbol{V})] \leq \mathrm{E}[g(\boldsymbol{V}')]$ for all supermodular functions ϕ, given that the expectations exist. A function $\phi : \mathbb{R}^m \to \mathbb{R}$ is supermodular if

$$\begin{aligned} & \phi(x_1, \ldots, x_i+\varepsilon, \ldots, x_j+\delta, \ldots, x_m) - \phi(x_1, \ldots, x_i+\varepsilon, \ldots, x_j, \ldots, x_m) \\ \geq \; & \phi(x_1, \ldots, x_i, \ldots, x_j+\delta, \ldots, x_m) - \phi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_m) \end{aligned}$$

holds for all $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$, $1 \leq i \leq j \leq m$ and all ε, $\delta > 0$. See Marshall and Olkin [308] for examples of supermodular functions. The supermodular order is used to compare random vectors $\boldsymbol{X}$ and $\boldsymbol{X}'$ with different levels of dependence. See, e.g., Shaked and Shanthikumar [427], Müller and Stoyan [335], or Denuit et al. [117] for details on supermodular ordering.

Some authors (e.g., Cossette and Marceau [95] and Wu and Yuen [477]) have analyzed the impact of the dependence on the riskiness of a portfolio on its associated adjustment coefficient. In the following proposition, we show that we can quantify the impact of the dependence relation between the lines of business of a portfolio based on its associated TVaR. From Example 13.3.1, it is clear that such a result cannot be obtained for the dynamic VaR.

Proposition 13.3.2. *If* $(M_1, \ldots, M_m) \leq_{\text{sm}} (M'_1, \ldots, M'_m)$, *then* $TVaR_\kappa(Z_n) \leq TVaR_\kappa(Z'_n)$ *for* $n \in \mathbb{N}$.

Proof: By Denuit et al. [118], $\left(M_1, \ldots, M_m\right) \leq_{\text{sm}} \left(M'_1, \ldots, M'_m\right)$ implies $\left(X_1, \ldots, X_m\right) \leq_{\text{sm}} \left(X'_1, \ldots, X'_m\right)$. Then, Bäuerle and Müller [47] show that $(X_1, \ldots, X_m) \leq_{\text{sm}} (X'_1, \ldots, X'_m)$ implies that $W \leq_{\text{icx}} W'$. Finally, by Proposition 13.2.5, since $W \leq_{\text{icx}} W'$, we have $TVaR_\kappa(Z_n) \leq TVaR_\kappa(Z'_n)$ for $n \in \mathbb{N}$ and $\kappa \in (0,1)$. ∎

Therefore, if the dependence relation between the lines of business becomes stronger according to the supermodular order, then the amount of initial capital that one needs to set aside increases.

13.4 Discrete-Time Risk Model with Random Income

13.4.1 Additional Definitions

In this section, we assume that the premium income is random (and positive). The aggregate claim amount in period $k \in \mathbb{N}_+$ corresponds to the rv W_k, and the aggregate premium income in period $k \in \mathbb{N}_+$ is denoted by the rv P_k. Then, $\{(W_k, P_k), k \in \mathbb{N}_+\}$ forms a sequence of i.i.d. random vectors distributed as the random vector (W, P). The components of (W, P) may be dependent meaning that within a period, the premium income and the aggregate claim amount may be dependent. This dependence relation may be negative or positive. The solvency condition is expressed by $\mathrm{E}[W - P] < 0$ or equivalently $\mathrm{E}[P] > \mathrm{E}[W]$.

We need to slightly adapt the definitions provided for the classical discrete-time risk model. The surplus process is defined by $\boldsymbol{U} = \{U_k, k \in \mathbb{N}\}$ with $U_0 = u =$ initial surplus and $U_k = u - \sum_{j=1}^{k} (W_k - P_k)$ for $k \in \mathbb{N}_+$. The net loss in period k is $L_k = (W_k - P_k)$ $(k \in \mathbb{N}_+)$, where $L_1, L_2, \ldots$ form a sequence of i.i.d. random variables distributed as $L = W - P$, with $\mathrm{E}[L] < 0$.

To compute the values of the dynamic VaR and TVaR and to simplify the presentation, we assume again that $W \in \{0, 1h, 2h, \ldots\}$ and $P \in \{0, 1h, 2h, \ldots\}$. The joint pmf of (W, P) is denoted by

$$f_{W,P}(j_1 h, j_2 h) = \mathrm{P}\{W = j_1 h, P = j_2 h\},$$

where $j_1, j_2 \in \mathbb{N}$. It also implies that $Z_n \in \{0, 1h, 2h, \ldots\}$, for $n \in \mathbb{N}^+$.

The pmf of the rv $L = W - P$ is given by

$$f_L(kh) = \sum_{j=\max(0,-k)}^{\infty} f_{W,P}((k+j)h, jh), \tag{13.4.1}$$

for $k \in \mathbb{Z}$. Then, using Eq. (13.4.1), the finite-time non-ruin probabilities are recursively computed with the following expression:

$$\overline{\psi}(kh, n) = \sum_{j=-\infty}^{k} \overline{\psi}((k-j)h, n-1) f_L(jh), \tag{13.4.2}$$

for $k \in \mathbb{N}$, $n \in \mathbb{N}_+$ and with $\overline{\psi}(kh, 1) = F_L(kh)$.

We illustrate the computation of the dynamic VaR and TVaR with (13.2.2), (13.2.3), (13.4.1), and (13.4.2) in the following example.

Example 13.4.1. Let the aggregate claim amount W follow a geometric distribution with parameter $\theta_1 = \frac{1}{1.5}$ (with $\mathrm{E}[W] = 1.5$) and the premium income follow a geometric distribution with parameter $\theta_2 = \frac{1}{1.8}$ (with $\mathrm{E}[P] = 1.8$). This implies that $\mathrm{E}[L] = -0.3$. Let the joint distribution of (W, P) be define with a Frank copula

$$C_\alpha(u_1, u_2) = \frac{-1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u_1} - 1)(e^{-\alpha u_2} - 1)}{(e^{-\alpha} - 1)}\right),$$

where $\alpha \in \mathbb{R} \backslash \{0\}$ is the dependence parameter. The bivariate cumulative distribution function $F_{W,P}$ of (W, P) with marginals F_W and F_P is defined as

$$F_{W,P}(k_1 h, k_2 h) = C(F_W(k_1 h), F_P(k_2 h)), \tag{13.4.3}$$

for $(k_1, k_2) \in \mathbb{N}_+ \times \mathbb{N}_+$. The joint pmf of (W, P) is given by

$$\begin{aligned} f_{W,P}(k_1 h, k_2 h) \quad = \quad & F_{W,P}(k_1 h, k_2 h) - F_{W,P}((k_1 - 1)h, k_2 h) \\ & - F_{W,P}(k_1 h, (k_2 - 1)h) + F_{W,P}((k_1 - 1)h, (k_2 - 1)h), \end{aligned} \tag{13.4.4}$$

for $(k_1, k_2) \in \mathbb{N}_+ \times \mathbb{N}_+$, and where $f_{W,P}(k_1 h, k_2 h) = 0$, if $k_1 = 0$ or $k_2 = 0$. As mentioned in Nelsen [355], (13.4.3) is defined on the support of (W, P). Also, a given discrete bivariate distribution does not lead to a unique copula. On the other hand, copulas can be used to construct discrete bivariate distributions since this type of structure allows the coupling of various marginals. See, e.g., in Joe [211], Trivedi & Zimmer [454], and Cossette et al. [94] for examples of applications.

In their review on copulas linking discrete distributions, Genest and Nešlehovà, J. [175] mention that dependence modeling with copulas as in Eq. (13.4.3) is a valid and attractive approach for constructing bivariate distributions. Many stochastic dependence properties of a copula are inherited by the bivariate model obtained in Eq. (13.4.3). Notably, stochastic ordering relations are preserved. See Genest and Nešlehovà [175] for further details.

The values of $VaR_k(Z_n)$ for $\kappa = 99\,\%$ are provided in the following table:

α	$VaR_\kappa(Z_1)$	$VaR_\kappa(Z_4)$	$VaR_\kappa(Z_{12})$
-20	4	7	11
0	3	6	9
20	1	2	3

We have also obtained the following values of $TVaR_k(Z_n)$ for $\kappa = 99\,\%$:

α	$TVaR_\kappa(Z_1)$	$TVaR_\kappa(Z_4)$	$TVaR_\kappa(Z_{12})$
-20	4.6173	8.0880	12.8217
0	4.2077	6.9513	10.4788
20	1.6975	2.7688	3.5781

We observe that, for a fixed n, $TVaR_k(Z_n)$ decreases as $\alpha \uparrow$. It means that the amount of capital required to be set aside initially decreases as the dependence between the aggregate claim amount and the aggregate premium income goes from $-\infty$ to ∞. ■

13.4.2 Dynamic TVaR and Concordance Order

We investigate the impact of the dependence between the random variables W and P on $TVaR_\kappa(Z_n)$ with the use of the concordance order. Indeed, (W, P) is less than (W', P') under the concordance order (denoted by $(W, P) \leq_{\text{co}} (W', P')$) if

$$\mathrm{P}\{W \leq a_1, P \leq a_2\} \leq \mathrm{P}\{W' \leq a_1, P' \leq a_2\}$$

or

$$\mathrm{P}\{W > a_1, P > a_2\} \leq \mathrm{P}\{W' > a_1, P' > a_2\}$$

$\forall a_1, a_2 \in \mathbb{R}$. See, e.g., Joe [211], Müller and Stoyan [335] or Shaked and Shanthikumar [427] for details on the concordance order.

Proposition 13.4.2. *If $(W', P') \leq_{\text{co}} (W', P')$, then $TVaR_\kappa(Z_n) \leq TVaR_\kappa(Z_n')$ for $n \in \mathbb{N}$.*

Proof: By Theorem 4 of Müller [332], if $(W, P) \leq_{\text{co}} (W', P')$ then $L' = W' - P' \leq_{\text{icx}} W - P = L$. Using the same arguments as the ones in the proof of Proposition 13.2.5, we obtain the desired result. ■

Given Proposition 13.4.2, we can state that as the concordance increases between W and P, one gets a better hedge of the aggregate claims by the premiums. One therefore needs less capital as the concordance between the aggregate claim amount W and the aggregate premium income P becomes stronger.

13.5 Acknowledgements

The research was financially supported by the Natural Sciences and Engineering Research Council of Canada and the Chaire d'actuariat de l'Université Laval. We would also like to thank the participants of the International Workshop on Stochastic Orders in Reliability and Risk Management (June 27–29, 2011 | Xiamen University, Xiamen, China) for their comments.

Chapter 14
Excess Wealth Transform with Applications

Subhash Kochar and Maochao Xu

Abstract: Shaked and Shanthikumar [425] introduced the excess wealth transform and the related excess wealth order. A lot of research activities have taken place on this topic lately. In this paper, we discuss some recent developments of this transform and illustrate how to use this transform in extreme value analysis. We also summarize the applications of excess wealth order in reliability theory, auction theory, and actuarial science. Some new research directions are mentioned as well.

S. Kochar
Fariborz Maseeh Department of Mathematics and Statistics,
Portland State University, Portland, OR, USA
e-mail: kochar@pdx.edu

M. Xu (✉)
Department of Mathematics, Illinois State University, Normal, IL, USA
e-mail: mxu2@ilstu.edu

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_14,

14.1 Introduction and Motivation

The concept of variability is a basic one in statistics, probability, and many other related areas, such as reliability theory, business, economics, and actuarial science, among others. Most of the classical measurers of variability are based only upon summary statistics such as variance and standard deviation which are usually quite noninformative though they are convenient to deal with. In the past two decades, several more refined transforms and stochastic orders, which measure and compare variabilities of random variables based on their entire distribution functions, have been introduced in the literature. Shaked and Shanthikumar [427] and Müller and Stoyan [335] present comprehensive discussions on most of those concepts and their properties. Among these, the most well known one is excess wealth (ew) transform as introduced by Shaked and Shanthikumar [425]. It was also independently proposed and studied by Fernández-Ponce et al. [160] and they called it as *right spread* transform. More specifically, for a random variable X with distribution F, the excess wealth transform is defined as

$$\begin{aligned} \mathrm{W}(p\,;F) &= \mathrm{E}\left[(X - F^{-1}(p))^{+}\right] \\ &= \int_{F^{-1}(p)}^{\infty} \bar{F}(x)\mathrm{d}x, \end{aligned}$$

where $F^{-1}(p) = \inf\{x : F(x) \geq p\}$ for $0 \leq p \leq 1$, $(X)^{+} = \max\{X, 0\}$, and $\bar{F} = 1 - F$ is the survival function of X. The scaled excess wealth transform can be defined as

$$\mathrm{SEW}(p\,;F) = \frac{\mathrm{W}(p\,;F)}{\mathrm{E}(X)},$$

where $\mathrm{E}(X) < \infty$. In the context of economics, $\mathrm{W}(p\,;F)$ can be thought of as the additional wealth (on top of the pth percentile income) of the richest $100(1-p)\,\%$ individuals in the population. Based on the excess wealth transform, a stochastic partial order has also been proposed in Shaked and Shanthikumar [425] and Fernández-Ponce et al. [160]. A random variable X with distribution F is said to be smaller than another random variable Y with distribution G in the *excess wealth order* (denoted by $X \leq_{\mathrm{ew}} Y$) if

$$\mathrm{W}(p\,;F) \leq \mathrm{W}(p\,;G), \quad 0 \leq p \leq 1.$$

One may refer to Shaked and Shanthikumar [427] for a comprehensive discussion of this order.

The total time on test transform, a dual form of excess wealth transform, proposed in Barlow and Campo [35] has been widely used in engineering and statistics as a powerful data analysis tool; see, for example, Bergman [59] and Klefsjö [240]. Although a lot of work has been done in the literature on excess wealth transform, it has not been yet used much as a data analysis tool. In this paper, we move a first step toward this direction. We illustrate how to use transformed excess wealth transform to detect heavy tails in distributions. We also summarize some of the latest development on excess wealth order with emphasis on its applications in various areas, such as actuary science, reliability theory, and auction theory. It should be pointed out that the list of the topics discussed in this paper is by no means exhaustive. It only reflects our personal interests in these areas.

The organization of the paper is as follows. In the next section, we give some properties of ew transform and ew order. In Sect. 14.3, we discuss the EW plot and scaled EW plot. In Sect. 14.4, applications of ew transform/order in extreme risk analysis, reliability theory, auction theory and actuarial science are discussed. In the last section, we mention some new directions for future research on this topic.

14.2 Properties of Excess Wealth Transform and Order

It is known that the excess wealth transform is closely related to variance. Fernández-Ponce et al. [160] showed that

$$\mathrm{Var}[X] = \int_0^1 \left[\frac{\mathrm{W}(p\,;F)}{1-p}\right]^2 \mathrm{d}p. \tag{14.2.1}$$

Note that the scaled excess wealth transform is closely related to the concept of coefficient of variation $\mathrm{CV}(X) = \sqrt{\mathrm{Var}(X)}/\mathrm{E}(X)$. It is easy to see that

$$\mathrm{CV}^2(X) = \int_0^1 \left[\frac{\mathrm{SEW}(p\,;F)}{1-p}\right]^2 \mathrm{d}p.$$

The following proposition reveals that the excess wealth transform is closely related to the truncated variance.

Proposition 14.2.1. *For any* $p_0 \in [0,1)$,

$$\mathrm{Var}\left[X|X > F^{-1}(p_0)\right] = \frac{1}{1-p_0}\int_{p_0}^1 \left[\frac{\mathrm{W}(p\,;F)}{1-p}\right]^2 \mathrm{d}p.$$

Proof: Denote the distribution of $[X|X > F^{-1}(p_0)]$ by F_{p_0}, then

$$\begin{aligned}
\mathrm{W}(p\,; F_{p_0}) &= \int_{F_{p_0}^{-1}(p)}^{\infty} \bar{F}_{p_0}(x)\mathrm{d}x \\
&= \int_{F^{-1}(p^*)}^{\infty} \bar{F}(x)\mathrm{d}x \cdot \frac{1}{1-p_0} \\
&= \frac{\mathrm{W}(p^*\,; F_{p_0})}{1-p_0},
\end{aligned}$$

where $p^* = 1-(1-p)(1-p_0)$. From Eq. (14.2.1), it holds that

$$\begin{aligned}
\mathrm{Var}\left[X|X > F^{-1}(p_0)\right] &= \int_0^1 \left[\frac{\mathrm{W}(p\,; F_{p_0})}{1-p}\right]^2 \mathrm{d}p \\
&= \int_0^1 \left[\frac{\mathrm{W}(p^*\,; F)}{1-p_0}\right]^2 \frac{1}{(1-p)^2}\mathrm{d}p \\
&= \frac{1}{1-p_0}\int_{p_0}^1 \left[\frac{\mathrm{W}(p\,; F)}{1-p}\right]^2 \mathrm{d}p.
\end{aligned}$$

■

The ew transform could also be used to test nonparametric life distribution classes such as increasing failure rate (IFR), decreasing mean residual life (DMRL), and new better than used in expectation (NBUE) as pointed out in Fernández-Ponce et al. [160].

Proposition 14.2.2 (Fernández-Ponce et al. [160]).

(i) X is IFR if and only if $\mathrm{W}(p\,; F)$ *is a convex function of* p.

(ii) X is DMRL if and only if $\mathrm{W}(p\,; F)$ *is convex at* $p = 1$.

(iii) X is NBUE if and only if $\mathrm{W}(p\,; F) \leq (1-p)\mathrm{E}(X)$.

In what follows, we review some important properties of ew ordering, which are pertinent to the discussion below. One may refer to Shaked and Shanthikumar [427] for a comprehensive discussion on this order.

A desirable property of ew ordering is that it is location independent.

Proposition 14.2.3. $X \overset{\text{ew}}{=} Y \Longleftrightarrow X \overset{\text{st}}{=} Y + c$ *for some real constant* c*, where "*$\overset{st}{=}$*" represents both sides have the same distribution.*

The excess wealth transform is preserved under the increasing convex transformations.

Proposition 14.2.4. $X \leq_{\text{ew}} Y \Longrightarrow \phi(X) \leq_{\text{ew}} \phi(Y)$ *for any increasing convex function* ϕ*.*

The other desirable property of ew order shown in Shaked and Shanthikumar [425] is as follows.

Proposition 14.2.5. *Let* X *and* Y *be two continuous random variables, each having a possible mass at the origin which is assumed to be the common left endpoint of their supports. If* $X \leq_{\text{ew}} Y$*, then*

$$\mathrm{Var}[h(X)] \leq \mathrm{Var}[h(Y)] \text{ for any increasing convex function } h : [0, \infty) \to \mathbb{R},$$

for which the variances exists.

14.3 Excess Wealth Plot

As mentioned in the introduction, the ew transform, compared to the TTT (total time on test) transform, has not received that much attention as an efficient data analysis tool. The only contribution in the literature on this topic is by Belzunce et al. [55], where they mentioned the idea of using ew transform to detect NBUE (NWUE) aging property; see also Denuit et al. [119]. However, it has not been systematically studied as a data analysis tool. In this section, we introduce the excess wealth plot, which may be used as a powerful tool in data analysis.

Let $X_1, \cdots, X_n$ be a random sample from X. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the order statistics corresponding to $X_1, \cdots, X_n$. The empirical distribution is defined as

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{I}(X_i \leq t).$$

Hence, the empirical version of excess wealth transform is, for $0 \leq i \leq n-1$,

$$W_i = W\left(\frac{i}{n}; F_n\right) = \sum_{j=i}^{n-1} \int_{F_n^{-1}(j/n)}^{F_n^{-1}((j+1)/n)} \bar{F}_n(x)\mathrm{d}x$$
$$= \sum_{j=i}^{n-1} \frac{n-j}{n} \left(X_{j+1:n} - X_{j:n}\right),$$

which may be written as, for $i = 1, \ldots, n-1$,

$$W\left(\frac{i}{n}; F_n\right) = \sum_{j=i+1}^{n} \frac{n-j+1}{n} \left(X_{j:n} - X_{j-1:n}\right),$$

which is the dual form of total time on test transform ([36, 55]). It is observed that

$$W_0 = \bar{X}, \qquad W_{i+1} = W_i - \frac{n-i}{n}(X_{i+1:n} - X_{i:n}), \quad 0 \le i \le n-1. \tag{14.3.1}$$

A visual tool, called EW plot, might be constructed as follows:

(a) Order the sample: $X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$.

(b) Compute W_i as defined in Eq. (14.3.1) for $i = 0, \cdots, n-1$.

(c) Plot the pairs $(i/n, W_i)$, $i = 0, \cdots, n$, where $(1, W_n) = (1, 0)$ and connect the points by line segments.

The scaled EW plot might be constructed by replacing W_i by $W_i/\bar{X}$ in the above procedure.

Similar to the TTT transform [36], if F^{-1} is continuous, using Glivenko-Cantelli theorem, it holds that

$$W(i/n; F_n) \longrightarrow W(p; F), \quad n \to \infty, \quad i/n \to p,$$

uniformly on $[0, 1]$ with probability one. Further, we have

$$\frac{W(i/n; F_n)}{\bar{X}} \longrightarrow \frac{W(p; F)}{\mathrm{E}X}, \quad n \to \infty, \quad i/n \to p,$$

uniformly on $[0, 1]$ with probability one, which is a direct consequence of Glivenko-Cantelli theorem and strong law of large numbers.

The following example is used for illustration.

Example 14.3.1. Let X be an exponential random variable with rate 2, from which 200 samples are generated for EW plot and scaled EW (SEW) plot. The EW plot and SEW plot are displayed in Fig. 14.1.

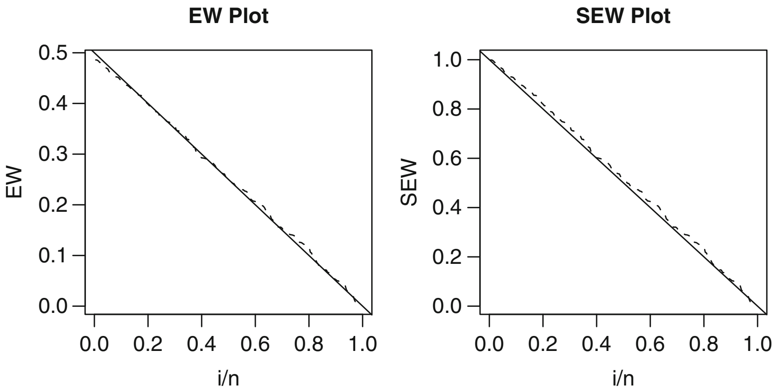


Figure 14.1: EW/SEW plots for exponential distribution, population (*solid*) and sample (*dashed*)

■

14.4 Applications

14.4.1 Extreme Risk Analysis

In extreme value theory, one popular method to study heavy-tailed data is the so-called peaks over threshold (POT). In the POT model the excess losses over high thresholds are modeled with the generalized Pareto distribution (GPD), which provides a unifying approach to the modeling of the tail of a severity distribution. For more details, one may refer to Embrechts et al. [150] and Resnick [387].

Prior to analyzing the data using extreme value theory, it is important for us to check whether the GPD is suitable for the tail. Recall that a random variable X has a GPD if it has a cumulative distribution function of the form

$$G_{\xi,\beta}(x) = \begin{cases} 1-(1+\xi x/\beta)^{-1/\xi} & \xi \neq 0, \\ 1-\exp(-x/\beta) & \xi = 0 \end{cases}$$

where $\beta > 0$ and $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\beta/\xi$ if $\xi < 0$. The parameters ξ and β are referred to as the shape and scale parameters, respectively.

It is known that the first and most important step of POT method is to select a proper threshold. The mean excess (ME) function is a tool popularly used to aid this choice of the proper threshold μ and also to determine the adequacy of the GPD model in practice (see Davison and Smith [108]).

$$M(\mu) = \mathrm{E}[X - \mu \mid X > \mu],$$

if $\mathrm{E}(X) < \infty$. It is seen that if $\xi < 1$, then

$$M(u) = \frac{\beta}{1-\xi} + \frac{\xi}{1-\xi}u.$$

Davison and Smith [108] used this property and suggested that if the ME plot is close to linear for high values of the threshold, then there is no evidence against the use of a GDP model. For more discussion on this plot in practice, please refer to Embrechts et al. [150].

One intriguing property of scaled EW transform on GPD is

$$\mathrm{SEW}(p; F) = (1-p)^{1-\xi}, \quad \xi < 1.$$

Hence, the following transformed EW transform (TEW) can be used to determine the adequacy of the GPD model in practice:

$$\mathrm{TEW}(p; F) = \frac{\ln(\mathrm{SEW}(p; F))}{\ln(1-p)} = 1 - \xi.$$

It can be shown that, if $n \to \infty, i/n \to p$,

$$\mathrm{TEW}(i/n; F_n) \longrightarrow \mathrm{TEW}(p; F)$$

uniformly on $[0, 1]$ with probability one. Hence, we may construct the following TEW plot to determine the adequacy of GPD model in practice:

$$\left(\frac{i}{n}, 1 - \mathrm{TEW}\left(\frac{i}{n}; F_n\right)\right).$$

Example 14.4.1. We generated 1,000 observations from the GPD distribution with $\beta = 0.9$ and $\xi = 0.7$, i.e.,

$$G_{\xi,\beta}(x) = 1 - (1 + .7x/.9)^{-1/.7}.$$

The TEW plot and ME plot (POT package in R) are displayed in Fig. 14.2. It is seen that the TEW plot is very informative and also provides a rough estimate of ξ around 0.7.

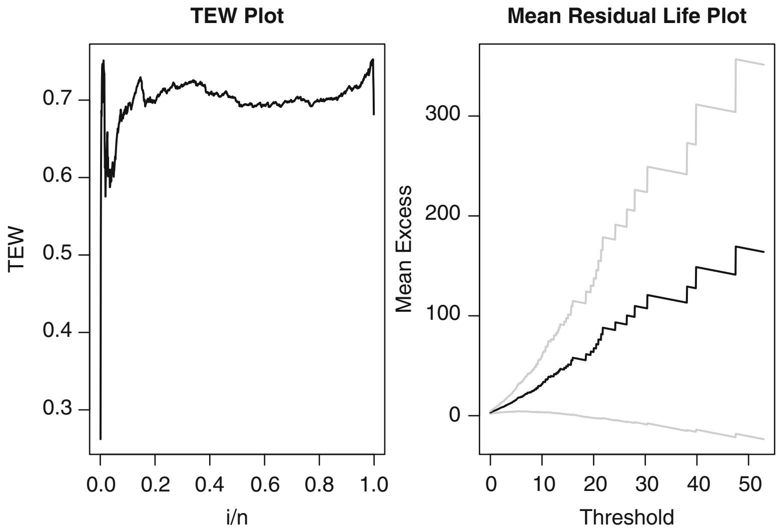


Figure 14.2: TEW plot and mean excess (ME) plot of GPD distribution with $\beta = 0.9$ and $\xi = 0.7$

■

14.4.2 Reliability Theory

Let $X_1, \ldots, X_n$ be a random sample from distribution F and let $Y_1, \ldots, Y_n$ be the other independent random sample from distribution G. For $k = 1, \ldots, n$, let $X_{k:n}$ denote the kth order statistic of the X-sample and $Y_{k:n}$ that of the Y-sample. Please refer to David and Nagaraja [107] and Balakrishnan and Rao [29, 30] for the comprehensive discussion on order statistics. In reliability theory, an n component system that works if and only if at least k of the n components work is called a k-out-of-n system. Both parallel and series systems are special cases of the k-out-of-n system. The lifetime of a k-out-of-n system can be represented as $X_{n-k+1:n}$.

Kochar et al. [250] proved the following result which states that if two distributions are ordered according to ew ordering, then the lifetimes of the parallel systems made of such i.i.d. components are also ordered according to ew ordering.

Theorem 14.4.2. $X \leq_{\text{ew}} Y \Longrightarrow X_{n:n} \leq_{\text{ew}} Y_{n:n}$

Now we consider the case when the underlying random variables are independent but not necessarily identical. Kochar and Xu [258] proved the following result when the lifetimes are exponentially distributed.

Theorem 14.4.3. *Let $X_1, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$. Let $Y_1, \ldots, Y_n$ be a random sample of size n from an exponential distribution with common hazard rate λ. Then*

$$\lambda^* \leq \lambda \Longleftrightarrow Y_{n:n} \leq_{\mathrm{ew}} X_{n:n},$$

where

$$\lambda^* = \sum_{i=1}^{n} \frac{1}{i} \left[\sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 \leq \cdots \leq i_k \leq n} \frac{1}{\sum_{j=1}^{k} \lambda_{i_j}} \right]^{-1}. \tag{14.4.1}$$

The above result was extended to the proportional hazard rate model as follows.

Theorem 14.4.4. *Let $X_1, \ldots, X_n$ be independent random variables with X_i having survival function $\bar{F}^{\lambda_i}$, $i = 1, \ldots, n$, and let $Y_1, \ldots, Y_n$ be another random sample with the common survival function $\bar{F}^{\lambda}$. If F is a DFR (decreasing failure rate) distribution, then*

$$\lambda^* \leq \lambda \Longrightarrow Y_{n:n} \leq_{\mathrm{ew}} X_{n:n},$$

where λ^* is given by Eq. (14.4.1).

A natural question is to examine whether Theorem 14.4.3 can be extended to other order statistics. While we don't know the answer for the general case, Zhao, Li, and Da [492] proved the next two results for the second order statistics.

Theorem 14.4.5. *Let $X_1, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$. Let $Y_1, \ldots, Y_n$ be a random sample of size n from an exponential distribution with common hazard rate λ. Then $Y_{2:n} \leq_{\mathrm{ew}} X_{2:n}$ if and only if*

$$\lambda \geq \lambda_U = \frac{2n-1}{n(n-1)\left(\sum_{i=1}^{n} \frac{1}{\Lambda_i} - \frac{n-1}{\Lambda(1)}\right)} \tag{14.4.2}$$

where

$$\Lambda(1) = \sum_{i}^{n} \lambda_i \text{ and } \Lambda_i = \Lambda(1) - \lambda_i.$$

Theorem 14.4.6. *Let $X_1, \ldots, X_n$ be independent random variables with X_i having survival function $\bar{F}^{\lambda_i}$, $i = 1, \ldots, n$, and let $Y_1, \ldots, Y_n$ be another random sample with the common survival function $\bar{F}^{\lambda}$. If F is a DFR distribution, then*

$$\lambda \geq \lambda_U \Longrightarrow Y_{2:n} \leq_{\mathrm{ew}} X_{2:n},$$

where λ_U is given by Eq. (14.4.2).

For comparing other k-out-of-n systems according to ew ordering, the only available results are due to Kochar and Xu [262] as given in the next two theorems.

Theorem 14.4.7. *Let $X_1, \ldots, X_p$ be independent exponential random variables with a common hazard rate λ_1, and let $X_{p+1}, \ldots, X_n$ be another set of independent and identically distributed exponential random variables with hazard rate λ_2. Let $Y_1, \ldots, Y_n$ be independent exponential random variables with a common hazard rate $\tilde{\lambda}$, satisfying the condition that*

$$\begin{aligned}\tilde{\lambda} \;\geq\; & \sum_{j=1}^{k} \frac{1}{n-j+1} \left[\sum_{j=n-k+1}^{n} (-1)^{j-n+k-1} \binom{j-1}{n-k} \right]^{-1} \cdot \\ & \left[\sum_{m \in \mathcal{M}} \binom{p}{m} \binom{n-p}{j-m} \frac{1}{m\lambda_1 + (j-m)\lambda_2} \right]^{-1}, \qquad (14.4.3)\end{aligned}$$

where $\mathcal{M} = \{m : \max\{j-n+p, 0\} \leq m \leq \min\{p, j\}\}$. Then for $1 \leq k \leq n$, the following equivalent statements hold:

(i) $X_{k:n} \geq_{\mathrm{ew}} Y_{k:n}$.

(ii) $\mathrm{E}[X_{k:n}] \geq \mathrm{E}[Y_{k:n}]$.

Theorem 14.4.8. *Let $X_1, \ldots, X_p$ be independent random variables with a common survival function $\bar{F}^{\lambda_1}$, and let $X_{p+1}, \ldots, X_n$ be another set of independent random variables with a common survival function $\bar{F}^{\lambda_2}$. Let $Y_1, \ldots, Y_n$ be a random sample from a distribution with survival function $\bar{F}^{\tilde{\lambda}}$ where $\tilde{\lambda}$ is given by Eq. (14.4.3). If F is DFR, then*

$$X_{k:n} \geq_{\mathrm{ew}} Y_{k:n}, \qquad \text{for } 1 \leq k \leq n.$$

14.4.3 Auction Theory

Let $D_{k:n} = X_{k:n} - X_{k-1:n}$ be the kth spacing from a random X-sample $X_1, \ldots, X_n$, $k = 1, \ldots, n$, and $H_{k:n} = Y_{k:n} - Y_{k-1:n}$ be the kth spacing of the random Y-sample $Y_1, \ldots, Y_n$, where $X_{0:\,n} = Y_{0:\,n} \equiv 0$. Spacings have been found in many applications in statistics, engineering, and economics, among others. One may refer to Kochar and Xu [261] and references therein. Recently, it is found that the spacings have interesting applications in auction theory. In this section, we discuss the application of ew order on the spacings in auction theory. One may refer to Krishna [267] for a comprehensive discussion of auction theory.

In a sealed auction, bidders are not aware of each other's offers, and all bidders, respectively, submit their own bids for the good. The most favorable one will be awarded the good at a price that is some function of the submitted bids. Let $X_1, \ldots, X_n$ be the valuations of the bidders. If the prices are bid in an ascending sequence by individual bidders until the highest bidder remains, and the price paid by the winner is the $(n-k+1)$-th largest price reached in the sequence, this is called a k-price buyer's auction. The *rent* of the winner, which is the difference between the largest price reached and the k-th largest price reached from bidders, can be characterized as $X_{n:n} - X_{k:n}$. If the prices are bid in a descending sequence by individual bidders until only one bidder remains, then the lowest bidder is awarded the good at a price corresponding to the k-th largest price reached in the sequence; this is called a k-price reverse auction. The *rent* of the winner is then the difference between the k-th smallest price and the smallest price from bidders, that is, $X_{k:n} - X_{1:n}$. In practice, $D_{n:n} = X_{n:n} - X_{n-1:n}$ and $D_{2:n} = X_{2:n} - X_{1:n}$ are of particular interest since they represent auction rents in buyer's auction and reverse auction in the second-price business auction.

Li [285] proved the following result.

Theorem 14.4.9. *If $X \leq_{\mathrm{ew}} Y$, then* $\mathrm{E}[D_{n:n}] \leq \mathrm{E}[H_{n:n}]$.

The above result states that in the second-price buyer's auction, if the bid is increasing in the sense of ew order, then the expected winner's rent is increasing.

This theorem was generalized in Kochar et al. [251], where they proved the following result.

Theorem 14.4.10. *If $X \leq_{\text{ew}} Y$, then $X_{n:\,n} - X_{k:\,n} \leq_{\text{icx}} Y_{n:\,n} - Y_{k:\,n}$, here "$\leq_{\text{icx}}$" means the increasing and convex order.*

Theorem 14.4.10 gives a stronger conclusion that in a k-price buyer's auction, an increase of the bid in the sense of ew order will result in an increase of the winner's rent in the sense of increasing convex order.

Now, we consider situations in which bidders are asymmetric. That is, different bidders' values are drawn from different distributions. There are only few results in the literature. The next result due to Kochar and Xu [261] gives necessary and sufficient conditions for comparing the sample spacings of two samples according to ew ordering for exponential distributions.

Theorem 14.4.11. *Let $X_1, \ldots, X_n$ be independent exponential random variables with X_i having hazard rate λ_i, $i = 1, \ldots, n$, and $Y_1, \ldots, Y_n$ be a random sample of size n from an exponential distribution with common hazard rate λ. Then, for $k \geq 2$,*

(a) $H_{k:n} \leq_{\text{ew}} D_{k:n}$ *or*

(b) $\mathrm{E}[H_{k:n}] \leq \mathrm{E}[D_{k:n}]$

if and only if

$$\frac{1}{\lambda} \leq (n-k+1) \sum_{\boldsymbol{r}} \frac{\prod_{i=1}^{n} \lambda_i}{\sum_{j=k}^{n} \lambda_{r_j} \prod_{i=1}^{n} \sum_{j=i}^{n} \lambda_{r_j}},$$

where $\boldsymbol{r}$ extends over all of the permutations of $\{1, 2, \ldots, n\}$.

14.4.4 Actuarial Science

In actuarial science, people are always interested in the following question: how much can we expect to lose with a given probability? This introduces the concept of Value-at-Risk (VaR), which has become the benchmark risk measure. For more details about VaR, please refer to Denuit et al. [117]. The VaR is defined as

$$\mathrm{VaR}[X;\, p] = F^{-1}(p).$$

As the VaR at a fixed level only gives local information about the underlying distribution, actuaries proposed the so-called *expected shortfall* to overcome this shortcoming. Expected shortfall at probability level p is the stop-loss premium with retention $\mathrm{VaR}[X;p]$, that is, $\mathrm{E}\,(X - \mathrm{VaR}[X;p])_+$, which is just the excess wealth transform of X. Hence, excess wealth order provides a natural way to compare the risks. Sordo [439] proved the following interesting result, which is closely related to Propositions 14.2.3 and 14.2.4.

Theorem 14.4.12. *Let X and Y be two random variables with respective distribution functions F and G. Then*

$$X \le_{\mathrm{ew}} Y \Longleftrightarrow H_{\phi,p}(X) \le H_{\phi,p}(Y), \quad 0 < p < 1,$$

where

$$H_{\phi,p}(X) = \mathrm{E}\left[\phi(X - \mathrm{E}(X_p))|X > F^{-1}(p)\right],$$

and ϕ is a convex function, and $X_p = (X|X > F^{-1}(p))$. $H_{\phi,p}(Y)$ is similarly defined.

As a direct consequence, the following result follows.

Corollary 14.4.13. *Let X and Y be two random variables with respective distribution functions F and G. Then*

$$X \le_{\mathrm{ew}} Y \Longrightarrow \mathrm{Var}[X_p] \le \mathrm{Var}[Y_p], \quad 0 < p < 1.$$

In fact, this result follows directly from Proposition 14.2.1.

Let $X_1, \ldots, X_n$ be independent random variables and $Y_1, \ldots, Y_n$ be another set of independent random variables. We consider the following individual risk model

$$S_X = X_1 + \ldots + X_n,$$

where X_i's are considered as risks. It is of particular interest to study the property of aggregated risk S_X.

The following result due to Hu et al. [197] states that the ew order is preserved under convolution under suitable conditions.

Theorem 14.4.14. *Let (X_i, Y_i), $i = 1, \ldots, n$ be independent pairs of random variables such that $X_i \le_{\mathrm{ew}} Y_i$, $i = 1, 2 \ldots, n$. If X_i, Y_i all have logconcave densities, except possibly one X_l and one Y_k $(l \ne k)$, then*

$$\sum_{i=1}^{n} X_i \le_{\mathrm{ew}} \sum_{i=1}^{n} Y_i.$$

In practice, the risks may not be identically distributed. Kochar and Xu [259, 263] studied the aggregated risk S_X composed of several subclaims which come from different gamma distributions. Recall that X_λ is a gamma random variable $\Gamma(a, \lambda)$ with shape parameter $a > 0$ and scale parameter $\lambda > 0$, if its density function can be written as

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} \exp\{-\lambda x\}, \quad x > 0.$$

Kochar and Xu [263] showed that more heterogeneity within the scale parameters of gamma distributions leads to increased uncertainty for the aggregated risk.

Theorem 14.4.15. *Let $X_1, \dots, X_n$ be independent gamma random variables $\Gamma(a_1, \lambda), \dots, \Gamma(a_n, \lambda)$, respectively. If $1 \le a_1 \le a_2 \le \dots \le a_n$, then*

$$(\beta_1, \cdots, \beta_n) \overset{w}{\succeq} \left(\beta_1', \cdots, \beta_n'\right) \Longrightarrow \sum_{i=1}^{n} \beta_{(i)} X_i \ge_{\text{ew}} \sum_{i=1}^{n} \beta_{(i)}' X_i,$$

where "$\overset{w}{\succeq}$" means the weak majorization [312].

The following result due to Kochar and Xu [264] provides a sufficient and necessary condition to compare heterogeneous gamma risks.

Theorem 14.4.16. *Let $X_1, \dots, X_n$ be independent gamma random variables $\Gamma(a_1, \lambda), \dots, \Gamma(a_n, \lambda)$, respectively. Then,*

$$\sum_{i=1}^{n} \beta_i X_i \ge_{\text{ew}} \beta \sum_{i=1}^{n} X_i \Longleftrightarrow \beta \le \frac{\sum_{i=1}^{n} \beta_i a_i}{\sum_{i=1}^{n} a_i}.$$

14.5 Remarks

The research on ew transform/order is far from complete. There are many ways to continue this topic. We list some topics of interest.

1. TEW plot. There are several plots used to detect heavy tails, such as Hill plot, Pickand plot, and QQ plot ([387]) in the literature. It is worth carrying out the comparison of TEW plot to the existed ones. Meanwhile, it is interesting to prove limiting results for TEW plot similar to Ghosh and Resnick [183] and Das and Ghosh [104].

2. Dependence. In the literature, the results on ew order are under the assumption that the underlying random variables are independent. However, in practice, the dependence also exists between random variables. Hence, it is of interest to study the dependence cases. For example, can we extend Theorems 14.4.3 and 14.4.15 to the dependence cases?

3. Heterogeneity. Statistics, such as order statistics, spacings, and convolutions, from heterogeneous samples have attracted more attention lately, since they are more realistic in practice. Unfortunately, there are few results on the ew order in this aspect, although it is an important and interesting topic. More research in this direction is needed. For example, can we generalize the results in Sects. 4.2 4.3 and 4.4 from exponential or gamma distributions to other general distributions?

4. Nonparametric testing. In the literature, several attempts have been made to develop nonparametric tests for the excess wealth order. Belzunce et al. [55] established L-statistics to test the right spread order based on Proposition 14.2.3

$$H_0 : X \overset{\text{ew}}{=} Y$$

vs. the alternative,

$$H_1 : X <_{\text{ew}} Y.$$

Denuit et al. [119] proposed a Kolmogorov-Smirnov-type test for the shortfall dominance against parametric alternatives, where the *shortfall order* is equivalent to the *excess wealth order* with replacing p by $1-p$. It is interesting to develop some nonparametric tests, such as Kolmogorov-Smirnov-type, Cramer-von Mises type or Anderson-Darling-type test, for ew order in two-sample cases; see, for example, [41, 119].

Part V

Applications

Chapter 15

Intermediate Tail Dependence: A Review and Some New Results

Lei Hua and Harry Joe

Abstract: The concept of intermediate tail dependence is useful if one wants to quantify the degree of positive dependence in the tails when there is no strong evidence of presence of the usual tail dependence. We first review existing studies on intermediate tail dependence, and then we report new results to supplement the review. Intermediate tail dependence for elliptical, extreme value, and Archimedean copulas is reviewed and further studied, respectively. For Archimedean copulas, we not only consider the frailty model but also the recently studied scale mixture model; for the latter, conditions leading to upper intermediate tail dependence are presented, and it provides a useful way to simulate copulas with desirable intermediate tail dependence structures.

L. Hua (✉)
Division of Statistics, Northern Illinois University, DeKalb, IL 60115, United States
e-mail: hua@math.niu.edu

H. Joe
Department of Statistics, University of British Columbia, Vancouver, BC V6T1Z4, Canada
e-mail: Harry.Joe@ubc.ca

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_15,

15.1 Introduction

For applications in many areas such as environmetrics, actuarial science, and quantitative finance, in addition to prudent examinations of univariate margins, careful modeling of various dependence patterns in corresponding distributional tails is often very important. For statistical modeling of dependence structures between random variables, a very useful approach is to employ a copula function to combine the univariate margins together to get the joint distribution. A copula $C : [0,1]^d \to [0,1]$ for a d-dimensional random vector can be defined as $C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))$, where F is the joint cumulative distribution function (cdf), F_i is the univariate cdf for the ith margin, and F_i^{-1} is the generalized inverse function defined as $F_i^{-1}(u) = \inf\{x : F_i(x) \ge u\}$. To avoid technical complexity, throughout the article, the univariate cdfs F_i's are assumed to be supported on $[0, \infty)$, and F_i's are continuous and thus the copula C is unique due to Sklar's theorem [437]. We refer the readers to [211] and [355] for references of copulas. Moreover, all distribution functions and density functions are assumed to be ultimately monotone to the left and right endpoints; this condition is very mild and satisfied by all the commonly used distributions.

The so-called tail dependence parameters (also called tail dependence coefficients or tail dependence index) have been studied as a summary quantity to capture the degree of tail dependence. Let λ be the upper tail dependence parameter, then $\lambda := \lim_{u \to 0^+} \widehat{C}(u, \ldots, u)/u$ (provided that the limit exists), where $\widehat{C}$ is the survival copula of C; that is, $\widehat{C}(u_1, \ldots, u_d) := \overline{C}(1-u_1, \ldots, 1-u_d)$, and $\overline{C}$ is the survival function of C and defined as $\overline{C}(u_1, \ldots, u_d) := 1 + \sum_{\emptyset \ne I \subseteq \{1,\ldots,d\}} (-1)^{|I|} C_I(u_i, i \in I)$, where C_I is the copula for the I-margin.

The concept of intermediate tail dependence arises when one wants to quantify the strength of such dependence in the tails but the usual tail dependence parameter $\lambda = 0$. In this case, one needs to find another quantity to capture the strength of dependence in the tails.

In [193], the concept of tail order (κ) is suggested as a quantity to capture the leading information of dependence in the tails when $\lambda = 0$. When we look at the decay of a copula function along the diagonal as a function in u, then a mild assumption is that $C(u, \ldots, u) \sim u^\kappa \ell(u)$ as $u \to 0^+$, where ℓ is a slowly varying function [61] and the notation $g \sim h$ means that the functions g and h are asymptotically equivalent; that is, $\lim_{t\to t_0} g(t)/h(t) = 1$ with t_0 being the corresponding limiting point that is usually 0 or ∞. The leading parameter κ is referred to as the *lower tail order* of copula C. In parallel, the upper tail order of copula C can be defined as the κ that satisfies that $\widehat{C}(u,\dots,u) \sim u^\kappa \ell(u), u \to 0^+$. Clearly, $\kappa = 1$ corresponds to the usual tail dependence and then the limit of the slowly varying function $\ell(u)$ is used as the quantity to capture the degree of tail dependence.

If copulas C_1 and C_2 have lower tail orders κ_1 and κ_2, respectively, and $\kappa_1 < \kappa_2$, then $C_1(\boldsymbol{u}) \geq C_2(\boldsymbol{u})$ for all $\boldsymbol{u}$ in a neighborhood of $\mathbf{0}$. Similarly, this holds for the upper tail orders in terms of survival functions. That is, the tail order for comparing copulas implies a form of multivariate stochastic order in the joint lower or upper tails.

To the best of our knowledge, [276] is the first paper that employs a regularly varying function to study the weaker dependence in the tails with the tail dependence parameter $\lambda = 0$. More specifically, for a bivariate random vector (X_1, X_2), where X_1 and X_2 are unit Fréchet distributed with cdf $F_i(x) = e^{-1/x}, x \geq 0, i = 1, 2$, and are nonnegatively associated, assume $\mathrm{P}\{X_1 > r, X_2 > r\} \sim \ell(r) r^{-1/\eta}, r \to \infty$, where $1/2 \leq \eta \leq 1$. It can be verified that the tail order κ corresponds to $1/\eta$ of Ledford and Tawn's representation, and η is called the residual dependence index in [190] and references therein. A lot of research has been done following this direction. We refer to [93, 191, 275, 276, 383] for further development of this idea.

Although the concept of tail order is defined with respect to a copula, which is a more intuitive way, the κ itself or some functional forms of κ simply describe the relative speed of decay of the joint tail probability to certain functional forms of the tail probability of one of the standardized margins. How to standardize the margins and how to choose the functional forms depend on how to make such relative speed of decay meaningful and be able to capture the leading information of dependence in the tails.

Another notion that is close to the concept of tail order is a tail dependence measure in the sense of [93], in which an upper tail dependence measure for a bivariate copula is defined as

$$\bar{\chi} := \lim_{u\to 0^+} 2\log(u)/\log(\overline{C}(1-u, 1-u)) - 1.$$

So, $\bar{\chi} = 2/\kappa_U - 1$. Note that if $C(u,\dots,u) \sim u^\kappa \ell(u)$ as $u \to 0^+$, letting $C'(u,\dots,u) := d(C(u,\dots,u))/du$ and $[u^\kappa \ell(u)]' := \mathrm{d}(u^\kappa \ell(u))/\mathrm{d}u$, then by the l'Hopital's rule and the Monotone Density Theorem [61]:

$$\begin{aligned}\lim_{u\to 0^+}\frac{\log(C(u,\ldots,u))}{\log(u)} &= \lim_{u\to 0^+}\frac{C'(u,\ldots,u)}{[u^\kappa \ell(u)]'}\times\frac{u[u^\kappa \ell(u)]'}{C(u,\ldots,u)}\\ &= \lim_{u\to 0^+}\frac{\kappa u^\kappa \ell(u)}{C(u,\ldots,u)}=\kappa. \qquad (15.1.1)\end{aligned}$$

Therefore, in some cases it may be easier to obtain the lower tail order by applying Eq. (15.1.1); for example, it will be used in the proof of Proposition 15.4.3 for a bivariate elliptical copula. Similarly, the upper tail order can be calculated as $\kappa_U = \lim_{u\to 0^+}[\log(\overline{C}(1-u,\ldots,1-u))]/[\log(u)]$.

In what follows, we will introduce the concepts of tail order and intermediate tail dependence in Sect. 15.2. Some detailed results for extreme value, elliptical, and Archimedean copulas are presented in Sects. 15.3, 15.4, and 15.5, respectively. In particular, in Sect. 15.5, we will study intermediate tail dependence through two different stochastic representations of Archimedean copulas: the frailty model and the scale mixture model. The study of upper intermediate tail dependence of Archimedean copulas derived from the scale mixture model is new in the literature. References are given for existing results, and proofs are only provided for new results. Sect. 15.6 will conclude the article.

15.2 Tail Order and Intermediate Tail Dependence

The theory of regular variation will be applied throughout the article. We refer the reader to [61], [386, 387], [113], and [173] for references. A measurable function $g:\mathbb{R}_+\to\mathbb{R}_+$ is regularly varying at ∞ with index $\alpha\neq 0$ (written $g\in \mathrm{RV}_\alpha$) if for any $t>0$, $\lim_{x\to\infty}(g(xt)/g(x)) = t^\alpha$. If the above equation holds with $\alpha=0$ for any $t>0$, then g is said to be slowly varying at ∞ and written as $g\in\mathrm{RV}_0$. For regularly varying at 0, that is, $\lim_{x\to 0^+}(g(xt)/g(x)) = t^\alpha$ for any $t>0$, the notation is $g\in\mathrm{RV}_\alpha(0^+)$, and slow variation of ℓ at 0 is written as $\ell\in\mathrm{RV}_0(0^+)$. For any $g\in\mathrm{RV}_\alpha, \alpha\in\mathbb{R}$, there exists an $\ell\in\mathrm{RV}_0$ such that $g(x)=x^\alpha\ell(x)$.

Other notation: a bold letter is used to represent a transposed vector, e.g., $\boldsymbol{x} := (x_1,\ldots,x_d)$, and $(u\mathbf{1}_d)$ represents $(u,\ldots,u)$ with d components of u's and $(u\boldsymbol{w})$ represents $(uw_1,\ldots,uw_d)$. We use $(\boldsymbol{w}_I) := (w_i; i\in I)$ when I is a subset of $\{1,\ldots,d\}$, and sometimes $(w_1,\ldots,w_d)$ is abbreviated as $(\boldsymbol{w}_d)$.

Definition 15.2.1. Suppose C is a d-dimensional copula. If there exists a real constant $\kappa_L(C) > 0$ and $\ell \in \mathrm{RV}_0(0^+)$ such that

$$C(u\mathbf{1}_d) \sim u^{\kappa_L(C)}\ell(u), \quad u \to 0^+,$$

then we refer to $\kappa_L(C)$ as the *lower tail order* of C and refer to $\lambda_L(C) = \lim_{u\to 0^+} \ell(u)$ as the *lower tail order parameter*, provided that the limit exists. Similarly, the *upper tail order* is defined as $\kappa_U(C)$ that satisfies $\overline{C}((1-u)\mathbf{1}_d) \sim u^{\kappa_U(C)}\ell(u), u \to 0^+$, with the *upper tail order parameter* being $\lambda_U(C) = \lim_{u\to 0^+} \ell(u)$, provided that the limit exists.

The notion of tail order is especially useful for bivariate copulas or for multivariate permutation-symmetric (i.e., exchangeable) copulas. Otherwise, the study of tail order involves more technical issues. All the copulas studied in this article are assumed to be permutation symmetric, in order to illustrate the main ideas and key results without involving too much technical discussion. Under such assumptions, intermediate tail dependence simply means that the corresponding tail order κ satisfies $1 < \kappa < d$.

Definition 15.2.2. Suppose C is a d-dimensional copula and $C(u\mathbf{1}_d) \sim u^{\kappa}\ell(u)$, $u \to 0^+$ for some $\ell \in \mathrm{RV}_0(0^+)$. The *lower tail order function* $b : \mathbb{R}^d_+ \to \mathbb{R}_+$ is defined as

$$b(\boldsymbol{w}; C, \kappa) = \lim_{u\to 0^+} \frac{C(uw_j, 1 \le j \le d)}{u^{\kappa}\ell(u)},$$

provided that the limit function exists. In parallel, if $\overline{C}((1-u)\mathbf{1}_d) \sim u^{\kappa}\ell(u)$, $u \to 0^+$ for some $\ell \in \mathrm{RV}_0(0^+)$, the *upper tail order function* $b^* : \mathbb{R}^d_+ \to \mathbb{R}_+$ is defined as

$$b^*(\boldsymbol{w}; C, \kappa) = \lim_{u\to 0^+} \frac{\overline{C}(1 - uw_j, 1 \le j \le d)}{u^{\kappa}\ell(u)},$$

provided that the limit function exists. If $\ell(u) \to h \neq 0$, then $hb(\boldsymbol{w}; C, 1)$ and $hb^*(\boldsymbol{w}; C, 1)$ become the tail dependence functions proposed in [213]. That is, the definition of the tail order function here absorbs a constant into ℓ so that $b(\mathbf{1}_d; C, \kappa) = 1$.

The following are some elementary properties of the lower and upper tail order functions b and b^*. Obvious properties of tail order for $\widehat{C}$ are the following: $\kappa_L(C) = \kappa_U(\widehat{C})$, $\kappa_U(C) = \kappa_L(\widehat{C})$, $b(\boldsymbol{w}; C, \kappa) = b^*(\boldsymbol{w}; \widehat{C}, \kappa)$, and $b^*(\boldsymbol{w}; C, \kappa) = b(\boldsymbol{w}; \widehat{C}, \kappa)$.

Proposition 15.2.3. *A lower tail order function $b(\boldsymbol{w}) = b(\boldsymbol{w}; C, \kappa)$ has the following properties: (1) $b(\boldsymbol{w}) = 0$ if there exists an $i \in \{1, \ldots, d\}$ with $w_i = 0$; (2) $b(\boldsymbol{w})$ is increasing in $w_i, i \in \{1, \ldots, d\}$; (3) $b(\boldsymbol{w})$ is positive homogeneous of order κ; that is, for any fixed $t > 0$, $b(t\boldsymbol{w}) = t^\kappa b(\boldsymbol{w})$; (4) if $b(\boldsymbol{w})$ is partially differentiable with respect to each w_i on $(0, +\infty)$, then $b(\boldsymbol{w}) = \frac{1}{\kappa}\sum_{j=1}^d \frac{\partial b}{\partial w_j} w_j, \forall \boldsymbol{w} \in \mathbb{R}_+^d$.*

Other properties of tail order and tail order functions can be found in [193]. Due to the limitation of space, in what follows, we will focus on the study of intermediate tail dependence for three important copula families.

15.3 Extreme Value Copula

If a copula C satisfies $C(u_1^t, \ldots, u_d^t) = C^t(u_1, \ldots, u_d)$ for any $(u_1, \ldots, u_d) \in [0, 1]^d$ and $t > 0$, then we refer to C as an extreme value copula. For any extreme value copula C, there exists a function $A : [0, \infty)^d \to [0, \infty)$ such that

$$C(u_1, \ldots, u_d) = \exp\{-A(-\log u_1, \ldots, -\log u_d)\}, \tag{15.3.1}$$

where A is convex, positive homogeneous of order 1 and satisfies

$$\max(x_1, \ldots, x_d) \le A(x_1, \ldots, x_d) \le x_1 + \cdots + x_d.$$

We refer to [375], [447], and Chap. 6 of [211] for references.

For the lower tail of an extreme value copula C, it can be verified [193] that $C(u, \ldots, u) = u^{A(1,\ldots,1)}$. That is, for any extreme value copula C, the lower tail order is $\kappa_L(C) = A(1, \ldots, 1)$, and there is intermediate lower tail dependence except for the boundary cases, such as the independence copula and the comonotonicity copula, where $A(1, \ldots, 1) = d$ and 1, respectively.

For the upper tail of an extreme value copula C, write the survival copula as $\widehat{C}(u\mathbf{1}_d) := \overline{C}((1-u)\mathbf{1}_d) = 1 + \sum_{\emptyset \ne I \subseteq \{1,\ldots,d\}} (-1)^{|I|} C_I((1 - u)\mathbf{1}_{|I|})$. Since each I-margin of C with $2 \le |I| \le d$ is also an extreme value copula [447], let A_I be the corresponding function in the sense of Eq. (15.3.1) for the extreme value copula C_I, then,

$$\widehat{C}(u, \ldots, u) = \overline{C}(1-u, \ldots, 1-u) = 1 + \sum_{\emptyset \ne I \subseteq \{1,\ldots,d\}} (-1)^{|I|} C_I(1 - \boldsymbol{u}_{|I|})$$

$$= 1 - d + du + \sum_{I \subseteq \{1,\ldots,d\}, |I| \geq 2} (-1)^{|I|} C_I(1 - \boldsymbol{u}_{|I|})$$
$$= 1 - d + du + \sum_{I \subseteq \{1,\ldots,d\}, |I| \geq 2} (-1)^{|I|} (1 - u)^{A_I(\mathbf{1}_{|I|})}.$$

By the l'Hopital's rule, as $u \to 0^+$,

$$\widehat{C}(u, \ldots, u) \sim u \left(d - \sum_{I \subseteq \{1,\ldots,d\}, |I| \geq 2} (-1)^{|I|} A_I(\mathbf{1}_{|I|}) \right) =: u\lambda.$$

So if $\lambda \neq 0$, then C has usual upper tail dependence with tail dependence parameter $\lambda = d - \sum_{I \subseteq \{1,\ldots,d\}, |I| \geq 2} (-1)^{|I|} A_I(\mathbf{1}_{|I|})$. If $\lambda = 0$, then it is unclear if an extreme value copula can have upper intermediate tail dependence or not; certain structures of those A_I functions are needed in this regard.

15.4 Elliptical Copula

Since a copula is invariant to a strict increasing transformation on margins, for the study of elliptical copula, we may omit the location and scale parameters of joint elliptical distributions. Intermediate tail dependence depends just on the radial random variable, and the condition on the radial random variable can be seen from the bivariate case, so the main result in this section is bivariate, and it can be easily extended to exchangeable multivariate elliptical copulas. Now consider the following representation: let $\boldsymbol{X} := (X_1, X_2)$ be an elliptical random vector such that

$$\boldsymbol{X} \stackrel{d}{=} RA\boldsymbol{U}, \tag{15.4.1}$$

where the radial random variable $R \geq 0$ is independent of $\boldsymbol{U}$, $\boldsymbol{U}$ is a bivariate random vector uniformly distributed on the surface of the unit hypersphere $\{\boldsymbol{z} \in \mathbb{R}^2 | \boldsymbol{z}\boldsymbol{z}^{\mathrm{T}} = 1\}$, and A is a 2×2 matrix such that $AA^{\mathrm{T}} = \Sigma$ where the entries of Σ are $\Sigma_{11} = \Sigma_{22} = 1$ and $\Sigma_{12} = \Sigma_{21} = \varrho$ with $-1 < \varrho < 1$, e.g., $A = \begin{pmatrix} 1 & 0 \\ \varrho & \sqrt{1 - \varrho^2} \end{pmatrix}$. For such an elliptical distribution, the margins have the same cdf assumed to be F.

For the usual tail dependence case, [413] proved that when the radial random variable R has a regularly varying tail, then X_1 and X_2 are tail dependent, and thus the tail order of the corresponding elliptical copula is $\kappa = 1$.

Example 15.4.1 (Student's t copula)**.** The radial random variable R for Student's t distributions is a generalized inverse Gamma distribution such that R^2 follows an inverse Gamma distribution with the shape and scale parameters being $\nu/2$, where ν is the degree of freedom. It can be verified that $\overline{F}_R \in \mathrm{RV}_{-\nu}$ (see Example 3 of [194]). So the tail order for Student's t copula is $\kappa = 1$. ■

For univariate tail heaviness, one often uses a concept referred to as maximum domain of attraction (MDA) of a univariate extreme value distribution. The following well-known result characterizes the distributions that belong to the MDA of Gumbel, which is relevant to intermediate tail dependence of elliptical copulas. We refer to [150] and [113] for more details about MDA. For notation, Λ is the Gumbel extreme value distribution for maxima, and Φ_α is the Fréchet extreme value distribution with parameter $\alpha > 0$.

Theorem 15.4.2. *A random variable X with cdf F is said to belong to the Gumbel MDA (denoted as $X \in \mathrm{MDA}(\Lambda)$ or $F \in \mathrm{MDA}(\Lambda)$) if and only if there exists a positive auxiliary function $a(\cdot)$ such that*

$$\lim_{x\to\infty} \frac{\overline{F}(x + ta(x))}{\overline{F}(x)} = e^{-t}, \quad t \in \mathbb{R}, \tag{15.4.2}$$

where $a(\cdot)$ can be chosen as $a(x) = \int_x^\infty \overline{F}(t)/\overline{F}(x)\mathrm{d}t$.

For the case where R has lighter tails than any regularly varying tails, some asymptotic study has been conducted for elliptical distributions where $R \in \mathrm{MDA}(\Lambda)$. We refer to [188], [190], and [189] for relevant references. Now we are ready to present a result that is useful to find the tail order of a bivariate elliptical copula where the radial random variable R belongs to Gumbel MDA, and this result is cited from [192]. [190] has a version of this result in Theorem 2.1 but the proof here is different.

Proposition 15.4.3. *Let C be the copula for an elliptical random vector $\boldsymbol{X} := (X_1, X_2)$ constructed as Eq. (15.4.1), and $b_\varrho = \sqrt{2/(1+\varrho)}$. If $R \in \mathrm{MDA}(\Lambda)$, then the upper and lower tail orders of C is*

$$\kappa = \lim_{r\to\infty} \frac{\log\left(1 - F_R(b_\varrho r)\right)}{\log\left(1 - F_R(r)\right)}, \tag{15.4.3}$$

provided that the limit exists.

Proof: Letting $r := F^{-1}(1-u)$ and $b_\varrho = \sqrt{2/(1+\varrho)}$, then by Example 6.2 (i) of [188], as $u \to 0^+$ and thus $r \to \infty$,

$$\begin{aligned}
&\overline{C}(1-u, 1-u) \\
&= \mathrm{P}\left[X_1 > F^{-1}(1-u), X_2 > F^{-1}(1-u)\right] = \mathrm{P}[X_1 > r, X_2 > r] \\
&= (1+o(1))\frac{(1-\varrho^2)^{3/2}}{2\pi(1-\varrho)^2}[a(b_\varrho r)/r][1-F_R(b_\varrho r)], \qquad (15.4.4)
\end{aligned}$$

where F_R is the cdf of R and $a(\cdot)$ is an auxiliary function of R with respect to the Gumbel MDA in the sense of Eq. (15.4.2). As $u \to 0^+$, i.e., $r \to \infty$, both $a(b_\varrho r)/r \to 0$ (see Theorems 3.3.26 and 3.3.27 of [150]) and $1 - F_R(b_\varrho r) \to 0$. Let $G(x) := 1/[1-F_R(x)]$, then $G : \mathbb{R} \to \mathbb{R}_+$ is increasing and the condition of Eq. (15.4.2) is equivalent to that $G \in \Gamma$-varying with auxiliary function $a(\cdot)$ [111, Definition 1.5.1]. The inverse function of a Γ-varying function is a Π-varying function [112, Corollary 1.10]. Therefore, $G^{-1} \in \Pi$-varying. Assuming that an auxiliary function of G^{-1} is $a_0(\cdot)$, by Lemma 1.2.9 of [113], the auxiliary function $a_0(\cdot)$ of the Π-varying function G^{-1} is slowly varying at ∞. Moreover, $a_0(t) = a(G^{-1}(t))$ [112, Corollary 1.10]. So, $a(x) = a_0(G(x))$. Then in Eq. (15.4.4),

$$a(b_\varrho r)/r = a_0(G(b_\varrho r))/r = a_0(1/[1-F_R(b_\varrho r)])/r,$$

while $1 - F_R(b_\varrho r)$ is rapidly varying in r at ∞ due to the fact that G is Γ-varying and any Γ-varying function is rapidly varying [111, Theorem 1.5.1]. Therefore,

$$1 - F_R(b_\varrho r) = 1 - F_R\left(\sqrt{2/(1+\varrho)}F^{-1}(1-u)\right)$$

dominates the tail behavior of Eq. (15.4.4) as $u \to 0$ and thus determines the corresponding tail order of the elliptical copula. By the definition of tail order in Definition 15.2.1, we may also obtain the upper tail order by the following:

$$\kappa = \lim_{u\to 0^+} \frac{\log \overline{C}(1-u, 1-u)}{\log u} = \lim_{r\to\infty} \frac{\log(1-F_R(b_\varrho r))}{\log(1-F(r))}.$$

By Example 6.2 (iii) of [188], as $r \to \infty$,

$$\mathrm{P}\{X_1 > r\} = (1+o(1))(2\pi)^{-1/2}[a(r)/r]^{1/2}[1-F_R(r)].$$

Due to the similar argument as before, $1 - F_R(r)$ dominates the tail behavior of $\mathrm{P}[X_1 > r]$, as $r \to \infty$. Therefore, we may write

$$\kappa = \lim_{r\to\infty} \frac{\log\left(1 - F_R(b_\varrho r)\right)}{\log\left(1 - F_R(r)\right)},$$

which completes the proof. ■

It is very convenient to apply this method to derive the tail order if we know the tail behavior of R, and R belongs to MDA of Gumbel. By Theorem 3.1 of [188], this result can also be extended to multivariate cases. For d-dimensional exchangeable elliptical copula, of which the off-diagonal entries of Σ are all ϱ and the diagonals are all 1's, the tail order is Eq. (15.4.3) where b_ϱ is replaced by $b_{\varrho,d} = \sqrt{d/[1 + (d-1)\varrho]}$.

Example 15.4.4 (Bivariate symmetric Kotz-type [158] copula)**.** The density generator

$$g(x) = Kx^{N-1}\exp\{-\beta x^\xi\}, \quad \beta, \xi, N > 0,$$

where K is a normalizing constant. By Theorem 2.9 of [158], the density function of R is $f_R(x) = 2\pi x g(x^2) = 2K\pi x^{2N-1}\exp\{-\beta x^{2\xi}\}$. So, the survival function is

$$\begin{aligned}
1 - F_R(x) &= \int_x^\infty 2K\pi t^{2N-1}\exp\{-\beta t^{2\xi}\}\mathrm{d}t \\
&= \int_{\beta x^{2\xi}}^\infty \frac{K\pi}{\xi}\beta^{-N/\xi} w^{N/\xi - 1}\exp\{-w\}\mathrm{d}w \\
&= \frac{K\pi}{\xi}\beta^{-N/\xi}\Gamma(N/\xi, \beta x^{2\xi}),\ \Gamma(\cdot,\cdot)\text{incomplete Gamma function} \\
&\sim \frac{K\pi}{\xi}\beta^{-1}x^{2N-2\xi}\exp\{-\beta x^{2\xi}\}, \quad x \to \infty,
\end{aligned}$$

where the asymptotic relation is referred to Sect. 6.5 of [2]. Then by Eq. (15.4.3), we can easily get that

$$\kappa = b_\varrho^{2\xi} = [2/(1+\varrho)]^\xi.$$

Therefore, the tail order for the symmetric Kotz-type copula is $\kappa = [2/(1+\varrho)]^\xi$. Gaussian copula belongs to this class with $\xi = 1$, so its tail order is $2/(1+\varrho)$ which is consistent to Example 1 of [193]. ■

15.5 Archimedean Copula

An Archimedean copula C has the following typical form:

$$C(u_1, \ldots, u_d) = \psi(\psi^{-1}(u_1), \cdots, \psi^{-1}(u_d)), \tag{15.5.1}$$

where ψ is referred to as the generator of the Archimedean copula C, and ψ needs to satisfy certain conditions (some papers or books also refer $\varphi = \psi^{-1}$ as the generator; e.g., [355]). Most of the commonly used Archimedean copulas correspond to ψ being a Laplace Transform (LT) of a positive random variable X [211, 212]; that is, $\psi(s) := \int_0^\infty \exp\{-st\} F_X(\mathrm{d}t)$, where F_X is the cdf of X. In this case, ψ is completely monotonic, $\psi(x)$ is decreasing in x, $\psi(0) = 1$, and $\psi(\infty) = 0$. When ψ is of such a LT, it can generate Archimedean copulas of any dimension. However, this condition is not necessary for a d-dimensional Archimedean copula with d being given and finite. A necessary and sufficient condition for finite-dimensional Archimedean copulas has been given in [301] and [315].

The above two sets of conditions on ψ correspond to two types of stochastic representations of Archimedean copulas, respectively. One is the well-known frailty model [310, 362], and the other is the recently studied scale mixture model [315] (or the resource-sharing model in the sense of [176]) for finite-dimensional Archimedean copulas. In this chapter, we will discuss conditions that lead to intermediate tail dependence of Archimedean copulas for both stochastic representations. The former has been studied in the literature [193], and the latter is only studied for the lower tail [272]. In what follows, we will first review existing studies on intermediate tail dependence for Archimedean copulas, then we will present our findings of conditions that lead to upper intermediate tail dependence of Archimedean copulas through the scale mixture model.

15.5.1 Resilience or Frailty Models

Denote

$$G_j(u) := \exp\{-\psi^{-1}(u)\}, \quad 0 \le u \le 1, \text{ for } j = 1, \ldots, d,$$

then clearly G_j's are identical univariate cdfs. Then,

$$C(u_1, \ldots, u_d) = \int_0^\infty \prod_{j=1}^d G_j^\xi(u_j)\, F_H(\mathrm{d}\xi), \tag{15.5.2}$$

where F_H is the cdf of the resilience random variable H, and

$$\psi(s) = \psi_H(s) = \int_0^\infty e^{-s\xi} F_H(\mathrm{d}\xi).$$

The mixture representation means that there are random variables $X_1, \ldots, X_d$ such that given $H = \xi$, they are conditionally independent with respective cdfs $G_1^\xi, \ldots, G_d^\xi$. Let $F_j(1-u) := \overline{G}_j(u)$, then F_j's are also cdfs, and

$$C(u_1, \ldots, u_d) = \int_0^\infty \prod_{j=1}^d \overline{F}_j^\xi(1-u_j)\, F_H(\mathrm{d}\xi).$$

Assume that F_j is the cdf of Y_j for each j. Then we can also look at the copula C as the survival copula for the vector $(Y_1, \ldots, Y_d)$ that are independent conditioning on the frailty random variable H; that is why we refer to this representation as a frailty model [176]. Actually, from different perspectives, the same random variable H can be referred to as either a resilience or a frailty random variable. We refer to [311] for more discussion about the concepts of resilience and frailty.

We now use the resilience model (15.5.2) to study how tail heaviness of H affects the tail behavior of the corresponding Archimedean copula, and we refer to [193] for relevant studies. For any given $(u_1, \ldots, u_d), 0 \le u_i \le 1$, a larger value of ξ leads to a smaller value of $C(u_1, \ldots, u_d)$, and thus a heavier right tail of H tends to generate stronger positive dependence in the upper tail. Sufficient conditions on the tail heaviness of H have been given in order to get an intermediate upper tail-dependent Archimedean copula. If we use $M_H := \sup\{m \ge 0 : \mathrm{E}[H^m] < \infty\}$ to describe the degree of tail heaviness of H, then under certain regularity conditions, $k < M_H < k+1$ with $k \in \{1, \ldots, d-1\}$ leads to upper intermediate tail dependence. The next result is presented in [193].

Proposition 15.5.1. *Suppose ψ is the LT of a positive random variable Y with $k < M_Y < k+1$ for some $k \in \{1, \ldots, d-1\}$, and $|\psi^{(k)}(0) - \psi^{(k)}(\cdot)|$ is regularly varying at 0^+ with the associated slowly varying function ℓ satisfying $\lim_{s\to 0^+} \ell(s) < \infty$. Then the Archimedean copula C_ψ has upper intermediate tail dependence, and the corresponding upper tail order is $\kappa_U = M_Y$.*

There are some upper intermediate tail-dependent Archimedean copulas that have a simple form. One example is the Archimedean copula constructed by LT [214]

$$\psi(s) = \int_s^\infty \exp\{-v^\alpha\}\mathrm{d}v/\Gamma(1+\alpha^{-1}), \quad 0 < \alpha < 1,$$

and the upper tail order is $\kappa_U = 1 + \alpha$ and the lower tail order is $\kappa_L = d^\alpha$. Another one-parameter Archimedean copula that has a very flexible upper tail dependence structure is the following:

Example 15.5.2 (Archimedean copula based on inverse Gamma LT (ACIG) [193])**.** Let $Y = X^{-1}$ and X follows Gamma$(\alpha, 1)$ with $\alpha > 0$, and then $M_Y = \alpha$ and the LT of the inverse Gamma distributed Y is

$$\psi(s;\alpha) = \frac{2}{\Gamma(\alpha)} s^{\alpha/2} K_\alpha(2\sqrt{s}), \quad s \geq 0,\ \alpha > 0, \tag{15.5.3}$$

where K_α is the modified Bessel function of the second kind. For $\alpha \in (0, +\infty)$ noninteger, $\kappa_U(C_\psi) = \max\{1, \min\{\alpha, d\}\}$ and $\kappa_L(C_\psi) = \sqrt{d}$. ■

The interesting property of ACIG is that it captures a very wide range of upper tail dependence patterns by involving only one parameter. That is, when $0 < \alpha \leq 1$, there is usual tail dependence in the upper tail; when $1 < \alpha < d$, intermediate upper tail dependence is present; when $d \leq \alpha$, the upper tail becomes tail quadrant independent (i.e., $\kappa = d$).

The condition of M_Y in Proposition 15.5.1 being noninteger seems to be unnecessary. In the next subsection, the restriction will be relaxed when we study the tail behavior of Archimedean copulas through the scale mixture model.

15.5.2 Scale Mixture Models

An Archimedean copula can also be represented as the survival copula for a random vector [315]

$$\boldsymbol{X} := (X_1, \ldots, X_d) \stackrel{d}{=} R \times (S_1, \ldots, S_d), \tag{15.5.4}$$

where R and S_i are independent for $i = 1, \ldots, d$, R is a positive random variable, and $(S_1, \ldots, S_d)$ is uniformly distributed on the simplex $\{\boldsymbol{x} \in \mathbb{R}_+^d : \sum_i x_i = 1\}$. The relationship between R in Eq. (15.5.4) and H in Eq. (15.5.2) is given in Proposition 1 of [316]; that is,

$$R \stackrel{d}{=} E_d/H,$$

where E_d is independent of H and is Erlang(d) distributed (i.e., E_d follows Gamma$(d, 1)$). The Archimedean copula can be constructed as

$$C_{\psi,d}(u_1, \ldots, u_d) := \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)), \tag{15.5.5}$$

where the generator ψ is the Williamson d-transform of cdf F_R with $F_R(0) = 0$ [475]; that is,

$$\psi(s) = \int_s^{\infty} (1 - s/r)^{d-1} F(\mathrm{d}r), \quad s \in [0, \infty).$$

Note that Williamson d-transform of a positive random variable can also lead to a generator ψ that is completely monotonic (see Example 15.5.5). Throughout this subsection, assume $\varphi(x) := \psi^{-1}(x)$. If the joint cdf of $(U_1, \ldots, U_d)$ is $C_{\psi,d}$, then

$$(X_1, \ldots, X_d) \stackrel{d}{=} (\varphi(U_1), \ldots, \varphi(U_d)) \stackrel{d}{=} R_{\psi,d} \times (S_1, \ldots, S_d);$$

that is, for each margin X_i, $\mathrm{P}\{X_i > x\} = \psi(x)$. From the proof of Theorem 1 in [272], we know that $X_i \stackrel{d}{=} RY$ for $i = 1, \ldots, d$, where $Y \sim$ Kumaraswamy$(1, d-1)$; that is,

$$F_Y(x) = 1 - (1 - x)^{d-1}, \qquad x \in [0, 1].$$

We refer to [438] for more discussion of margins of a L_p-norm uniform distribution.

Recently, [272] has studied the tail behavior of Archimedean copulas via the scale mixture representation. In Sect. 5 of [272], some tail dependence patterns have been derived except for the intermediate upper tail dependence case. We will fill the gap in this subsection.

The lower intermediate tail dependence of bivariate Archimedean copulas has been studied in Proposition 7 of [272], and the conditions needed on ψ is essentially the same as Theorem 3.3 of [84]. A more intuitive and fairly general pattern of ψ has been considered in [193]; that is, as $s \to \infty$,

$$\psi(s) \sim T(s) = a_1 s^q \exp\{-a_2 s^{1-\beta}\} \quad \text{and} \quad \psi'(s) \sim T'(s).$$

In this case, if $0 < \beta < 1$, then C_ψ has lower intermediate tail dependence with $1 < \kappa_L(C_\psi) = d^{1-\beta} < d$.

With survival copula for Eq. (15.5.4), the upper tail of R may influence the lower tail of the corresponding Archimedean copula.

Using the notation of lower tail order κ and the relation (15.1.1), the result about the lower intermediate tail dependence in [272] becomes the following:

Proposition 15.5.3. *Let C be a d-dimensional Archimedean copula constructed as the survival function of the random vector in Eq. (15.5.4) and the associated ψ in Eq. (15.5.5) is differentiable. Assume further that the radial part R in Eq. (15.5.4) satisfies $R \in \mathrm{MDA}(\Lambda)$ with the auxiliary function $a(\cdot)$ of R satisfying $a \in \mathrm{RV}_\beta$ for some $0 < \beta < 1$ and Eq. (15.1.1) holds, then $\kappa_L = d^{1-\beta}$ and thus C has lower intermediate tail dependence.*

Based on Eq. (15.5.4), the lower tail of R, or equivalently, the upper tail of $1/R$ may affect the upper tail of the associated Archimedean copula. In what follows, we will prove that $1/R$ belonging to the MDA of Fréchet (written as $1/R \in \mathrm{MDA}(\Phi_\alpha)$) may lead to upper intermediate tail dependence for the associated Archimedean copula; here the condition of the tail order being an integer is relaxed.

Proposition 15.5.4. *If $1/R \in \mathrm{MDA}(\Phi_\alpha)$ with $k \le \alpha < k+1$, where $k \in \{1, \ldots, d-1\}$ is a positive integer, and ψ is the Williamson d-transform of F_R such that $F_R(0) = 0$, and if $k = d-2$ or $d-1$, then further require that ψ is $(d+1)$-monotone and $(d+2)$-monotone, respectively, then the Archimedean copula constructed by ψ as Eq. (15.5.5) has upper tail order $\kappa_U = \alpha$ and thus upper intermediate tail dependence if $1 < \alpha < d$.*

Proof: First note that either ψ being $(d+1)$-monotone or $(d+2)$-monotone can imply ψ being d-monotone [315]. So we can still apply the Williamson d-transform for the two cases where $k = d-2$ or $d-1$. By definition, the Williamson d-transform of R is

$$
\begin{aligned}
\psi(s) &= \int_s^\infty (1 - s/r)^{d-1} F_R(\mathrm{d}r) \\
&= \int_s^\infty \sum_{i=0}^{d-1} \binom{d-1}{i} (-s/r)^i F_R(\mathrm{d}r) \\
&= 1 - F_R(s) + \sum_{i=1}^{d-1} \binom{d-1}{i} (-1)^i s^i \int_s^\infty r^{-i} F_R(\mathrm{d}r) \\
&= 1 - F_R(s) + \sum_{i=1}^{d-1} \binom{d-1}{i} (-1)^i s^i \left(-s^{-i} F_R(s) + i \int_s^\infty F_R(r) r^{-i-1} \mathrm{d}r \right)
\end{aligned}
$$

$$= 1 - F_R(s) + \sum_{i=1}^{d-1} \binom{d-1}{i} (-1)^i \left(-F_R(s) + is^i \int_0^{1/s} F_R(1/y) y^{i-1} \mathrm{d}y \right)$$
$$= 1 + \sum_{i=1}^{d-1} \binom{d-1}{i} (-1)^i \left(is^i \int_0^{1/s} F_R(1/y) y^{i-1} \mathrm{d}y \right)$$
$$=: 1 + \sum_{i=1}^{d-1} \binom{d-1}{i} (-1)^i m_i(s). \tag{15.5.6}$$

For any $1 \le \alpha < d$, there exists a positive integer $k \in \{1, 2, \ldots, d-1\}$ such that

$$k \le \alpha < k + 1.$$

To study the upper tail for an Archimedean copula, we now investigate the behavior of the functions $m_i(s)$ in Eq. (15.5.6) for $i = k, \ldots, d-1$, as $s \to 0^+$. Depending on the value of i, we consider the following cases:

Case 1: $\alpha < k + 1 \le i \le d - 1$.

The condition $1/R \in \mathrm{MDA}(\Phi_\alpha)$ implies that $\mathrm{P}\{1/R > \cdot\} \in \mathrm{RV}_{-\alpha}$, and $F_R = \mathrm{P}\{R \le \cdot\} \in \mathrm{RV}_\alpha(0^+)$. Write $F_R(s) := s^\alpha \ell_R(s)$, where $\ell_R \in \mathrm{RV}_0(0^+)$. Since $y \mapsto F_R(1/y) \in \mathrm{RV}_{-\alpha}$, we have $y \mapsto F_R(1/y)y^{i-1} \in \mathrm{RV}_{i-\alpha-1}$. By Karamata's theorem (e.g., [387]), $i \ge k + 1 > \alpha$ implies that

$$\int_0^{1/s} F_R(1/y) y^{i-1} \mathrm{d}y \sim \frac{1}{i-\alpha} F_R(s) s^{-i}, \quad s \to 0^+,$$

and thus

$$m_i(s) := is^i \int_0^{1/s} F_R(1/y) y^{i-1} \mathrm{d}y \sim \frac{i}{i-\alpha} F_R(s) = \frac{i}{i-\alpha} s^\alpha \ell_R(s), \quad s \to 0^+. \tag{15.5.7}$$

Case 2: $1 \le i < \alpha$. The condition $F_R \in \mathrm{RV}_\alpha(0^+)$ implies that

$$\ell_i(s) := i \int_0^{1/s} F_R(1/y) y^{i-1} \mathrm{d}y \nearrow \mathrm{E}[R^{-i}] < \infty, \quad \text{as } s \to 0^+.$$

Therefore, $m_i(s) \sim \mathrm{E}[R^{-i}] s^i$ as $s \to 0^+$.

Case 3: $i = k = \alpha$. Let $\ell_k(s) := k \int_0^{1/s} F_R(1/y) y^{k-1} \mathrm{d}y$. Similar to the derivation of Eq. (15.5.7), by Karamata's theorem (e.g., Theorem 2.1 (a) of [387]), $\ell_k \in \mathrm{RV}_0(0^+)$, and hence, $m_k \in \mathrm{RV}_k(0^+)$.

Then $\psi(s) = 1 + \sum_{i=1}^{d-1} \binom{d-1}{i}(-1)^i m_i(s)$ implies that

if α is a positive noninteger, that is, $k < \alpha$, then $\ell_i(s) \to \mathrm{E}[R^{-i}]$ for $i = 1, \ldots, k$, and

$$\begin{aligned}\psi(s) \sim 1 &+ \sum_{i=1}^{k} \binom{d-1}{i} (-1)^i s^i \ell_i(s) \\ &+ \sum_{i=k+1}^{d-1} \binom{d-1}{i} (-1)^i \frac{i}{i-\alpha} s^\alpha \ell_R(s), \quad s \to 0^+; \qquad (15.5.8)\end{aligned}$$

if α is a positive integer, that is, $k = \alpha$, then $\ell_i(s) \to \mathrm{E}[R^{-i}]$ for $i = 1, \ldots, k-1$, and

$$\begin{aligned}\psi(s) \sim 1 &+ \mathbf{1}\{k > 1\} \times \sum_{i=1}^{(k-1)\vee 1} \binom{d-1}{i} (-1)^i s^i \ell_i(s) \\ &+ \binom{d-1}{k} (-1)^k s^k \ell_k(s) \\ &+ \sum_{i=k+1}^{d-1} \binom{d-1}{i} (-1)^i \frac{i}{i-\alpha} s^\alpha \ell_R(s), \quad s \to 0^+. \qquad (15.5.9)\end{aligned}$$

Therefore, $1 \le \alpha$ implies that the map $1 - \psi(\cdot) \in \mathrm{RV}_1(0^+)$. Write $1 - \psi(s) = s\ell_0(s)$, where $\ell_0 \in \mathrm{RV}_0(0^+)$.

Assume that there exists a constant κ and a slowly varying function $\ell_* \in \mathrm{RV}_0(0^+)$ such that $\overline{C}(1-u, \ldots, 1-u) \sim u^\kappa \ell_*(u)$ as $u \to 0^+$. If we can prove that $\kappa = \alpha$, get the expression of ℓ_*, and prove that such an ℓ^* is a slowly varying function, then the proof is finished.

Let $s := \psi^{-1}(1-u)$, then $1 - \psi(s) = s\ell_0(s)$ implies that

$$\begin{aligned}1 &= \lim_{u \to 0^+} \frac{\overline{C}(1-u, \ldots, 1-u)}{u^\kappa \ell_*(u)} \\ &= \lim_{u \to 0^+} \frac{1 + \sum_{j=1}^{d} (-1)^j \binom{d}{j} \psi(j\psi^{-1}(1-u))}{u^\kappa \ell_*(u)} \\ &= \frac{\lim_{s \to 0^+} \left\{1 + \sum_{j=1}^{d} (-1)^j \binom{d}{j} \psi(js)\right\}}{s^\kappa \ell_0^\kappa(s) \ell_*(1-\psi(s))}.\end{aligned}$$

The locally uniform convergence of ℓ_* at 0^+, together with $1 - \psi(\cdot) \in \mathrm{RV}_1(0^+)$, implies that, for any given $t > 0$,

$$\lim_{s\to0^+}\frac{\ell_*(1-\psi(ts))}{\ell_*(1-\psi(s))}=\lim_{s\to0^+}\frac{\ell_*(\frac{1-\psi(ts)}{1-\psi(s)}\times(1-\psi(s))}{\ell_*(1-\psi(s))}=\lim_{u\to0^+}\frac{\ell_*(tu)}{\ell_*(u)}=1.$$

Hence, $\ell_*(1-\psi(\cdot))\in \mathrm{RV}_0(0^+)$. Let $\ell(s):=\ell_0^\kappa(s)\ell_*(1-\psi(s))$, then $\ell\in\mathrm{RV}_0(0^+)$, due to Proposition 1.3.6 of [61].

Applying the Monotone Density Theorem [61, Theorem 1.7.2b] $k-1$ times on Eq. (15.5.6) implies that $\psi^{(i)}(0)<\infty$ for $i=0,\dots,k-1$. Moreover, choosing $w_i\equiv 1$ for each i in Lemma 2 of [193] implies that $\sum_{j=1}^d(-1)^j\binom{d}{j}j^i\equiv 0$ for any positive integer i that is less than d. Therefore, $\sum_{j=1}^d(-1)^j\binom{d}{j}j^i\psi^{(i)}(0)=0$ for $i=0,\dots,k-1$. Define $[x]:=\max\{z \text{ integer}; z<x\}$, and for any $y>0$, $y!:=y\times(y-1)\times\cdots\times(y-[y])$. By the l'Hopital's rule,

$$1=\lim_{s\to0^+}\frac{\sum_{j=1}^d(-1)^j\binom{d}{j}j^{k-1}\psi^{(k-1)}(js)}{[\kappa!/(\kappa-k+1)!]s^{\kappa-k+1}\ell(s)}. \tag{15.5.10}$$

If α is a noninteger, that is, if $k<\alpha<k+1$, then applying the Monotone Density Theorem two more times for Eq. (15.5.10) with respect to Eq. (15.5.8) leads to

$$(-1)^{k+1}\psi^{(k+1)}(s)\sim\left|\sum_{i=k+1}^{d-1}\binom{d-1}{i}(-1)^i\frac{i}{i-\alpha}\right|[\alpha!/(\alpha-k-1)!]s^{\alpha-k-1}\ell_R(s); \tag{15.5.11}$$

$$\begin{aligned}1&=\lim_{s\to0^+}\frac{\sum_{j=1}^d(-1)^j\binom{d}{j}j^k\psi^{(k)}(js)}{[\kappa!/(\kappa-k)!]s^{\kappa-k}\ell(s)}\\&=\lim_{s\to0^+}\frac{\sum_{j=1}^d(-1)^{j-k-1}\binom{d}{j}j^{k+1}[(-1)^{k+1}\psi^{(k+1)}(js)]}{[\kappa!/(\kappa-k-1)!]s^{\kappa-k-1}\ell(s)}\end{aligned} \tag{15.5.12}$$

Combining Eqs.(15.5.11) and (15.5.12) leads to $\kappa=\alpha$, and then

$$\ell(s)=\sum_{j=1}^d(-1)^{j-k-1}\binom{d}{j}j^\alpha\left|\sum_{i=k+1}^{d-1}\binom{d-1}{i}(-1)^i\frac{i}{i-\alpha}\right|\ell_R(js),$$

which is a slowly varying function. Thus, the slowly varying function ℓ^* can be chosen accordingly, which proves the case where α is a positive noninteger.

If α is a positive integer, that is, $\alpha=k$, then applying the Monotone Density Theorem one more time for Eq. (15.5.10) with respect to Eq. (15.5.9) leads to $\kappa=k=\alpha$, and similarly, the slowly varying function ℓ^* can be obtained accordingly. ∎

Example 15.5.5 (ACIG copula). For the ACIG copula studied in Example 4 of [193], $H^{-1} \sim \text{Gamma}(\alpha, 1)$ for $\alpha > 0$. Therefore, R has the same distribution as the product of two independent Gamma random variables with scale parameter 1 and respective shape parameters d and α. The product follows a K-distribution and the density function is

$$\begin{aligned} f_R(x; d, \alpha) &= \frac{2}{\Gamma(d)\Gamma(\alpha)} x^{(\alpha+d)/2-1} K_{d-\alpha}(2\sqrt{x}), \\ & \qquad x \in [0, \infty); d \text{ positive integer}; \alpha > 0, \end{aligned} \tag{15.5.13}$$

where K is the modified Bessel function of the second kind. We refer to [206] for the reference of density functions of a K-distribution. Because the left tail behavior of R affects the upper tail dependence pattern of the corresponding Archimedean copula derived from Eq. (15.5.4), we need to study the behavior of Eq. (15.5.13) at 0. If α is not an integer, then, as $s \to 0^+$,

$$K_{d-\alpha}(s) \sim \frac{1}{2}\left(\Gamma(d-\alpha)(s/2)^{\alpha-d} + \Gamma(\alpha-d)(s/2)^{d-\alpha}\right).$$

Therefore, as $x \to 0^+$,

$$f_R(x; d, \alpha) \sim \frac{\Gamma(d-\alpha)}{\Gamma(d)\Gamma(\alpha)} x^{\alpha-1} + \frac{\Gamma(\alpha-d)}{\Gamma(d)\Gamma(\alpha)} x^{d-1}.$$

If $1 < \alpha < d$, then the term $x^{\alpha-1}$ dominates the tail behavior at 0; this is an upper intermediate tail dependence case. When $d - \alpha$ is a positive integer, then by [2],

$$\begin{aligned} K_{d-\alpha}(s) &\sim \tfrac{1}{2}(s/2)^{\alpha-d} \sum_{k=0}^{d-\alpha-1} \frac{(-1)^k (d-\alpha-k-1)!}{k!} (s/2)^{2k} \\ &\sim \frac{(d-\alpha-1)!}{2}(s/2)^{\alpha-d}, \quad s \to 0^+. \end{aligned}$$

Therefore,

$$f_R(x; d, \alpha) \sim \frac{(d-\alpha-1)!}{\Gamma(d)\Gamma(\alpha)} x^{\alpha-1} = \frac{\Gamma(d-\alpha)}{\Gamma(d)\Gamma(\alpha)} x^{\alpha-1}, \quad x \to 0^+,$$

which is the same as α being a noninteger. Combining these two cases, $1 < \alpha < d$ implies upper intermediate tail dependence of the ACIG copula. ∎

Proposition 15.5.4 is actually very useful to guide us to find a scale mixture random vector whose survival copula is an upper intermediate tail-dependent Archimedean copula. For example, let R follow the positive Weibull distribution; that is, $F_R(x) = 1 - \exp\{-x^\alpha\} \sim x^\alpha, x \to 0^+$ with $1 < \alpha < d$. Then Proposition 15.5.4 implies that $\kappa_U = \alpha$ and thus the corresponding Archimedean copula has upper intermediate tail dependence. Another example of Archimedean copula that has upper intermediate tail dependence is presented in the following:

Example 15.5.6 (Dagum-simplex mixture)**.** The Dagum distribution is also referred to as an inverse Burr distribution, and it is a special case of generalized beta distribution of the second kind (e.g., [241]). Let the cdf of the radial random variable R be Dagum, then

$$F_R(x) = \left[1 + (x/\sigma)^{-\alpha}\right]^{-\beta}, \quad x > 0, \quad \alpha, \beta, \sigma > 0.$$

We choose $\sigma = 1$ for simulation as the scale parameter does not affect the associated Archimedean copula. It can be derived that $F_R \in \mathrm{RV}_{\alpha\beta}(0^+)$. By Proposition 15.5.4, if $1 < \alpha\beta < 2$, then the copula $C(u, v) := \psi(\psi^{-1}(u) + \psi^{-1}(v))$ should have upper tail order $\kappa_U = \alpha\beta$. The simulated scatter plots are illustrated in Fig. 15.1, where the left plot is for uniform margins and the right plot is for standard normal margins. The sample size was $2,000$ for the simulations.

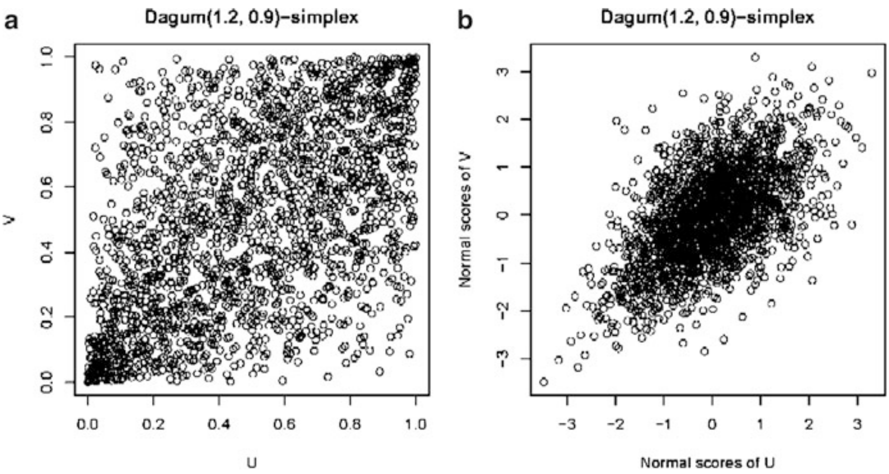
In Fig. 15.1, the upper tail order is $\kappa_U = 1.08$, which belongs to upper intermediate tail dependence. For the lower tail, $\overline{F}_R(x) = 1 - (1 + x^{-\alpha})^{-\beta}$, and as $x \to \infty$,

$$\overline{F}_R^{(1)}(x) = -\alpha\beta(1 + x^{-\alpha})^{-\beta-1}x^{-\alpha-1} \sim -\alpha\beta x^{-\alpha-1}.$$

So $\overline{F}_R \in \mathrm{RV}_{-\alpha}$. By Corollary 2 of [272], there is lower tail dependence and the tail order parameter $\lambda_L = 2^{-\alpha}$. ■

15.6 Remark and Future Work

The notion of tail order provides a quantity to evaluate the degree of dependence in the tails of joint distributions, especially when intermediate tail dependence appears. We first review fundamental concepts and existing results of intermediate tail dependence. Throughout the review, some new properties of intermediate tail dependence have been given to supplement existing results. The new results mainly consist



Dagum $(1.2, 0.9)$ Dagum $(1.2, 0.9)$

Figure 15.1: Simulation of Dagum-simplex copula (**a**) Dagum $(1.2, 0.9)$ (**b**) Dagum $(1.2, 0.9)$

of an easy way to derive the tail order of a bivariate intermediate tail-dependent elliptical copula, and the study of intermediate upper tail dependence for Archimedean copulas constructed from a scale mixture model.

Proposition 15.5.4 is helpful for constructing an upper intermediate tail-dependent Archimedean copula. However, the scale mixture approach can often only give us a simple way to simulate desired tail dependence structures, but not a simple closed-form parametric copula family. So, how to apply the scale mixture model to provide various desirable models and to make statistical inference efficiently will be a very interesting topic for future research.

As mentioned earlier in Sect. 15.1, there is a link between tail orders and multivariate stochastic orders. Future research also includes whether certain forms of multivariate stochastic orders [426, Chap. 6] can be adapted to tail forms for comparing the strength of dependence in the tails of copulas.

Chapter 16
Second-Order Conditions of Regular Variation and Drees-Type Inequalities

Tiantian Mao

Abstract: There are a variety of concepts extending regular variation, among which are the extended regular variation (ERV), second-order regular variation (2RV), and second-order extended regular variation (2ERV). In this paper, we reexamine the connections from 2ERV to 2RV and recover and strengthen the main result in Neves (2009, [358]) by using a different but straightforward approach. We also present new Drees-type inequalities in which the original auxiliary functions of a 2ERV or ERV function are not replaced by other ones.

T. Mao (✉)
Department of Statistics and Finance, School of Management, University of Science and Technology of China, Hefei, Anhui 230026, China
e-mail: mttiy@mail.ustc.edu.cn

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_16,

16.1 Introduction

Regular variation (RV) has become one of the key notions which appears in a natural way in applied probability, statistics, risk management, telecommunications networks, and other fields. There are a variety of concepts extending RV, among which are the *extended regular variation* (ERV), *second-order regular variation* (2RV) , and *second-order extended regular variation* (2ERV). Here, RV and ERV are termed as the first-order conditions, and 2RV and 2ERV are termed as the second-order conditions. The second-order conditions can be used to study the speed of convergence of certain estimators in the extreme value theory. Standard references on RV and its different extensions are given by [61, 113, 114, 194, 387].

A measurable function $h : \mathbb{R}_+ \to \mathbb{R}$ that is eventually positive is said to be of RV at infinity with index $\alpha \in \mathbb{R}\backslash\{0\}$, denoted by $h \in \mathrm{RV}_\alpha$, if for any $x > 0$,

$$\lim_{t\to\infty} \frac{h(tx)}{h(t)} = x^\alpha. \tag{16.1.1}$$

If Eq. (16.1.1) holds with $\alpha = 0$ for any $x > 0$, then h is said to be slowly varying at infinity and denoted by $h \in \mathrm{RV}_0$. $h : \mathbb{R}_+ \to \mathbb{R}$ is said to be of ERV, denoted by $h \in \mathrm{ERV}_\gamma$, if there exists a function $a : \mathbb{R}_+ \to \mathbb{R}_+$ such that for some $\gamma \in \mathbb{R}$ and all $x > 0$,

$$\lim_{t\to\infty} \frac{h(tx) - h(t)}{a(t)} = \frac{x^\gamma - 1}{\gamma}, \tag{16.1.2}$$

where, for $\gamma = 0$, the right hand side in Eq. (16.1.2) is interpreted as $\log x$. The function a is called an auxiliary function for h. $h \in \mathrm{ERV}_0$ is also written as $h \in \Pi$.

In order to specify the inherent rate of the convergence in Eqs. (16.1.1) and (16.1.2), second-order conditions are needed. A measurable function $h : \mathbb{R}_+ \to \mathbb{R}$ that is eventually positive is said to be of 2RV with the first-order parameter $\gamma \in \mathbb{R}$ and the second-order parameter $\rho \le 0$, denoted by $h \in 2\mathrm{RV}_{\gamma,\rho}$, if there exists some ultimately positive or negative function $A(t)$ with $A(t) \to 0$ as $t \to \infty$ such that

$$\lim_{t\to\infty} \frac{h(tx)/h(t) - x^\gamma}{A(t)} = x^\gamma \int_1^x u^{\rho-1}\mathrm{d}u, \quad \forall\, x > 0. \tag{16.1.3}$$

Here, $A(t)$ is referred to as an auxiliary function of h, and ρ governs the speed of convergence in Eq. (16.1.1). $h : \mathbb{R}_+ \to \mathbb{R}$ is said to be of 2ERV with the first-order parameter $\gamma \in \mathbb{R}$ and the second-order parameter $\rho \leq 0$, denoted by $h \in 2\mathrm{ERV}_{\gamma,\rho}$, if there exists some positive function $a(t)$ and some ultimately positive or negative function $A(t)$ with $A(t) \to 0$ as $t \to \infty$ such that

$$\lim_{t\to\infty} \frac{\frac{h(tx)-h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} = H_{\gamma,\rho}(x) \tag{16.1.4}$$

with

$$H_{\gamma,\rho}(x) = \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} \mathrm{d}u \mathrm{d}s, \quad \forall\, x > 0.$$

Here, $a(t)$ and $A(t)$ are referred to as first-order and second-order auxiliary functions of h, respectively, and ρ governs the speed of convergence in Eq. (16.1.2).

There are closed connections between different notions of RV. For example, ERV can be characterized by RV (see Theorems B.2.2 and B.2.12 in [113]), and 2ERV can be characterized by ERV (see Theorems B.3.6 in [113] or Theorem 2 in [114]). More recently, Fraga Alves et al. (2007, [163]) and Neves (2009, [358]) established connections from 2ERV to 2RV.

Drees (1998, [135]) and Cheng and Jiang (2001, [88]) established uniform bounds for the class of RV [resp. ERV, 2RV, and 2ERV] functions in the sense that the difference between the terms in both sides of Eq. (16.1.1) [resp. (16.1.2), Eq. (16.1.3), and Eq. (16.1.4)] is uniformly bounded under appropriately chosen auxiliary functions a_0 and/or A_0 instead of auxiliary functions a and/or A. For example, if $h \in \mathrm{ERV}_\gamma$ with auxiliary function a and $\gamma \in \mathbb{R}$, then for all $\varepsilon, \delta > 0$, there exists $t_0 = t_0(\varepsilon, \delta)$ such that for any $t, tx > t_0$,

$$\left| \frac{h(tx) - h(t)}{a_0(t)} - \frac{x^\gamma - 1}{\gamma} \right| \leq \varepsilon x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}, \tag{16.1.5}$$

where

$$a_0(t) = \begin{cases} \gamma h(t), & \gamma > 0; \\ h(t) - t^{-1}\int_0^t h(s)\mathrm{d}s, & \gamma = 0; \\ -\gamma(h(\infty) - h(t)), & \gamma < 0. \end{cases} \tag{16.1.6}$$

Here, $h(\infty) := \lim_{t\to\infty} h(t)$ is finite when $\gamma < 0$. If $h \in 2\mathrm{ERV}_{\gamma,\rho}$ with first-order and second-order auxiliary functions a and A, respectively, $\gamma \in \mathbb{R}$, $\rho \leq 0$, then for all $\varepsilon, \delta > 0$, there exists $t_0 = t_0(\varepsilon, \delta)$ such that for any $t, tx > t_0$,

$$\left| \frac{\frac{h(tx)-h(t)}{a_0(t)} - \frac{x^\gamma-1}{\gamma}}{A_0(t)} - H_{\gamma,\rho}(x) \right| \le \varepsilon x^{\gamma+\rho} \max\left\{x^\delta, x^{-\delta}\right\}, \qquad (16.1.7)$$

where a_0 and A_0 are chosen specifically, taking complicated forms such that $A_0(t) \sim A(t)$ and $a_0(t)/a(t) - 1 = o(A(t))$ as $t \to \infty$. These kinds of inequalities are referred to as Drees-type inequalities.

The purposes of this paper are twofold. First, we reexamine the connections from 2ERV to 2RV and recover and strengthen the main result in [163, 358] by using a different but straightforward approach. Second, we present new Drees-type inequalities, similar to Eqs. (16.1.5) and (16.1.7), in which the original auxiliary functions a and A are not replaced by a_0 and A_0, respectively. These kinds of inequalities may have potential applications.

Throughout, the notation "$g(t) \sim h(t)$, $t \to t_0$" means asymptotic equivalence; that is, $\lim_{t\to t_0} g(t)/h(t) = 1$. For any increasing function h, define its generalized inverse $h^{\leftarrow}$ by

$$h^{\leftarrow}(x) = \inf\{t : h(t) \ge x\}, \quad \forall\, x.$$

16.2 Connections Between 2ERV and 2RV

First, note that $H_{\gamma,\rho}$ in Eq. (16.1.4) can be written as

$$H_{\gamma,\rho}(x) = \begin{cases} \frac{1}{\rho}\left(\frac{x^{\gamma+\rho}-1}{\gamma+\rho} - \frac{x^\gamma-1}{\gamma}\right), & \rho \ne 0; \\ \frac{1}{\gamma}\left(x^\gamma \log x - \frac{x^\gamma-1}{\gamma}\right), & \rho = 0, \gamma \ne 0; \\ \frac{1}{2}(\log x)^2, & \gamma = \rho = 0. \end{cases} \qquad (16.2.1)$$

Also,

- If a satisfies (16.1.2) for $\gamma \in \mathbb{R}$, then $a \in \mathrm{RV}_\gamma$; see Theorem B.2.1 in [113].
- If the functions a and A satisfy (16.1.4), then $|A| \in \mathrm{RV}_\rho$ and $a \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function A; see Theorem B.3.1 in [113].
- For $\gamma \in \mathbb{R}$, $h \in \mathrm{ERV}_\gamma$ if and only if $\tilde{h} \in \mathrm{RV}_\gamma$, where

$$\tilde{h}(t) = \begin{cases} h(t), & \gamma > 0; \\ h(t) - \frac{1}{t}\int_{t_0}^t h(s)\mathrm{d}s, & \gamma = 0; \\ h(\infty) - h(t), & \gamma < 0, \end{cases}$$

  for any fixed $t > 0$; see Theorem B.2.2 in [113].

- For $\gamma \in \mathbb{R}$ and $\rho \le 0$, $h \in 2\mathrm{RV}_{\gamma,\rho}$ if and only if $t^{-\gamma}h(t) \in \mathrm{ERV}_\rho$. In particular, for $\rho \le 0$, $h \in 2\mathrm{RV}_{0,\rho}$ if and only if $h \in \mathrm{ERV}_\rho$.

Obviously, $h \in 2\mathrm{RV}_{\gamma,\rho}$ implies $h \in 2\mathrm{ERV}_{\gamma,\rho}$ for $\gamma \neq 0$ and $\gamma+\rho \neq 0$. While, for $\gamma = 0$ or $\gamma+\rho = 0$, $h \in 2\mathrm{RV}_{\gamma,\rho}$ does not imply $h \in 2\mathrm{ERV}_{\gamma,\rho}$. In the next result, we will establish some other connections from ERV to 2RV. First, we give the following useful lemma, whose proof follows from Theorem 3.6.6 in [61].

Lemma 16.2.1. *For $\gamma \in \mathbb{R}$ and $\rho < 0$, $h \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function $A(t)$ if and only if there exists a constant $c > 0$ such that*

$$h(t) = ct^\gamma \left[1 + \frac{1}{\rho}A(t) + o(A(t))\right], \quad t \to \infty. \tag{16.2.2}$$

Proposition 16.2.2. *Let $h \in 2\mathrm{ERV}_{\gamma,\rho}$ with $\gamma \neq 0$, $\rho \le 0$ and with first-order and second-order auxiliary functions $a(t)$ and $A(t)$, respectively, i.e., Eq. (16.1.4) holds. Then $\tilde{h} \in 2\mathrm{RV}_{\gamma,\rho}$ with an auxiliary function B, where the functions $\tilde{h}$ and B are given as follows according to different cases:*

(i) *For $\rho \le 0$ and $\rho + \gamma > 0$,*

$$\tilde{h}(t) = h(t) \quad \text{and} \quad B(t) = \frac{\gamma}{\rho+\gamma}A(t).$$

(ii) *For $\rho \le 0$ and $\rho + \gamma < 0$,*

$$\tilde{h}(t) = -\frac{\gamma}{|\gamma|}(w_\infty - h(t)) \quad \text{and} \quad B(t) = \frac{\gamma}{\rho+\gamma}A(t),$$

where $w_\infty = h(\infty) \in \mathbb{R}$ for $\rho = 0$, $\lim_{t\to\infty} t^{-\gamma}a(t) = c \in (0,\infty)$ and

$$w_\infty := \lim_{t\to\infty}\left[h(t) - c\frac{t^\gamma}{\gamma}\right] = \lim_{t\to\infty}\left[h(t) - \frac{a(t)}{\gamma}\right] \tag{16.2.3}$$

exist and are finite for $\rho < 0$.

(iii) *For $\rho < 0$ and $\rho + \gamma = 0$, w_∞ defined by Eq. (16.2.3) exists and $w_\infty \in (0, \infty]$.*

If $w_\infty = +\infty$, then choose

$$\tilde{h}(t) = h(t) \quad \textit{and} \quad B(t) = \gamma t^\rho \int_1^t u^{\gamma-1} A(u) \mathrm{d}u; \tag{16.2.4}$$

and if $w_\infty < \infty$, then choose

$$\tilde{h}(t) = |w_\infty - h(t)| \quad \textit{and} \quad B(t) = -\gamma t^\rho \int_t^\infty u^{\gamma-1} A(u) \mathrm{d}u.$$

Proof: From Remark B.3.7 in [113], it follows that $h \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function A for $\rho = 0$ and $\gamma > 0$, and $h(\infty) - h \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function A for $\rho = 0$ and $\gamma < 0$, where $h(\infty)$ exists and is finite when $\gamma < 0$. This means that Part (i) and Part (ii) hold for the case $\rho = 0$. Thus, to prove the desired result, it suffices to prove the case $\rho < 0$.

Suppose that $\rho < 0$ and $\gamma \neq 0$. Again from Remark B.3.7 in [113], $\lim_{t\to\infty} t^{-\gamma} a(t) = c \in (0, \infty)$ and

$$\lim_{t\to\infty} \frac{h(tx) - c\frac{(tx)^\gamma}{\gamma} - \left(h(t) - c\frac{t^\gamma}{\gamma}\right)}{a(t)A(t)/\rho} = \frac{x^{\gamma+\rho} - 1}{\gamma + \rho}. \tag{16.2.5}$$

Now, we consider three cases:

<u>*Case 1: $\gamma + \rho > 0$*</u> Applying Theorem B.2.2 in [113] to Eq. (16.2.5), we have

$$h(t) - c\frac{t^\gamma}{\gamma} \sim \frac{1}{\rho(\rho+\gamma)} a(t)A(t) \sim \frac{c}{\rho(\rho+\gamma)} t^\gamma A(t), \quad t \to \infty,$$

or, equivalently,

$$h(t) = \frac{c}{\gamma} t^\gamma \left[1 + \frac{\gamma}{\rho(\rho+\gamma)} A(t) + o(A(t))\right], \quad t \to \infty. \tag{16.2.6}$$

Since $|A| \in \mathrm{RV}_\rho$ and $A(t) \to 0$ as $t \to \infty$, it follows from Eq. (16.2.6) that $h \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function $\frac{\gamma}{\gamma+\rho} A(t)$.

<u>*Case 2: $\gamma + \rho < 0$*</u> Again applying Theorem B.2.2 in [113] to Eq. (16.2.5), we get that

$$\lim_{t\to\infty} \left[h(t) - c\frac{t^\gamma}{\gamma}\right] = w_\infty \text{ exists and is finite,}$$

and

$$\lim_{t\to\infty}\frac{w_\infty-\left(h(t)-c\frac{t^\gamma}{\gamma}\right)}{a(t)A(t)/\rho}=-\frac{1}{\rho+\gamma}.$$

The last equation reduces to

$$w_\infty-h(t)=-\frac{c}{\gamma}t^\gamma\left[1+\frac{\gamma}{\rho(\rho+\gamma)}A(t)+o(A(t))\right],\quad t\to\infty.$$

This implies $-\frac{\gamma}{|\gamma|}(w_\infty-h)\in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function $\frac{\gamma}{\gamma+\rho}A(t)$. The second equality of Eq. (16.2.3) follows from Lemma 16.2.1.

Case 3: $\gamma+\rho=0$ From Eq. (16.2.5), we know that $h(t)-ct^\gamma/\gamma\in\Pi$ with auxiliary function $a(t)A(t)/\rho$. By Corollary B.2.13 in [113], w_∞ defined by Eq. (16.2.3) exists and $w_\infty\in(0,\infty]$. By Theorem B.2.12 in [113], there exists a function $\varphi\in\mathrm{RV}_0$ such that

$$\varphi(t)\sim\frac{1}{\rho}a(t)A(t)\sim\frac{c}{\rho}t^\gamma A(t),\quad t\to\infty,\tag{16.2.7}$$

and

$$h(t)-c\frac{t^\gamma}{\gamma}=\varphi(t)+\int_1^t\frac{\varphi(u)}{u}\mathrm{d}u.\tag{16.2.8}$$

Define

$$g(t)=\frac{\gamma}{ct^\gamma}\left[\varphi(t)+\int_1^t\frac{\varphi(u)}{u}\mathrm{d}u\right],\quad t>0.$$

Then

$$h(t)=c\frac{t^\gamma}{\gamma}+\frac{c}{\gamma}t^\gamma g(t)=\frac{c}{\gamma}t^\gamma\left[1+g(t)\right].$$

Since $\varphi\in\mathrm{RV}_0$, we have

$$\lim_{t\to\infty}\frac{\varphi(t)}{\int_1^t\varphi(u)/u\mathrm{d}u}=0\tag{16.2.9}$$

by Karamata's theorem (see Theorem B.1.5 in [113]). Hence, in terms of $\gamma>0$ and Eq. (16.2.7), exploiting L'Hôpital's rule yields that

$$\lim_{t\to\infty}g(t)=\lim_{t\to\infty}\frac{\gamma}{c}\frac{\int_1^t\varphi(u)/u\mathrm{d}u}{t^\gamma}=\lim_{t\to\infty}\frac{\varphi(t)/t}{ct^{\gamma-1}}=\lim_{t\to\infty}\frac{A(t)}{\rho}=0.\tag{16.2.10}$$

Two subcases arise: $w_\infty=\infty$ and $0<w_\infty<\infty$.

First, consider the subcase $w_\infty = \infty$. In view of Eq. (16.2.9), we conclude from Eq. (16.2.8) that

$$\int_1^\infty \frac{\varphi(u)}{u}\mathrm{d}u = \infty.$$

Define

$$A_*(t) = \int_1^t u^{\gamma-1}A(u)\mathrm{d}u.$$

Then $A_*(t) \to \infty$ as $t \to \infty$ since

$$\lim_{t\to\infty}\frac{A_*(t)}{\int_1^t \varphi(u)/u\mathrm{d}u} = \lim_{t\to\infty}\frac{t^{\gamma-1}A(t)}{\varphi(t)/t} = \frac{\rho}{c} > 0$$

by L'Hôpital's rule. Again, applying L'Hôpital's rule and by Eq. (16.2.7), we have

$$\lim_{t\to\infty}\frac{\int_1^{tx}\varphi(u)/u\mathrm{d}u}{cA_*(t)/\rho} = \lim_{t\to\infty}\frac{\varphi(tx)}{ct^\gamma A(t)/\rho} = x^\gamma \lim_{t\to\infty}\frac{A(tx)}{A(t)} = 1, \quad x > 0. \tag{16.2.11}$$

Therefore, for any $x > 0$,

$$\begin{aligned}
&\lim_{t\to\infty}\frac{h(tx)/h(t) - x^\gamma}{\gamma t^\rho A_*(t)} \\
= \;& x^\gamma \lim_{t\to\infty}\frac{g(tx)-g(t)}{\gamma t^\rho A_*(t)}\cdot\frac{1}{1+g(t))} \\
= \;& x^\gamma \lim_{t\to\infty}\frac{g(tx)-g(t)}{\gamma t^\rho A_*(t)} \qquad \text{[by (16.2.10)]} \\
= \;& x^\gamma \lim_{t\to\infty}\frac{(tx)^\rho\left(\varphi(tx)+\int_1^{tx}\varphi(u)/u\mathrm{d}u\right) - t^\rho\left(\varphi(t)+\int_1^t\varphi(u)/u\mathrm{d}u\right)}{ct^\rho A_*(t)} \\
= \;& x^\gamma \lim_{t\to\infty}\frac{[x^\rho\varphi(tx)-\varphi(t)] + \left[x^\rho\int_1^{tx}\varphi(u)/u\mathrm{d}u - \int_1^t\varphi(u)/u\mathrm{d}u\right]}{cA_*(t)} \\
= \;& x^\gamma \lim_{t\to\infty}\frac{x^\rho\int_1^{tx}\varphi(u)/u\mathrm{d}u - \int_1^t\varphi(u)/u\mathrm{d}u}{cA_*(t)} \qquad \text{[by (16.2.9)]} \\
= \;& x^\gamma\frac{x^\rho-1}{\rho}. \qquad \text{[by (16.2.11)]}
\end{aligned}$$

This means that $h \in 2\mathrm{RV}_{\gamma,\rho}$ with an auxiliary function $\gamma t^\rho A_*(t)$.

Next, we consider the subcase $0 < w_\infty < \infty$. Then, in terms of Eqs. (16.2.8) and (16.2.9), we have

$$w_\infty = \int_1^\infty \frac{\varphi(u)}{u}\mathrm{d}u < \infty \quad \textit{and} \quad A_*(\infty) < \infty$$

and, hence,

$$A^*(t) := \int_t^\infty u^{\gamma-1}A(u)\mathrm{d}u \longrightarrow 0, \quad t \to \infty.$$

By Karamata's theorem,

$$\lim_{t\to\infty} \frac{\varphi(t)}{A^*(t)} = \lim_{t\to\infty} \frac{\varphi(t)}{t^\gamma A(t)} \cdot \frac{t^\gamma A(t)}{A^*(t)} = 0. \tag{16.2.12}$$

Applying L'Hôpital's rule and by Eq. (16.2.7), we have

$$\lim_{t\to\infty} \frac{\int_{tx}^\infty \varphi(u)/u\mathrm{d}u}{cA_*(t)/\rho} = \lim_{t\to\infty} \frac{\rho\varphi(tx)}{ct^\gamma A(t)} = 1, \quad x > 0. \tag{16.2.13}$$

Since $w_\infty - h(t)$ has the same sign as $\rho A(t)$ for t large enough, we have, for any $x > 0$,

$$\begin{aligned}
&\lim_{t\to\infty} \frac{\frac{w_\infty - h(tx)}{w_\infty - h(t)} - x^\gamma}{\gamma t^\rho A^*(t)} \\
&= \lim_{t\to\infty} \frac{c(tx)^\gamma[g(tx) - g(t)] - \gamma w_\infty(1 - x^\gamma)}{ct^\gamma[1 + g(t)] - w_\infty \gamma} \cdot \frac{1}{\gamma t^\rho A^*(t)} \\
&= \lim_{t\to\infty} \frac{c(tx)^\gamma[g(tx) - g(t)] - \gamma w_\infty(1 - x^\gamma)}{c\gamma A^*(t)} \qquad \text{[by (16.2.10)]} \\
&= \lim_{t\to\infty} \frac{[\varphi(tx) - \varphi(t)x^\gamma] + \int_1^{tx}\varphi(u)/u\mathrm{d}u - x^\gamma \int_1^t \varphi(u)/u\mathrm{d}u - w_\infty(1 - x^\gamma)}{cA^*(t)} \\
&= \lim_{t\to\infty} \frac{\int_1^{tx}\varphi(u)/u\mathrm{d}u - x^\gamma \int_1^t \varphi(u)/u\mathrm{d}u - w_\infty(1 - x^\gamma)}{cA^*(t)} \qquad \text{[by (16.2.12)]} \\
&= \lim_{t\to\infty} \frac{x^\gamma \int_t^\infty \varphi(u)/u\mathrm{d}u - \int_{tx}^\infty \varphi(u)/u\mathrm{d}u}{cA^*(t)} \\
&= \frac{x^\gamma - 1}{\rho} = -x^\gamma \frac{x^\rho - 1}{\rho}, \qquad \text{[by (16.2.13) and } \rho + \gamma = 0]
\end{aligned}$$

which means that $|w_\infty - h| \in 2\mathrm{RV}_{\gamma,\rho}$ with an auxiliary function $-\gamma t^\rho A^*(t)$. This completes the proof of the proposition. ■

Neves (2009, [358]) conducted a similar study on the relationships from 2ERV to 2RV under a restriction $\gamma \neq \rho$ upon the parameters. The conclusions of Theorem 1 in [358] and of Proposition 16.2.2 are the same under the conditions (1) $\gamma > 0$ and $\gamma + \rho > 0$, and (2) $\gamma > 0$, $\gamma + \rho < 0$ and $w_\infty = 0$, respectively. For convenience of comparison, we list the other conclusions of [358] in the following proposition.

Proposition 16.2.3 (Neves [358]). *Let h be any measurable (eventually) positive function, and let w_∞ be as defined by Eq. (16.2.3) for $\rho < 0$. Assume that $h \in 2\mathrm{ERV}_{\gamma,\rho}$ with first-order and second-order auxiliary functions $a(t)$ and $A(t)$, respectively. Then*

(i) *For $\gamma > 0$, $\gamma + \rho < 0$, and $w_\infty \neq 0$, $h \in 2\mathrm{RV}_{\gamma,-\gamma}$ with auxiliary function $a(t)/h(t) - \gamma$*

(ii) *For $\gamma > 0$ and $\gamma + \rho = 0$, $h \in 2\mathrm{RV}_{\gamma,-\gamma}$ with auxiliary function $a(t)/h(t) - \gamma$*

(iii) *For $\gamma \leq 0$, $h \in \mathrm{ERV}_\gamma$*

Proposition 16.2.3(iii) is trivial since $h \in 2\mathrm{ERV}_{\gamma,\rho}$ implies $h \in \mathrm{ERV}_\gamma$. Proposition 16.2.3(i) and (ii) can be derived from Proposition 16.2.2 as follows.

Proposition 16.2.2(ii) $\Longrightarrow$ Proposition 16.2.3(i): From Lemma 16.2.1, it is easy to see that, for $\gamma > 0$, $\gamma + \rho < 0$, and any constant $d \neq 0$,

$$h \in 2\mathrm{RV}_{\gamma,\rho} \Longrightarrow h(t) + d \in 2\mathrm{RV}_{\gamma,-\gamma} \ \text{ with auxiliary function } -\frac{\gamma d}{c} t^{-\gamma}, \tag{16.2.14}$$

where $c = \lim_{t\to\infty} t^{-\gamma} h(t)$. Now, suppose that $h \in 2\mathrm{ERV}_{\gamma,\rho}$ with first-order and second-order auxiliary functions $a(t)$ and $A(t)$, respectively. For $\gamma > 0$, $\gamma + \rho < 0$, and $w_\infty \neq 0$, by Proposition 16.2.2 (ii), we have $h(t) - w_\infty \in 2\mathrm{RV}_{\gamma,\rho}$. Thus, in view of the observation (16.2.14), $h = (h - w_\infty) + w_\infty \in 2\mathrm{RV}_{\gamma,-\gamma}$ with auxiliary function $B(t)$ given by

$$B(t) = -\gamma \frac{\lim_{s\to\infty}(h(s) - a(s)/\gamma)}{\lim_{s\to\infty} s^{-\gamma} h(s)} t^{-\gamma} \sim -\gamma \frac{(h(t) - a(t)/\gamma)}{t^{-\gamma} h(t)} t^{-\gamma} = \frac{a(t)}{h(t)} - \gamma$$

as $t \to \infty$. ■

Proposition 16.2.2(iii) $\Longrightarrow$ Proposition 16.2.3(ii): First, consider the case $\gamma > 0$, $\rho + \gamma = 0$, and $w_\infty = \infty$; Proposition 16.2.2(iii) implies $h \in 2\mathrm{RV}_{\gamma,-\gamma}$ with auxiliary function B given by Eq. (16.2.4). It suffices to prove that

$$B(t) \sim \frac{a(t)}{h(t)} - \gamma, \quad t \to \infty. \tag{16.2.15}$$

From Eq. (16.2.5), we know $h(t) - ct^\gamma/\gamma \in \Pi$. By Theorem B.2.12 in de Haan and Ferreira [113], there exists $\varphi \in \mathrm{RV}_0$ such that

$$\varphi(t) \sim \frac{a(t)A(t)}{\rho}, \quad t \to \infty, \tag{16.2.16}$$

and

$$h(t) - \frac{a(t)}{\gamma} \sim h(t) - \frac{c}{\gamma}t^\gamma = \varphi(t) + \int_1^t \frac{\varphi(s)}{s}\mathrm{d}s, \quad t \to \infty.$$

Since $w_\infty = \infty$ and $\varphi(t) = o\left(\int_1^t \varphi(s)/s\mathrm{d}s\right)$ as $t \to \infty$ by Karamata's theorem, it follows that $\int_1^t \varphi(s)/s\mathrm{d}s \to \infty$ as $t \to \infty$ and, hence,

$$h(t) - \frac{a(t)}{\gamma} \sim \int_1^t \frac{\varphi(s)}{s}\mathrm{d}s \sim \int_1^t \frac{a(s)A(s)}{s\rho}\mathrm{d}s, \quad t \to \infty,$$

where the second asymptotic equivalence follows from Eq. (16.2.16) by applying L'Hôpital's rule. Therefore,

$$\begin{aligned} \lim_{t\to\infty} \frac{B(t)}{a(t)/h(t) - \gamma} &= -\lim_{t\to\infty} \frac{h(t)}{\gamma}\frac{B(t)}{h(t) - a(t)/\gamma} \\ &= -\lim_{t\to\infty} \frac{c}{\gamma}\frac{\int_1^t u^{\gamma-1}A(u)\mathrm{d}u}{\int_1^t \frac{a(s)A(s)}{s\rho}\mathrm{d}s} = \lim_{t\to\infty}\frac{ct^\gamma}{a(t)} = 1, \end{aligned}$$

implying (16.2.15).

Next, consider the case $\gamma > 0$, $\rho + \gamma = 0$, and $0 < w_\infty < \infty$. Applying Corollary B.2.13 in [113] to Eq. (16.2.5) yields that

$$\frac{a(t)A(t)}{\rho} = o\left(w_\infty - \left[h(t) - \frac{c}{\gamma}t^\gamma\right]\right), \quad t \to \infty,$$

implying $a(t)A(t) \to 0$ as $t \to \infty$. Hence,

$$\lim_{t\to\infty} \frac{A(t)}{\gamma h(t)/a(t) - 1} = \frac{1}{\gamma} \lim_{t\to\infty} \frac{a(t)A(t)}{h(t) - a(t)/\gamma} = \frac{1}{\gamma} \lim_{t\to\infty} \frac{a(t)A(t)}{w_\infty} = 0.$$

Since $a \in 2\mathrm{RV}_{\gamma,\rho}$, by Lemma 16.2.1, we have

$$\begin{aligned} h(t) &= \frac{a(t)}{\gamma} + \left(h(t) - \frac{a(t)}{\gamma}\right) \\ &= \frac{a(t)}{\gamma}\left[1 + \left(\frac{\gamma h(t)}{a(t)} - 1\right)\right] \\ &= \frac{c}{\gamma} t^\gamma \left(1 + \frac{A(t)}{\rho} + o(A(t))\right)\left[1 + \left(\frac{\gamma h(t)}{a(t)} - 1\right)\right] \\ &= \frac{c}{\gamma} t^\gamma \left(1 + \left(\frac{\gamma h(t)}{a(t)} - 1\right)(1 + o(1))\right) \\ &= \frac{c}{\gamma} t^\gamma \left(1 + \frac{1}{-\gamma}\left(\frac{a(t)}{h(t)} - \gamma\right)(1 + o(1))\right). \end{aligned} \tag{16.2.17}$$

Note that

$$\left|\frac{a(t)}{h(t)} - \gamma\right| = \frac{\gamma}{|h(t)|} \cdot \left|h(t) - \frac{a(t)}{\gamma}\right| = \frac{w_\infty \gamma}{|h(t)|} \sim \frac{w_\infty \gamma^2}{a(t)} \in \mathrm{RV}_{-\gamma}.$$

Therefore, again, applying Lemma 16.2.1 to Eq. (16.2.17) yields Proposition 16.2.3(ii) for the case $w_\infty < \infty$. ■

Next, we are going to investigate the relationships between a survival function $\overline{F}$ and its tail quantile function $U = \left(1/\overline{F}\right)^{\leftarrow}$ when they exhibit some form of second-order condition.

Proposition 16.2.4. *Let $U \in 2\mathrm{ERV}_{\gamma,\rho}$ with first-order and second-order auxiliary functions $a(t)$ and $A(t)$, respectively, where $\gamma \neq 0$ and $\rho \leq 0$. Define*

$$w_\infty = \lim_{t\to\infty}\left(U(t) - \frac{a(t)}{\gamma}\right) \quad \textit{when } \rho < 0.$$

Then

(i) *For $\gamma > 0$ and $\rho + \gamma > 0$ or for $\gamma > 0$, $\rho + \gamma < 0$, and $w_\infty = 0$, $\overline{F} \in 2\mathrm{RV}_{-1/\gamma,\rho/\gamma}$ with an auxiliary function $\frac{1}{\gamma(\rho+\gamma)} A(1/\overline{F}(t))$*

(ii) *For $\gamma > 0$ and $\rho + \gamma = 0$ or for $\gamma > 0$, $\rho + \gamma < 0$, and $w_\infty \neq 0$, $\overline{F} \in 2\mathrm{RV}_{-1/\gamma,-1}$ with an auxiliary function $\frac{1}{\gamma}\left(\frac{a(1/\overline{F}(t))}{t} - 1\right)$*

(iii) *For $\gamma < 0$, $\overline{F}(\hat{x} - 1/\cdot) \in 2\mathrm{RV}_{1/\gamma, -\rho/\gamma}$ with an auxiliary function $-\frac{1}{\gamma(\rho+\gamma)} A(1/\overline{F}(\hat{x} - 1/t))$*

Proof:

(i) For $\gamma > 0$ and $\rho + \gamma > 0$, from Proposition 16.2.2(i),

$$U \in 2\mathrm{RV}_{\gamma,\rho} \text{ with auxiliary function } \frac{\gamma}{\rho+\gamma} A(t). \tag{16.2.18}$$

By Proposition 2.6 in [299], we have

$$\frac{1}{\overline{F}} \in 2\mathrm{RV}_{1/\gamma,\rho/\gamma} \text{ with auxiliary function } -\frac{1}{\gamma(\rho+\gamma)} A(1/\overline{F}(t)).$$

Then, by Proposition 2.5 in [299],

$$\overline{F} \in 2\mathrm{RV}_{-1/\gamma,\rho/\gamma} \text{ with auxiliary function } \frac{1}{\gamma(\rho+\gamma)} A(1/\overline{F}(t)). \tag{16.2.19}$$

It should point out that the conclusion of Proposition 2.6 in [299] is stated only for $\rho < 0$, but by Vervaat's Theorem in Appendix A of [113], one can prove that the conclusion is also valid for increasing function U with $\rho = 0$. In fact, Eq. (16.2.19) follows from Eq. (16.2.18) directly by Theorem 2.3.9 in [113].

(ii) The proof is similar to Part (i) by Proposition 16.2.3.

(iii) For $\gamma < 0$, from Proposition 16.2.2(ii),

$$\hat{x} - U(\cdot) \in 2\mathrm{RV}_{\gamma,\rho} \text{ with auxiliary function } \frac{\gamma}{\rho+\gamma} A(t).$$

This is equivalent to

$$\frac{1}{\hat{x} - U(\cdot)} \in 2\mathrm{RV}_{-\gamma,\rho} \text{ with auxiliary function } -\frac{\gamma}{\rho+\gamma} A(t).$$

Since the inverse function of $1/(\hat{x} - U(\cdot))$ is $1/\overline{F}(\hat{x} - 1/\cdot)$, then

$$\frac{1}{\overline{F}(\hat{x} - 1/\cdot)} \in 2\mathrm{RV}_{-1/\gamma,-\rho/\gamma} \text{ with auxiliary function } \frac{1}{\gamma(\rho+\gamma)} A\left(\frac{1}{\overline{F}(\hat{x} - 1/t)}\right)$$

by a similar argument to Part (i). Now, Part (ii) follows.

Therefore, we complete the proof of the proposition. ∎

16.3 Inequalities of Drees Type

In this section, we give new Drees-type inequalities for ERV and 2ERV functions. Compared with the ordinary Drees-type inequalities, the original auxiliary functions in these new inequalities are not replaced by other ones with special forms.

Proposition 16.3.1. *If $h \in \mathrm{ERV}_\gamma$ with auxiliary function $a(t)$ and $\gamma \in \mathbb{R}$, then for any $\varepsilon, \delta > 0$, there exists $t_0 = t_0(\varepsilon, \delta, \gamma)$ such that for all $t,\ tx \geq t_0$,*

$$\left|\frac{h(tx) - h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}\right| \leq \varepsilon \left(\left|\frac{x^\gamma - 1}{\gamma}\right| + x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right). \tag{16.3.1}$$

Proof: Suppose $h \in \mathrm{ERV}_\gamma$ with auxiliary function $a(t)$ and $\gamma \in \mathbb{R}$. Let a_0 be as defined by Eq. (16.1.6). Since $a(t) \sim a_0(t)$ as $t \to \infty$, for any $\varepsilon \in (0, 1)$ and $\delta > 0$, choose $t_0 = t_0(\varepsilon, \delta, \gamma)$ such that for all $t,\ tx \geq t_0$, Eq. (16.1.5) holds and

$$0 < (1 - \varepsilon)a(t) \leq a_0(t) \leq (1 + \varepsilon)a(t).$$

Now, consider two cases: $x \geq 1$ and $x \in (0, 1)$.

Case 1. $t > t_0$ and $x \geq 1$ Note that

$$\begin{aligned} h(tx) - h(t) &\leq a_0(t)\frac{x^\gamma - 1}{\gamma} + \varepsilon a_0(t) x^\gamma \max\left\{x^\delta, x^{-\delta}\right\} \\ &\leq a(t)\frac{x^\gamma - 1}{\gamma} + \varepsilon a(t)\left(\frac{x^\gamma - 1}{\gamma} + (1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right), \end{aligned}$$

implying

$$\frac{h(tx) - h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma} \leq \varepsilon \left(\frac{x^\gamma - 1}{\gamma} + (1 + \varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right). \tag{16.3.2}$$

Similarly,

$$\begin{aligned} h(tx) - h(t) &\geq a_0(t)\frac{x^\gamma - 1}{\gamma} - \varepsilon a_0(t) x^\gamma \max\left\{x^\delta, x^{-\delta}\right\} \\ &\geq a(t)\frac{x^\gamma - 1}{\gamma}(1 - \varepsilon) - \varepsilon a(t)(1 + \varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\} \\ &= a(t)\frac{x^\gamma - 1}{\gamma} - \varepsilon a(t)\left(\frac{x^\gamma - 1}{\gamma} + (1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right), \end{aligned}$$

implying

$$\frac{h(tx)-h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma} \geq -\varepsilon\left(\frac{x^\gamma-1}{\gamma} + (1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right). \tag{16.3.3}$$

Thus, Eq. (16.3.1) follows from Eqs. (16.3.2) and (16.3.3) by replacing the above ε by $\varepsilon/2$.

Case 2. $tx > t_0$ and $0 < x < 1$ Note that $(x^\gamma - 1)/\gamma < 0$. Then,

$$\begin{aligned} h(tx) - h(t) &\leq a(t)(1-\varepsilon)\frac{x^\gamma-1}{\gamma} + \varepsilon a(t)(1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\} \\ &\leq a(t)\frac{x^\gamma-1}{\gamma} + \varepsilon a(t)\left(-\frac{x^\gamma-1}{\gamma} + (1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right), \end{aligned}$$

implying

$$\frac{h(tx)-h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma} \leq \varepsilon\left(-\frac{x^\gamma-1}{\gamma} + (1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right).$$

Similarly,

$$\frac{h(tx)-h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma} \geq -\varepsilon\left(-\frac{x^\gamma-1}{\gamma} + (1+\varepsilon)x^\gamma \max\left\{x^\delta, x^{-\delta}\right\}\right).$$

Therefore, the desired result follows. ■

Note that $|\log x| \leq C \max\left\{x^\delta, x^{-\delta}\right\}$ for some constant C and all $x > 0$. Also, $h \in 2\mathrm{RV}_{\gamma,\rho}$ if and only if $t^{-\gamma}h(t) \in \mathrm{ERV}_\rho$ for $\gamma \in \mathbb{R}$ and $\rho \leq 0$. Two immediate consequences of Proposition 16.3.1 are the following two corollaries.

Corollary 16.3.2. *Let $h \in \Pi$ with auxiliary function $a(t)$. Then, for any $\varepsilon > 0$ and $\delta > 0$, there exists $t_0 = t_0(\varepsilon, \delta)$ such that for all $t,\ tx \geq t_0$,*

$$\left|\frac{h(tx)-h(t)}{a(t)} - \log x\right| \leq \varepsilon \max\left\{x^\delta, x^{-\delta}\right\}.$$

Corollary 16.3.3. *If $h \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary functions $A(t)$, $\gamma \in \mathbb{R}$ and $\rho \leq 0$, then, for any $\varepsilon, \delta > 0$, there exists $t_0 = t_0(\varepsilon, \delta, \gamma, \rho) > 0$ such that for all $t,\ tx \geq t_0$,*

$$\left|\frac{\frac{h(tx)}{h(t)} - x^\gamma}{A(t)} - x^\gamma \frac{x^\rho - 1}{\rho}\right| \leq \varepsilon\, x^\gamma\left(\left|\frac{x^\rho-1}{\rho}\right| + x^\rho \max\left\{x^\delta, x^{-\delta}\right\}\right).$$

Proposition 16.3.4. *If $h \in 2\mathrm{ERV}_{\gamma,\rho}$ with first-order and second-order auxiliary functions $a(t)$ and $A(t)$, respectively, then, for any $\varepsilon, \delta > 0$, there exists $t_0 = t_0(\varepsilon, \delta, \gamma, \rho) > 0$ such that for all $t,\ tx \geq t_0$,*

$$\left| \frac{\frac{h(tx)-h(t)}{a(t)} - \frac{x^\gamma - 1}{\gamma}}{A(t)} - H_{\gamma,\rho}(x) \right| \leq \varepsilon\, W_{\gamma,\rho,\delta}(x), \tag{16.3.4}$$

where $W_{\gamma,\rho,\delta}(x)$ is given by

$$W_{\gamma,\rho,\delta}(x) = \begin{cases} \left|\frac{x^{\gamma+\rho}-1}{\gamma+\rho}\right| + x^\gamma \left|\frac{x^\rho - 1}{\rho}\right| + x^{\gamma+\rho} \max\left\{x^\delta, x^{-\delta}\right\}, & \gamma \neq 0; \\ \left|\frac{x^\rho-1}{\rho}\right| + x^\rho \max\left\{x^\delta, x^{-\delta}\right\}, & \gamma=0,\ \rho\neq0; \\ \max\left\{x^\delta, x^{-\delta}\right\}, & \gamma=0,\ \rho=0. \end{cases}$$

Proof:

(1) First, consider the case $\gamma \neq 0$. Note that, for $x > 0$,

$$\begin{aligned} &\frac{h(tx) - h(t) - a(t)\frac{x^\gamma - 1}{\gamma}}{a(t)A(t)} \\ &= \frac{h(tx) - \frac{1}{\gamma}a(t)x^\gamma - \left(h(t) - \frac{1}{\gamma}a(t)\right)}{a(t)A(t)} \\ &= \frac{h(tx) - a(tx)\frac{1}{\gamma} - \left(h(t) - \frac{1}{\gamma}a(t)\right)}{a(t)A(t)} + \frac{x^\gamma}{\gamma}\, \frac{(tx)^{-\gamma}a(tx) - t^{-\gamma}a(t)}{t^{-\gamma}a(t)A(t)}. \end{aligned} \tag{16.3.5}$$

Since $a \in 2\mathrm{RV}_{\gamma,\rho}$ with auxiliary function $A(t)$, it follows that $t^{-\gamma}a(t) \in \mathrm{ERV}_\rho$ with auxiliary function $t^{-\gamma}a(t)A(t)$ and, hence, the second term in the right hand of Eq. (16.3.5) converges to $\frac{x^\gamma}{\gamma}\frac{x^\rho - 1}{\rho}$ as $t \to \infty$. Then the first term in the right hand of Eq. (16.3.5) converges to $-\frac{1}{\gamma}\frac{x^{\gamma+\rho}-1}{\gamma+\rho}$ as $t \to \infty$, i.e., $h - a/\gamma \in \mathrm{ERV}_{\gamma+\rho}$ with auxiliary function $-a(t)A(t)/\gamma$. So, by Proposition 16.3.1, for any $\varepsilon > 0$ and $\delta > 0$, there exists $t_1 = t_1(\varepsilon, \delta, \gamma, \rho)$ such that for all $t, tx \geq t_1$,

$$\left| \frac{h(tx) - a(tx)\frac{1}{\gamma} - \left(h(t) - \frac{1}{\gamma}a(t)\right)}{a(t)A(t)} + \frac{1}{\gamma}\frac{x^{\gamma+\rho} - 1}{\gamma + \rho} \right|$$

$$\leq \frac{\varepsilon}{2|\gamma|}\left(\left|\frac{x^{\gamma+\rho}-1}{\gamma+\rho}\right| + x^{\gamma+\rho}\max\left\{x^{\delta}, x^{-\delta}\right\}\right)$$

and

$$\left|\frac{(tx)^{-\gamma}a(tx) - t^{-\gamma}a(t)}{t^{\gamma}a(t)A(t)} - \frac{x^{\rho}-1}{\rho}\right| \leq \frac{\varepsilon}{2}\left(\left|\frac{x^{\rho}-1}{\rho}\right| + x^{\rho}\max\left\{x^{\delta}, x^{-\delta}\right\}\right).$$

Inserting these two inequalities into Eq. (16.3.5), we conclude (16.3.4) for the case $\gamma \neq 0$.

(2) Next, consider the case $\gamma = 0$ and $\rho < 0$. Since $a \in 2\mathrm{RV}_{0,\rho}$ with auxiliary function $A(t)$, we have $a(t) = c[1 + A(t)/\rho + o(A(t))]$ as $t \to \infty$ with $0 < c < \infty$ by Lemma 16.2.1. Then, for $x > 0$,

$$\begin{aligned} &\frac{h(tx) - h(t) - a(t)\log x}{a(t)A(t)} \\ = \quad &\frac{h(tx) - h(t) - c\left(1 + \frac{A(t)}{\rho} + o(A(t))\right)\log x}{a(t)A(t)} \\ = \quad &\frac{h(tx) - c\log(tx) - (h(t) - c\log t)}{a(t)A(t)} - \frac{c + o(1)}{\rho a(t)}\log x. \end{aligned}$$

By Theorem B.3.6 in de Haan and Ferreira [113] or by Eq. (16.2.5), we get $h(t) - c\log t \in \mathrm{ERV}_{\rho}$ with auxiliary function $a(t)A(t)/\rho$. So, by Proposition 16.3.1, for any $\varepsilon > 0$ and $\delta > 0$, there exists $t_1 = t_1(\varepsilon, \delta, \gamma, \delta)$ such that for all $t, tx \geq t_1$,

$$\begin{aligned} &\left|\frac{h(tx) - c\log tx - (h(t) - c\log t)}{a(t)A(t)} - \frac{x^{\rho}-1}{\rho^2}\right| \\ &\quad \leq \frac{\varepsilon}{2|\rho|}\left(\left|\frac{x^{\rho}-1}{\rho}\right| + x^{\rho}\max\left\{x^{\delta}, x^{-\delta}\right\}\right) \end{aligned}$$

and

$$\left|\frac{c + o(1)}{\rho a(t)}\log x + \frac{1}{\rho}\log x\right| \leq \frac{\varepsilon}{2|\rho|}\max\left\{x^{\delta}, x^{-\delta}\right\}.$$

Thus,

$$\left|\frac{h(tx) - h(t) - a(t)\log x}{a(t)A(t)} - H_{0,\rho}(x)\right| \leq \varepsilon\left(\left|\frac{x^{\rho}-1}{\rho^2}\right| + x^{\rho}\max\{x^{\delta}, x^{-\delta}\}\right).$$

(3) Finally, consider the case $\gamma = \rho = 0$. From Eq. (16.1.7), it follows that, for any $\varepsilon > 0$ and $\delta > 0$, there exists $t_1 = t_1(\varepsilon, \delta)$ such that for all $t, tx \geq t_1$,

$$\left| \frac{\frac{h(tx)-h(t)}{a_0(t)} - \log x}{A_0(t)} - \frac{1}{2}(\log x)^2 \right| \leq \frac{1}{4}\varepsilon \max\left\{x^{\delta}, x^{-\delta}\right\},$$

where a_0 and A_0 are chosen such that $A_0(t) \sim A(t)$ and $a_0(t)/a(t) - 1 = o(A(t))$ as $t \to \infty$. Note that

$$\begin{aligned} &\frac{\frac{h(tx)-h(t)}{a(t)} - \log x}{A(t)} \\ = \ & \frac{h(tx) - h(t) - a_0(t)\log x}{a_0(t)A_0(t)} \cdot \frac{a_0(t)A_0(t)}{a(t)A(t)} - \frac{a(t) - a_0(t)}{a(t)A(t)} \log x \\ \stackrel{\text{def}}{=} \ & J_1(t) + J_2(t), \end{aligned}$$

where

$$J_1(t) \longrightarrow \frac{1}{2}(\log x)^2 \quad \text{and} \quad J_2(t) \longrightarrow 0 \quad \text{as} \quad t \to \infty.$$

Also, note that $(\log x)^2 \leq C \max\left\{x^{\delta}, x^{-\delta}\right\}$ for all $x > 0$ and some constant C. Therefore, as t large enough,

$$\begin{aligned} &\left| J_1(t) - \frac{1}{2}(\log x)^2 \right| \\ \leq \ & \left| \frac{h(tx) - h(t) - a_0(t)\log x}{a_0(t)A_0(t)} - \frac{1}{2}(\log x)^2 \right| \cdot \frac{a_0(t)A_0(t)}{a(t)A(t)} \\ & + \left| \frac{1}{2}(\log x)^2 \right| \left| \frac{a_0(t)A_0(t)}{a(t)A(t)} - 1 \right| \\ \leq \ & \frac{1}{2}\varepsilon \max\left\{x^{\delta}, x^{-\delta}\right\}, \end{aligned}$$

and

$$\left| \frac{a(t) - a_0(t)}{a(t)A(t)} \log x \right| \leq \frac{1}{2}\varepsilon \max\left\{x^{\delta}, x^{-\delta}\right\}.$$

This means that Eq. (16.3.4) holds for this last case. We thus complete the proof of the whole proposition. ■

Acknowledgments

The research is supported by the Fundamental Research Funds for the Central Universities and the NNSF of China (Nos. 70821001).

Chapter 17
Individual and Moving Ratio Charts for Weibull Processes

Francis Pascual

Abstract: This chapter proposes methods for monitoring Weibull processes when data collection is restricted to one observation per sampling period. Because the Weibull distribution is an asymmetric distribution, it is not appropriate to apply normal-based individual and moving range charts to Weibull data. A transformation of Weibull to approximate normality prior to applying normal-based methods has been suggested in the literature. This chapter studies the run length properties of and the difficulties encountered with this approach. As an alternative, this chapter proposes combined individual and moving range charts for monitoring changes in either the Weibull scale or shape parameter. A method for computing the average run length ARL is discussed, and control limits are presented for ARL-unbiased charts. The proposed method is applied to a Weibull data set.

F. Pascual (✉)
Department of Mathematics, Washington State University, Pullman,
WA 99164-3113, USA
e-mail: jpascual@math.wsu.edu

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_17,

17.1 Introduction

The Weibull distribution is named in honor of Weibull who, in [472, 473], used it to describe the breaking strength distribution of materials. Its theoretical origin is rooted in extreme value theory. The time to failure of a series system with n parts is the failure time of the first failure. For instance, in composite materials such as ceramics, failure happens at the weakest flaw of the test specimen. For many distributions that describe failure times, the distribution of the "weakest link" approaches the Weibull distribution as the number n of parts or flaws becomes large. See [171]. The Weibull distribution is now often used to model lifetimes of test units or electronic products and strength distributions of materials. References [274, 317, 356, 357, 388] describe various Weibull applications in engineering.

There are practical situations, e.g., chemical processes, when samples are restricted to a single observation per sampling period because of budget, time, and resource constraints. Thus, there is not enough observations that define a "rational subgroup" from which to compute statistics (e.g., sample mean and variance) that are relevant to process performance. One approach is to construct respective charts for individual values or the moving ranges (absolute differences between consecutive individual values). These charts can be used individually or simultaneously.

Work on individual and moving range charts is predominantly on the normal distribution. There is a disproportionately smaller amount of work on these charts for non-normal or asymmetric distributions such as Weibull. Because of the Weibull distribution's importance in real-life applications, the Weibull versions of these charts are valuable tools for the practitioner in process monitoring.

17.1.1 Outline

This chapter is organized as follows. Section 17.2 describes methods in the literature for monitoring Weibull processes. Section 17.3 describes model assumptions and notation that will be used in the discussions. Section 17.4 describes a transformation proposed in the literature to approximate normality of the Weibull distribution that can be applied before normal-based control charts can be implemented. Properties of this approach are explored specifically when samples are restricted to single observations. Section 17.5 describes the proposed combined individual and moving ratio charts for Weibull processes. The following section describes computation of the average run length (ARL) for the proposed method. Control limits are tabulated in Sect. 17.7. The proposed method is applied to an example in Sect. 17.8.

17.2 Related Work

When normal-based median and range charts were applied to Weibull processes, [354] concluded that the rates of out-of-control signals were different from what were anticipated. Chen and Cheng [85] investigated the effect of non-normality on the control limits of the $\bar{X}$ chart and concluded that the effects of non-normality were significant and "should not be ignored." Benneyan [58] studied the consequences of applying normal-based individual and moving range charts on non-normal single-parameter distributions. The author remarked that the charts have a poor ability to detect process changes for geometric and exponential processes and, furthermore, noted that using 3σ or probability limits often resulted in ARL that was significantly different from what was anticipated.

If normal-based control charts are used on an exponential variable W, [353] suggested the transformation $W^{0.2777}$ because the respective skewness and kurtosis of $W^{0.2777}$'s distribution are close to those of the normal. As an improvement, [483] suggested the power 0.2654 instead of 0.2777 based on the Kullback-Leibler information. These results are extended by noting that if Y is a Weibull variable with shape parameter β, then Y^{β} is exponential. Normal-based control charts can then be applied to values of $Y^{0.2777\beta}$, as done in [45], or $Y^{0.2654\beta}$. See also [58].

Padgett and Spurrier [363] proposed Shewhart-type charts for monitoring quantiles of Weibull strength distributions. Nichols and Padgett [359] computed control limits for monitoring Weibull quantiles by Monte Carlo simulations of maximum likelihood estimates. Huang and Pascual [195] discussed ARL-unbiased control charts for monitoring the quantiles of a Weibull process when the shape parameter β was known and data were censored at the first failure. Ramalhoto and Morais [381, 382] studied EWMA and Shewhart control charts for the Weibull scale parameter when the Weibull shape and threshold parameters were known. Zhang and Chen [485] studied EWMA charts for monitoring the mean of Weibull lifetimes assuming that the shape parameter was known and stable.

Pascual and Zhang [368] studied ARL-unbiased control charts based on the sample range of log-Weibull [smallest extreme value (SEV)] data for monitoring β. For Weibull samples of size $n = 1$, [367] studied ARL-unbiased moving range charts for monitoring β. Pascual [366] studied methods for monitoring β based on the Page cumulative sum in [364] and the exponentially weighted moving average. The authors of the above articles showed that their charts depended on the sample size n and the ratio between the stable and true values of β and not on η. Their chart schemes were motivated by the need to check the stability of the shape parameter β that is assumed by other methods found in the literature for monitoring the Weibull mean, quantiles, or scale parameter η.

The body of work on individual and moving range charts is mostly for the normal distribution. They have been applied in various areas such as engineering, health care, chemistry, and computer science. References [10, 136] are early references that described the use of these charts. Crowder [97] discussed how to compute the ARL, while [460] (p. 172) used the former's results to tabulate constants for computing control limits that yield target stable-process ARLs.

17.3 Model Assumptions

Let $i = 1, 2, \ldots$ denote the sampling period. In period i, the practitioner takes a single measurement X_i assumed to follow a two-parameter Weibull distribution with scale and shape parameters $\eta > 0$ and $\beta > 0$, respectively. Write $X_i \sim \mathrm{WEI}(\eta, \beta)$. The probability density (pdf) and cumulative distribution (cdf) functions are, respectively,

$$f_X(x; \eta, \beta) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right], \quad F_X(x; \eta, \beta) = 1 - \exp\left[-\left(\frac{x}{\eta}\right)^{\beta}\right] \tag{17.3.1}$$

for $x \geq 0$. The Weibull mean and variance are

$$\mathrm{E}[X_i] = \eta\Gamma\left(1 + \frac{1}{\beta}\right), \quad \mathrm{Var}[X_i] = \eta^2\left[\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma^2\left(1 + \frac{1}{\beta}\right)\right] \tag{17.3.2}$$

where Γ denotes the gamma function.

For period $i = 2, 3, \ldots$, define the moving ratio as

$$R_i = \max\{X_{i-1}, X_i\} / \min\{X_{i-1}, X_i\}.$$

Of course, there is no moving ratio in period 1 because there is no prior observation. The results of [464] can be used to show that the cdf of R_i is

$$F_R(r;\beta) = 1 - \frac{2}{r^\beta + 1}, \quad r \geq 1$$

from which can be derived the p quantile given by

$$r_p = \left(\frac{1+p}{1-p}\right)^{1/\beta}, \tag{17.3.3}$$

for $0 < p < 1$. Observe that r_p is a strictly decreasing function of β. Thus, R_i is an appropriate statistic for monitoring shifts in β that does not depend on the actual value of η. Also, observe that $\log(X_i)$ has a SEV distribution, and $\log(R_i)$ is a moving range of two consecutive SEV variables. The distributional properties of $\log R_i$ are discussed in [367]. They studied SEV moving range control charts for monitoring changes in β.

Below, control charts based on the individual and moving ratio are explored. Equation (17.3.2) suggests that interpretation of out-of-control signals in the individual chart may be tricky because they may be attributed to either shifts in the scale η and/or the shape β. However, the distributional properties of R_i, e.g., dependence on β only, suggest that simultaneously implementing a moving range chart may provide valuable information in pinpointing the cause(s) for signals.

17.4 Normal Individual and Moving Range Charts

If W is an exponential random variable, then $W^{0.2777}$ is approximately normal with similar symmetry and kurtosis as normal's. See [353]. It can easily be shown that if $X \sim \text{WEI}(\eta, \beta)$, then X^β is exponential with mean η^β. Hence, $Y = X^{0.2777\beta}$ is approximately normal.

Based on the above results, [45] studied the application of normal-based individual and moving range charts on the process $Y_i = X_i^{0.2777\beta_S}$ where β_S is the target Weibull shape. The authors simulated values of the individuals Y_i and the moving range $MR_i = |Y_i - Y_{i-1}|$ assuming a stable process and obtained respective sample means and standard deviations. They then used the mean to approximate the centerline and the mean $\pm$ 3 standard deviations to approximate the upper/lower control limits. They did not recommend using the moving range chart because their simulation study suggested that it was not effective at all in detecting shifts in all scenarios that they considered. They recommended the individual chart but only in specific cases. The theoretical properties of the above approach are presented below.

17.4.1 Shewhart Control Limits

If $X_i \sim \text{WEI}(\eta, \beta)$ and $\lambda > 0$, then it can be shown that $Y_i = X_i^{\lambda\beta} \sim \text{WEI}(\eta^{\lambda\beta}, 1/\lambda)$ for any positive constant λ. Thus, the mean and variance of Y_i are

$$\mathrm{E}[Y_i] = \eta^{\lambda\beta}\Gamma(1+\lambda), \quad \mathrm{Var}[Y_i] = \eta^{2\lambda\beta}\left\{\Gamma(1+2\lambda) - [\Gamma(1+\lambda)]^2\right\}.$$

It can also be shown, using the results of [218] (Chap. 21), that

$$\mathrm{E}[MR_i] = 2\lambda\eta^{\lambda\beta}\Gamma(\lambda)\left(1 - \frac{1}{2^\lambda}\right).$$

Let η_S and β_S denote the stable-process parameters. The centerlines and control limits of the individuals Y and moving range MR charts are given by

$$CL_Y = \mathrm{E}[Y_i] = \eta_S^{\lambda\beta_S}\lambda\Gamma(\lambda),$$

$$UCL_Y = \eta_S^{\lambda\beta_S}\lambda\Gamma(\lambda)\left[1 - 5.319149\left(1 - \frac{1}{2^\lambda}\right)\right],$$

$$LCL_Y = \alpha^{\lambda\beta_S}\lambda\Gamma(\lambda)\left[1 + 5.319149\left(1 - \frac{1}{2^\lambda}\right)\right],$$

and

$$CL_{MR} = 2\eta_S^{\lambda\beta_S}\lambda\Gamma(\lambda)\left(1 - \frac{1}{2^\lambda}\right), \quad UCL_{MR} = 6.534\eta_S^{\lambda\beta_S}\lambda\Gamma(\lambda)\left(1 - \frac{1}{2^\lambda}\right).$$

There is no LCL for the MR chart because the lower 3-sigma limit is negative for λ values that yield close normal approximations. [45] considered $\lambda = 0.2777$ for which

$$CL_Y = 0.901119\eta_S^{0.2777\beta_S}, \; UCL_Y = 1.740382\eta_S^{0.2777\beta_S}, LCL_Y = 0.061856\eta_S^{0.2777\beta_S}$$

and

$$CL_{MR} = 0.315563\eta_S^{0.2777\beta_S}, \; UCL = 1.030944\eta_S^{0.2777\beta_S}.$$

[45] (in Tables IV–VII) assigned $\eta_S = 1$ and performed simulations to estimate the centerline and control limits. Hence, the UCL, CL, and LCL values found in the last four lines of their Table V (control limits for individual charts) are four sets of estimates of 1.740382, 0.901119, and 0.061856, respectively. The UCL and CL values in the last four lines of Table VII (control limits for moving range charts) are estimates of 1.030944 and 0.315563, respectively.

Given λ, $\mathrm{E}[Y_i]$ and $\mathrm{E}[MR_i]$ are directly proportional to each other. More specifically, the Y and MR charts, in essence, both monitor changes in the value of η^β. This could be troublesome for the practitioner because both η and β may shift significantly while changes in η^β may be relatively small. More importantly, changes in the Weibull mean and variance may not necessarily be suggested by changes in the value of η^β.

17.4.2 Run Length Properties of the Y and MR Charts

The run length L of a control chart is the number of sampling periods till the first OOC signal. Its expected value $\mathrm{E}[L]$ is known as the ARL. If the charted process statistics are independent from sample to sample, then L is a geometric random variable with success probability $p = \mathrm{P}\{\mathrm{OOC}\}$ and $ARL = 1/p$. This is true for the individual Y chart discussed above. On the other hand, this does not hold for the MR chart because the moving range spans two sampling periods and, hence, two consecutive moving ranges are not independent.

Let η_T and β_T be the true parameter values and assign $\rho_\eta = \eta_S/\eta_T$ and $\rho_\beta = \beta_S/\beta_T$. The Y chart has ARL given by

$$ARL^{-1}=\mathrm{P}\{\mathrm{OOC}\}=1-\exp\left\{-\rho_\eta^{\rho_\beta\beta_S}\left(LCL_Y/\eta_S^{\lambda\beta_S}\right)\right\}+\exp\left\{-\rho_\eta^{\rho_\beta\beta_S}\left(UCL_Y/\eta_S^{\lambda\beta_S}\right)\right\}.$$

Observe that ARL depends on ρ_η, ρ_β, and β_S because $LCL_Y/\eta_S^{\lambda\beta_S}$ and $UCL_Y/\eta_S^{\lambda\beta_S}$ do not depend on η_S. It suffices to assume for the discussions below that $\eta_S = 1$.

If the process is stable, i.e., $\rho_\eta = \rho_\beta = 1$, then $ARL \doteq 1462$ for $\lambda = 0.2777$. Figure 17.1 is a series of contour plots of ARL for $\beta_S = 0.5, 1, 2$. The solid lines indicate constant ARL curves. As expected the $ARL = 1462$ curve passes through the point $(\rho_\eta = 1, \rho_\beta = 1)$. The exponential (Rayleigh) case is given by $\beta_S = 1$ ($\beta_S = 2$), and, if the X_i stays within the exponential (Rayleigh) family, then the ARL values are given by the intersection of the contour curves with the horizontal line at $\rho_\beta = 1$. The curve denoted by "$- - -$" indicates the combinations of ρ_η and

ρ_β for which the Weibull mean is equal to the stable-process mean. Points to the left of this curve correspond to decreases from the target Weibull mean. The "$- \cdot -$" curve is when the Weibull variance is the same as the stable-process variance, and points to the left represent decreases from the target Weibull variance. A broader view of the exponential case ($\beta_S = 1$) is shown in Fig. 17.2 to show ARL behavior at extreme upward shifts in η.

The ARL for the MR chart is computed using a Fredholm integral equation of the second kind. The Appendix provides the details. ARL contour plots for the MR chart that are analogous to those of Fig. 17.1 are given in Fig. 17.3.

17.4.3 Discussion

For the (transformed individual) Y chart, we have the following observations:

- Recall that the region above (or to the left of) the equal-mean curve in Fig. 17.1 indicates increases from the stable-process Weibull mean. In general, the Y chart is effective in detecting increases in the mean because the ARL is mostly below the all-OK ARL of 1,462. It is also capable of detecting decreases in the mean around the region bounded by the equal-mean and the equal-variance curves in the first quadrant because, here, the ARL falls below the all-OK ARL. But it is ineffective in detecting decreases because ARL is above 1,462 in most cases.

- The region above the equal-variance curve corresponds to increases in the Weibull variance from the stable-process value. Because the equal-mean and equal-variance curves are both positively sloped, the conclusions in the previous item regarding the process mean is also true for the process variance. Thus, the Y chart is capable of detecting increases in the variance in general.

- A broader view of the exponential case is depicted in Fig. 17.2. This plot shows that for fixed ρ_β, the ARL eventually reaches a peak and decreases when ρ_η gets large enough. But, in this case, η_T has to be a lot smaller than η_S for the ARL to fall below 1,462. The ARL drops more quickly for larger β_S. Still, the inability to detect smaller changes can be a cause for concern.

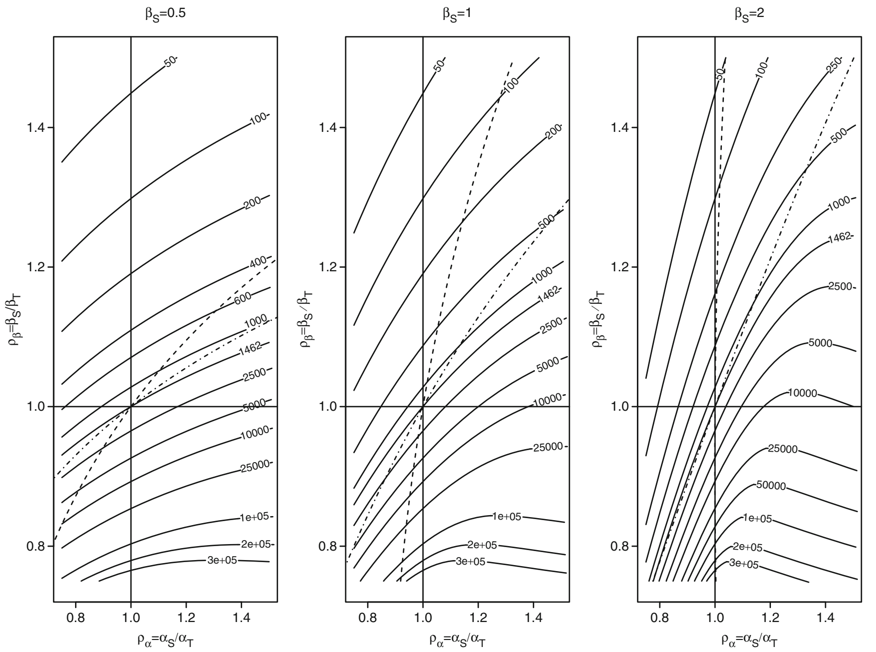


Figure 17.1: ARL contour of Y chart for $\lambda = 0.2777$, $\beta_S = 0.5, 1, 2$

- In summary, the Y chart is capable of detecting increases in both mean and variance. However, this is disconcerting particularly for reliability engineers because a decrease in, say, average product life or material strength is probably more detrimental than an increase in the variance.

For the MR chart, we have the following observations:

- The ARL trends as shown in Fig. 17.3 are similar to those for the Y chart, and remarks made above regarding the Y chart also apply to the MR chart. The MR chart is capable of detecting increases in Weibull mean and variance which may be problematic to reliability engineers.

- On the other hand, the MR chart's ARL does not exhibit the behavior of reaching a peak and coming back down as in the Y charts even for extremely large values of ρ_η. This suggests that using the MR chart may be more problematic than using the Y chart.

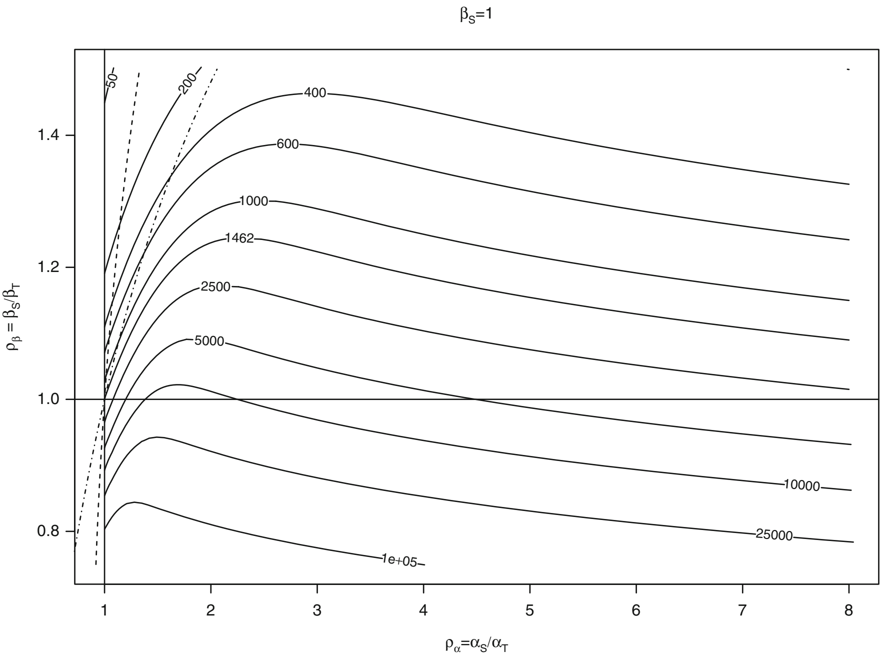


Figure 17.2: ARL of the exponential Y chart for $\lambda = 0.2777$ and different values of ρ_η and ρ_β

The above results agree with [45]'s recommendation not to use either Y or MR charts as described above in general. Below, a method simultaneously using individual and moving ratio charts of Weibull data is proposed. This method avoids the ARL and shift detection problems encountered with the Y and MR charts.

17.5 Individual and Moving Ratio Charts for Weibull Distribution

The objective here is to develop monitoring schemes that declare that a Weibull process is OOC when a shift in either the scale or shape parameter has occurred. For this, charts for the individual value (X) and the moving ratio (R) are proposed. Distributional properties of X and R are used to determine chart limits and run length properties without using simulations.

Let η_S and β_S be the stable-process parameter values. The individual (X) chart is a time-series plot of $X_1, X_2, \ldots$ with centerline CL_X,

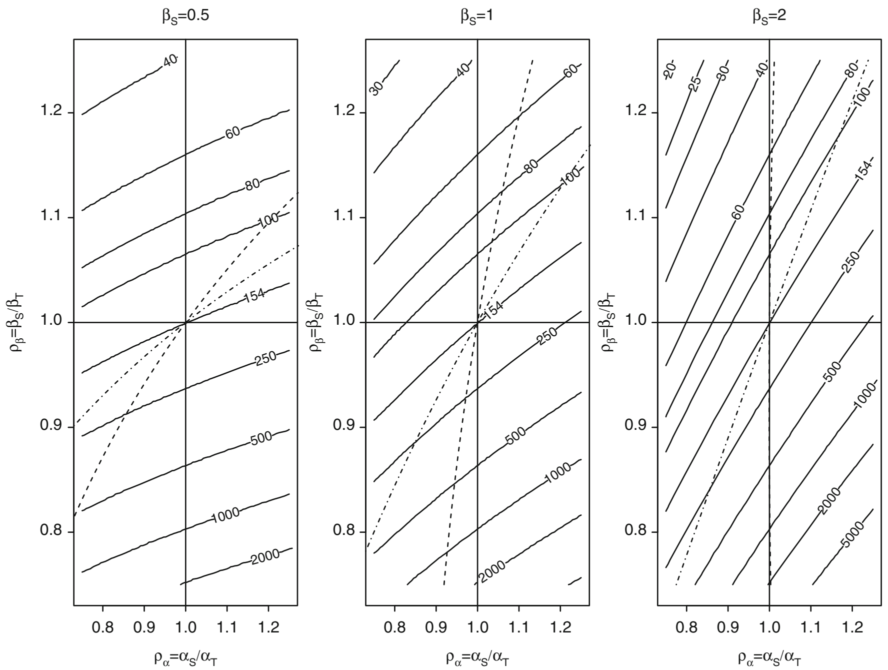


Figure 17.3: ARL of MR chart for $\lambda = 0.2777$, $\beta_S = 0.5, 1, 2$ and different values of ρ_η and ρ_β

upper control limits UCL_X, and lower control limits LCL_X given by

$$CL_X = \eta_S[-\log(0.5)]^{1/\beta_S}, \quad UCL_X{=}\eta_S(u_x)^{1/\beta_S}, \quad LCL_X{=}\eta_S(l_x)^{1/\beta_S} \tag{17.5.1}$$

so that an OOC signal occurs if $X_i > UCL_X$ or $X_i < LCL_X$ for $i = 1, 2, \ldots$. The centerline is at the median of WEI(η_S, β_S) but the mean given by Eq. (17.3.2) may be used as an alternative value. The constants l_x and u_x are the control limits when $\eta_S = \beta_S = 1$ and, hence, are referred to as the standardized X control limits. Also, UCL_X and LCL_X are the $(1 - e^{-u_x})$ and $(1 - e^{-l_x})$ quantiles, respectively, of WEI(η_S, β_S).

The moving ratio (R) chart is a time-series plot of $R_2, R_3, \ldots$ with centerline and control limits given by

$$CL_R = 3^{1/\beta_S}, \quad UCL_R = (u_r)^{1/\beta_S}, \quad LCL_R = (l_r)^{1/\beta_S} \tag{17.5.2}$$

so that an OOC signal occurs when $R_i > UCL_R$ or $R_i < LCL_R$ for $i = 2, 3, \ldots$. Equation (17.3.3) shows that the above chart values are, respectively, the 0.50, $[(u_r-1)/(u_r+1)]$, and $(l_r-1)/(l_r+1)$ quantiles of R_i under a stable process. The constants l_r and u_r are referred to as the standardized R chart control limits because they are the chart constants when $\beta_S = 1$.

The X chart is appropriate for monitoring changes in the process mean if the practitioner can assume that the shape parameter β is constant and stable. However, the X chart is not adequate to monitor changes in both parameters. Equation (17.3.2) suggests that there are infinitely many combinations of η and β for which the mean is constant. Furthermore, a constant mean does not necessarily mean a constant variance, and vice versa.

Suppose that the shift in η occurs immediately before implementing the R chart. The process statistic R does not depend on η because the effect of a scale parameter is multiplicative and taking the ratio of two measurements eliminates that effect. So, in this case, the R chart is not appropriate for monitoring shifts in both η and β. For this reason, the application of the R chart alone is not studied below. Instead, the reader is referred to [367].

If both scale and shape parameters shift over time, the above discussions do not suggest implementing the X and R charts separately. Instead, it makes intuitive sense to use these charts simultaneously to monitor a Weibull process. Below, the properties of the combined application of the X and R charts are investigated and evaluated with respect to run length.

17.6 Average Run Length and Unbiasedness

The number of periods L till the first OOC signal is called the run length of a control chart. Its expected value called the ARL is often used in the literature to study and compare the performance of control chart schemes.

17.6.1 Average Run Length

Let η_T and β_T be the true parameter values. If the process is stable, i.e., $\eta_T = \eta_S$ and $\beta_T = \beta_S$, the ARL is referred to as the stable-process ARL. Otherwise, the ARL is called an OOC ARL. Let L_0 denote the desired stable-process ARL. The X and R chart parameters l_x, u_x, l_r, and u_r are chosen so that $ARL = L_0$ when the process is stable. To describe the respective changes in parameters, define

$$\rho = \frac{\beta_T}{\beta_S}, \quad \Delta = \left(\frac{\eta_T}{\eta_S}\right)^{\beta_S}.$$

The percent changes in the scale η and shape β are given by $100[1 - \exp(\Delta/\beta_S)]\%$ and $100(\rho - 1)\%$, respectively.

For the combined X/R chart, the ARL is approximated using Markov chains. See the Appendix for details. A careful inspection of the results therein reveals that the required transition probabilities depend on ρ and Δ. Consequently, the ARL is uniquely determined by ρ and Δ as in the case of the X chart. This suggests that to study the ARL performance of a specific X/R chart, it suffices to assume that $\eta_S = \beta_S = 1$, and calculate the ARL under different combinations of ρ and Δ.

17.6.2 *ARL* Unbiasedness

It is possible that when the process is OOC, the ARL is actually longer than the desired stable-process ARL L_0. This produces an undesirable result because an OOC situation must be detected quickly. To address this issue, [376] studied the concept of ARL-unbiasedness. For the X/MR chart proposed here, this means finding chart parameters l_x, u_x, l_r, and u_r so that

- $ARL = L_0$ when $\rho = 1$ and $\Delta = 0$ (i.e., the process is stable)

- $ARL < L_0$ when $\rho \neq 1$ or $\Delta \neq 0$ (i.e., the process has shifted)

Below, the performance of the X/R charts in detecting shifts in both η and β is studied. All the X/R charts presented below are unbiased for shifts in both η and β. Unbiasedness is verified through contour plots of the ARL.

17.7 Numerical Results

In this section, standardized control limits for X/R charts are tabulated. Computations of the limits are carried out using computer programs written in the R language of [380]. The programs are available from the author.

17.7.1 Control Limits for the Combined X/R Charts

Table 17.1 gives the natural logarithms of the X/R standardized limits for desired all-OK ARL of $L_0 = 50$, 100, 200, 300, 400, and 500. These values are used with Eqs. (17.5.1) and (17.5.2) to compute the control limits in actual applications. Observe that $|l_x| \neq |u_x|$ and $|l_r| \neq |u_r|$ suggest that using just one constant (e.g., constant=3 for the normal case) to determine control limits for X_i is not appropriate for Weibull individual values or moving ranges.

Table 17.1: Standardized control limits for ARL-unbiased X/R charts

Control limits	L_0 50	100	200	300	400	500
$\log(l_x)$	−6.2955	−7.1106	−7.9180	−8.5834	−8.5824	−8.6850
$\log(u_x)$	2.1883	2.2930	2.3877	2.4919	2.4312	2.4165
$\log(l_r)$	0.0360	0.0180	0.0090	0.0060	0.0046	0.0036
$\log(u_r)$	6.7789	7.5064	8.2291	8.5070	9.1708	9.8827

17.7.2 ARL-Unbiasedness of X/R Charts

Unbiasedness of X/R charts can be checked graphically. For the application in Sect. 17.8 below, stable-process parameters are $\eta_S = 3.2$ and $\beta_S = 4.8$, and the X/R chart control limits are $LCL_X = 0.7274$, $UCL_X = 5.1596$, $LCL_R = 1.0038$, and $UCL_R = 4.7771$ so that the stable-process ARL is 100. Figure 17.4 is a contour plot of ARL values under this chart and for different combinations of true values for η and β. The plot suggests ARL unbiasedness because $ARL < 100$ when at least one of $\eta \neq 3.2$ and $\beta \neq 4.8$ is true. Similar graphs, not included here, also suggest unbiasedness for other values of L_0 in Table 17.1.

17.7.3 Sample Size Requirements for Phase I

With respect to how process parameters are obtained, there are two types of control charts, namely, standards-given and retrospective charts. See [460]. In standards-given charts, stable-process parameters are known, e.g., from prior experience with the process or engineering judgment. In this case, both parameters and control limits are fixed quantities. In retrospective charts, process parameters and control limits are estimated because, for example, the process may be related to a new product design or a pilot study. Thus, stable-process data

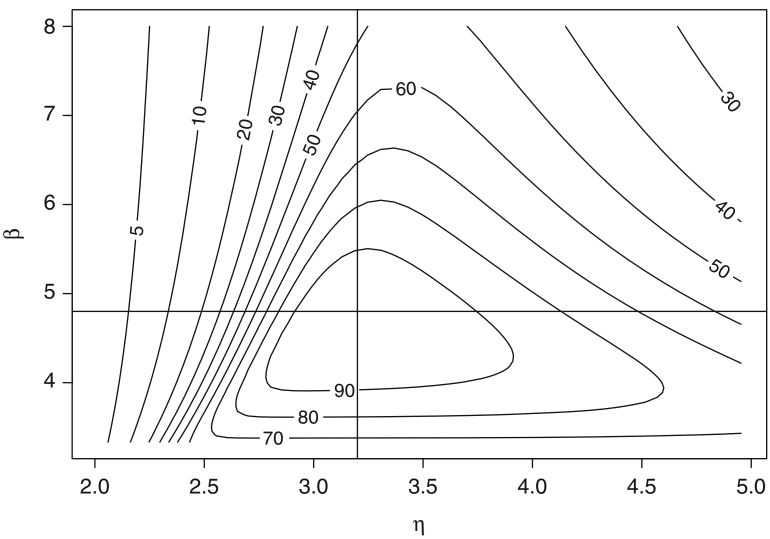


Figure 17.4: Contour plot of ARL for the unbiased X/R chart with $L_0 = 100$

from m periods are used to estimate process parameters from which control limits are derived. The prior m periods are referred to as Phase I. The subsequent monitoring of the process using the estimated control limits is referred to as Phase II. For retrospective charts, control limits are, in reality, random variables.

The respective run length distributions under the standards-given and retrospective scenarios are expected to be different. Thus, it is important to know how large m should be so that the run length distributions of standards-given and retrospective charts are reasonably similar. Simulations are performed to compare the empirical distributions of run lengths under stable-process conditions for different values of m. Maximum likelihood methods are used to estimate process parameters.

Based on the discussion of Sect. 17.6, it can be assumed without loss of generality that $\eta_S = 1$ and $\beta_S = 1$ for studying the properties of run lengths for X/R charts. Simulation results suggest that run length distributions of standards-given and retrospective charts are similar for $m \geq 150$ for the values of L_0 considered here.

17.8 An Application

In this section, the proposed X/R charts are applied to a data set on the strength of carbon fibers. Padgett and Spurrier [363] presented data on the strengths of the carbon fibers. The data set consisted of samples of size 5 from 20 inspection periods. It was known that the first 10 periods were obtained from a stable process, while the later 10 periods were from a process that had shifted. The stable-process distribution for carbon-fiber strengths was WEI($\eta = 3.2, \beta = 4.8$). Thus, $\eta_S = 3.2$ and $\beta_S = 4.8$. One observation is randomly chosen from each of the 20 sampling periods. These values and the corresponding moving ratio values are given under the X and R columns, respectively, of Table 17.2. The proposed method is applied to this data subset.

Table 17.2: Individual observations and moving ratio for strength distribution of carbon fibers

Period	X	R	Period	X	R
1	2.74	NA	11	1.36	1.5956
2	3.11	1.1350	12	3.68	2.7059
3	3.19	1.0257	13	1.73	2.1272
4	1.87	1.7059	14	1.71	1.0117
5	2.97	1.5882	15	1.18	1.4492
6	2.93	1.0137	16	4.38	3.7119
7	2.55	1.1490	17	0.39	11.2308
8	2.85	1.1176	18	4.70	12.0513
9	2.35	1.2128	19	2.03	2.3153
10	2.17	1.0829	20	3.65	1.7980

Suppose that the desired stable-process ARL is $L_0 = 100$. The control limits for $L_0 = 100$ in Table 17.1 yield $LCL_X = 0.7274$, $UCL_X = 5.1596$, $LCL_R = 1.0038$, and $UCL_R = 4.7771$ for the X/R chart. Figure 17.5 gives the X/R charts for the carbon-fiber data. Monitoring starts in Period 1. It is discernible from a quick inspection of each of the respective time-series plots of individual and moving ratio values that there is an increased variance in observations in the later 10 periods. The X/R charts signal OOC in period 17 (7 periods after shift occurs). These charts suggest a shift in, at least, the shape parameter β. Table 17.3 gives the different OOC ARL values for this application and different target stable-process ARL L_0.

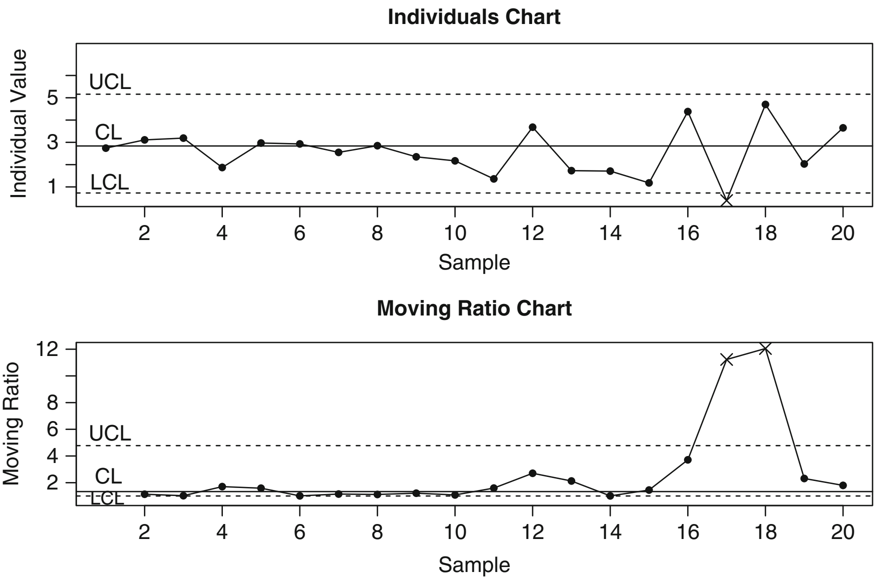


Figure 17.5: X/R chart with stable-process ARL $L_0 = 100$ for the strengths of carbon fibers

Table 17.3: ARL of X/R charts for the shift in carbon-fiber strength distribution

	L_0					
	50	100	200	300	400	500
ARL	20.2	35.5	62.15	92.7	101.9	112.7

17.9 Conclusion

The Weibull is a relevant distribution in many practical situations such as time-to-event applications, engineering reliability, and survival analysis. It is not recommended to apply normal-based individual and moving range charts using the transformation $X^{0.2777\beta}$ to Weibull data. The authors of [45] performed simulations that illustrated problems with this approach. This chapter provides theoretical results that agree with their findings. This chapter also presents ARL-unbiased combined individual and moving ratio X/R charts when the process is described by a Weibull distribution. The proposed X/R charts provide the practitioner a method of monitoring a Weibull process when shifts in either scale or shape are possible.

17.10 Appendix

17.10.1 *ARL* Computation for the Moving Range Chart

Let $X_i \sim \mathrm{WEI}(\eta_T, \beta_T)$, $\lambda > 0$, and $Y_i = X_i^{\lambda\beta_S}$. Then, Y_i is a $\mathrm{WEI}(\eta_T^{\lambda\beta_S}, 1/(\lambda\rho_\beta))$ variable. The moving range is given by $MR_i = |Y_i - Y_{i-1}|$ for $i = 2, 3, \ldots$. The MR chart signals OOC in period $i > 1$ if $MR_i > UCL_{MR} = 6.534\eta_S^{\lambda\beta_S}\lambda\Gamma(\lambda)(1 - 1/2^\lambda)$. Based on this, there is no OOC signal in period i if $\max\{0, Y_{i-1} - UCL_{MR}\} \leq Y_i \leq Y_{i-1} + UCL_{MR}$. Let $L(y_{i-1})$ be the expected number of observations to an OOC signal when the current value is y_{i-1}. For simplicity, assign $u = UCL_{MR}$. Table 17.4 enumerates the possible values of $L(y_{i-1})$ given the next individual value y_i. Thus, if $f(x)$ is the pdf of $\mathrm{WEI}(\eta_T^{\lambda\beta_S}, 1/(\lambda\rho_\beta))$, then

$$
\begin{aligned}
L(y_{i-1}) &= \int_0^{\max\{0, y_{i-1}+u\}} 1 f(x)\mathrm{d}x + \int_{\max\{0, y_{i-1}+u\}}^{y_{i-1}+u} [1 + L(x)] f(x)\mathrm{d}x + \int_{y_{i-1}+u}^{\infty} Q1 f(x)\mathrm{d}x \\
&= 1 + \int_{\max\{0, y_{i-1}+u\}}^{y_{i-1}+u} L(x) f(x)\mathrm{d}x
\end{aligned}
$$

This is a Fredholm integral equation of the second kind. The values of $L(y)$ can be approximated using the Fortran codes in [378]. The earliest possible OOC signal is when $i = 2$. Thus, it seems sensible to approximate the ARL of the MR chart by $L(\eta_T^{\lambda\beta_S}(-\log(0.5))^{\lambda\rho_\beta})$, i.e., we substitute the median of $\mathrm{WEI}(\eta_T^{\lambda\beta_S}, 1/(\lambda\rho_\beta))$ for the first observation y_1.

Table 17.4: Expected number of observations till an OOC signal

Next value $Y_i = y_i$	Expected number of steps to OOC
$y_i > u + y_{n-1}$	1
$0 < y_i < \max\{0, y_{i-1} + u\}$	1
$\max\{0, Y_{i-1} - u\} \leq Y_i \leq Y_{i-1} + u$	$1 + \mathrm{E}[L(y_i)]$

17.10.2 *X/R* ARL Computation by Markov Chains

Consider the X/R scheme described in Sect. 17.5. Let n denote the next period with response X_n. For the current period $n - 1$, let x_{n-1} denote the response such that $\eta_S(l_x)^{1/\beta_S} \leq x_{n-1} \leq \eta_S(u_x)^{1/\beta_S}$. Let NS (no signal) be the complement of OOC, i.e., NS is the event that OOC is not signaled in period n. Define the following quantities:

$$a_1 = \max\{x_{n-1}(u_r)^{-1/\beta_S}, \eta_S(l_x)^{1/\beta_S}\}, \quad a_2 = \min\{x_{n-1}(l_r)^{-1/\beta_S}, \eta_S(u_x)^{1/\beta_S}\},$$

$$b_1 = \max\{x_{n-1}(l_r)^{-1/\beta_S}, \eta_S(l_x)^{1/\beta_S}\}, \quad b_2 = \min\{x_{n-1}(u_r)^{-1/\beta_S}, \eta_S(u_x)^{1/\beta_S}\}.$$

Define the following intervals:

$$A = \begin{cases} [a_1, a_2] & \text{if } a_1 < a_2 \\ \phi & \text{otherwise} \end{cases}, \qquad B = \begin{cases} [b_1, b_2] & \text{if } b_1 < b_2 \\ \phi & \text{otherwise} \end{cases}.$$

It can be shown that $\mathrm{NS} = \{X_n \in A \cup B\}$.

Divide the interval $[\eta_S(l_x)^{1/\beta_S}, \eta_S(u_x)^{1/\beta_S}]$ into t intervals so that interval i is given by $(L_i, U_i]$ where

$$L_i = \eta_S(l_x)^{1/\beta_S} \left(\frac{u_x}{l_x}\right)^{\frac{i-1}{t\beta_S}}, \quad U_i = \eta_S(l_x)^{1/\beta_S} \left(\frac{u_x}{l_x}\right)^{\frac{i}{t\beta_S}}$$

for $i = 1, 2, \ldots, t$. Include L_1 in the first interval, that is, the first interval is closed on both ends. If X_n falls in interval i, let X_n assume the value of the geometric mean $m_i = \sqrt{L_i U_i}$, i.e., $X_n = m_i$.

For the Markov chain, define the following non-OOC states:

$$M_0 = \text{The state prior to period 1}, \quad M_i = \{X_n = m_i\}$$

for $i = 1, 2, \ldots, t$. The absorbing state is $M_{t+1} = \{\mathrm{OOC}\}$. Let p_{ij} be the transition probability that the next state is M_j given that the current state is M_i, that is, $p_{ij} = \mathrm{P}\{M_j | M_i\}$. Suppose that the true state is $Y_i \sim \mathrm{WEI}(\eta_T, \beta_T)$. The cdf of Y_i is $F_X(x; \eta_T, \beta_T)$ given by Eq. (17.3.1). Then,

$$\begin{aligned} p_{i0} &= 0 \\ p_{0j} &= F_X(u_j; \eta_T, \beta_T) - F_X(l_j; \eta_T, \beta_T) \end{aligned}$$

for $i = 0, 1, \ldots, t$ and $j = 1, 2, \ldots t$. Furthermore, for $i, j = 1, 2, \ldots, t$

$$p_{ij} = \mathrm{P}\{X_n \in (L_j, U_j] \cap \mathrm{NS}\}$$

which can be evaluated using the cdf F_X.

Let $\boldsymbol{R}$ be the $(t+1) \times (t+1)$ matrix whose (i, j) element is given by $p_{i,j}$ for $i, j = 0, 1, \ldots, t$. Define the vector

$$\boldsymbol{L} = [\boldsymbol{I} - \boldsymbol{R}]^{-1} \times \mathbf{1}$$

where $\boldsymbol{I}$ is the $(t+1) \times (t+1)$ identity matrix and $\mathbf{1}$ is the vector of 1's of length $t+1$. Element i of $\boldsymbol{L}$ is the expected number of periods to

OOC if the process is currently in state M_i. Thus, the all-OK ARL is the first element of $\boldsymbol{L}$. Computer codes written in the R language to carry out necessary computations are available from the author. For this chapter, simulations of run lengths suggest that $t = 50$ is adequate for approximating the ARL.

Chapter 18
On a Slow Server Problem

Vladimir Rykov

Abstract: The slow server problem is generalized for the case of additional cost structure. With the help of special partial ordering of the system state space it is shown that the optimal policy for the problem has a monotone property consisting in the following: an additional server should be switched on only in the case if the queue length exceeds some level depending on the system state, and in this case the server with minimal service cost should be used.

18.1 Introduction and Motivation

This chapter focuses on the meaning of usual ordering for stochastic systems control problems decision. It will be done with the example of a slow server problem (SSP). For a queueing system with heterogeneous servers there exists a problem of the sojourn time of customers in the system minimization. The problem can be considered and solved with special ordering in the system state space. There exists some special partial order on the phase space of the system, in which the

V. Rykov (✉)
Department of Applied Mathematics and Computational Modeling,
Gubkin Russian State University of Oil and Gas,
Moscow 119991, Leninsky prosp. 65, Russia
e-mail: vladimir_rykov@mail.ru

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_18,

optimal control policy possesses some special monotonicity properties that allows to describe its characterization and simplify their real construction.

The SSP has an enough long history that consists of the following. Using the traditional *conservative* service discipline in queueing systems (QS) with *heterogeneous* servers is *noneffective* because it leads to increasing the queue length and waiting time. This stimulates searching of the system service rules in order to minimize the *mean sojourn time*.

The SSP is interesting both from *theoretical* and *practical* points of view.

- Theoretically the problem is based on the theory of controllable queueing systems (see, e.g., Kitaev and Rykov [239] and Sennott [416]) and also introduces the new direction in the theory–investigation of *qualitative properties* of optimal control policies (Rykov [402]).

- In practice the knowledge of qualitative properties of optimal policies allows to *essentially simplify* their real construction.

- The last circumstance allows to use these models in different applications especially in telecommunications (see Pedro [369], Vishnevsky, and Semenova [463]).

The SSP has been initially stated and considered by Krishnamoorthy [268] in 1963 for the system with two servers. The problem for two servers has been studied in more details by Hajek [187] and by Lin and Kumar [282], where the monotonicity of optimal policy has been proved. Koole [265] in 1995 proposed a simplified proof with the same result also for two servers, and Weber [471] in 1993 formulated a conjecture that the result is true in general case. An appropriate solution has been provided by Rykov [403], where the suggested condition of the optimal service rule stability has been omitted, leading to incompleteness of the formal proof of the result. It has been noted in Verycourt and Zhou [462].

In Rykov and Efrosinin [404], it was shown that the monotonicity property of optimal rule takes also place for the queueing system with heterogeneous servers with additional structure of penalties for busy servers and customers waiting. In Efrosinin and Breuer [144] and Efrosinin [143], it was shown numerically that the monotonicity property of optimal service rule preserves also for retrial service systems and for the systems with PH-distributed inter-arrival times. An improved proof for the generalized case of the queueing system with heterogeneous servers with respect to mean lost minimization has been obtained in Rykov and Efrosinin [405].

In this chapter these results will be summarized. The chapter is organized as follows: in the next section the problem formulation is presented. The optimality equation is discussed in the third section and its transformation to more convenient form for the problem solution is detailed in the fourth section. The main result of the chapter, a theorem about monotonicity properties of optimal policies, is presented in the fifth section. The chapter is concluded with some new problems formulation.

18.2 Problem Formulation

Consider an $M/M/K/N-K$ $(K \leq N \leq \infty)$ QS, presented in Fig. 18.1, with:

- K heterogeneous exponential servers with intensities μ_k, $1 \leq k \leq K$
- $N - K$ places in the buffer
- A Poisson input with intensity λ
- A cost $C^Q(q) = c_0 q$ for q customers holding in the queue
- A cost $C^U(\mu_k) = c_k, \quad 1 \leq k \leq K$, for using the k-th server with intensity μ_k per unit of time

In the case of infinite buffers for existence of the stationary regime it is supposed that

$$\lambda < \sum_{1 \leq k \leq K} \mu_k \equiv M.$$

The control consists of switching the servers on and off at the control times S_n that coincide with the *arrival* and *service completion* times. The goal is to *minimize* the average long-run working cost of the system per unit of time. For example, for the problem of minimizing the mean number of customers in the system, it is necessary to put

$$c_0 = c_k = 1. \tag{18.2.1}$$

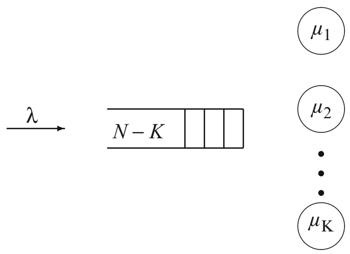


Figure 18.1: Multi-server queueing system with heterogeneous servers

For modeling of the system operation we consider the controllable process

$$\{Z(t)\} = \{(X(t), U(t))\}$$

with the observed process

$$\{X(t)\} = \{(Q(t), D(t))\},$$

where $Q(t)$ is the queue length at time t and $D(t) = (D_1(t), \dots, D_K(t))$ describes the states of servers at this time,

$$D_i(t) = \begin{cases} 0, & \text{when the } i\text{-th server is idle at the time } t \text{ and} \\ 1, & \text{otherwise.} \end{cases}$$

Denote the state space of the observed process by

$$E = \mathbb{N} \times \{0, 1\}^K$$

with the denumerable set $\mathbb{N}$. For each state $x = (q, d_1, \dots, d_K)$ denote by

$$J_0(x) = \{j : d_j(x) = 0\}, \qquad J_1(x) = \{j : d_j(x) = 1\}$$

the sets of indices j for which components $d_j = 0$ or $d_j = 1$, respectively.

As a controlling process, consider the process $\{U(t) : \ t \geq 0\}$, which takes their values in the sets $A(x)$, where

$$A(x) = \begin{cases} J_0(x) \cup \{0\} & \text{for } x \text{ with } q(x) < N, \\ J_0(x) & \text{for } x \text{ with } q(x) = N. \end{cases}$$

Its values $a = k$ denote the server's number k which must be switched on or do not do it if $a = 0$ at the nearest to t control time.

Thus for this model the decision set $A = \{0, 1, \dots, K\}$ consists of $K + 1$ points, and control $a = 0$ denotes do not switch on any server, while control $a = k$ denotes to switch on k-th server in times of customers arrival and service completion.

Under the considered assumptions, the process

$$\{Z(t)\} = \{(X(t), U(t))\}$$

is a Markov decision one with phase space $E = \mathbb{N} \times \{0, 1\}^K$ and a control space $A(x)$ depending on the state $x \in E$.

Consider the shift operators S_0, S_j on the phase space E,

$$S_0 x = x + e_0 1_{\{q(x) < N\}}, \qquad S_j x = x + e_j 1_{\{j \in J_0(x)\}}, \tag{18.2.2}$$

where e_i is $(K+1)$-dimensional vector with the i-th component being one (beginning from 0-th), and zero otherwise. Denote by

$$S_j^{-1} x, \ (j = 0, 1, \dots, K)$$

the inverse operators for such points $x \in E$, for which it exists and put $S_j^{-1} x = x$ in another cases, i.e.,

$$S_0^{-1} x = x, \quad \text{if} \quad q(x) = 0, \qquad S_j^{-1} x = x, \quad \text{if} \quad j \in J_0(x). \tag{18.2.3}$$

Put

$$G_j = S_0^{-1} S_j^{-1}$$

and spread out the act of operators S_j for the functions $h(\cdot)$ on E, as usual by the relations

$$S_j h(x) = \begin{cases} h(S_j x), & \text{for } j \in \{0\} \cup J_0(x), \\ h(x), & \text{for } j \in J_1(x). \end{cases} \tag{18.2.4}$$

With the above notations, the transition intensities of the process $\{Z(t)\}$ get the form for $a \in A(x)$

$$\lambda_{xy}(a) = \begin{cases} \lambda, & \text{for } \ y = S_a x, \\ \mu_j, & \text{for } \ y = S_a G_j x, \quad j \in J_1(x), \\ 0, & \text{otherwise}. \end{cases}$$

In accordance with the given cost structure, the objective (loss) functional has the form

$$Y(t) = \int_0^t \left[c_0 Q(u) + \sum_{1 \le k \le K} c_k D_k(u) \right] \mathrm{d}u.$$

As usual (see, e.g., Kitaev and Rykov [239]), define

- A strategy δ
- The probability distribution P_x^δ of the process $\{Z(t)\}$ given initial state x and strategy δ
- The expectation E_x^δ with respect to this probability distribution
- The expectation E_x^a on the control interval (transition period) $T_n = S_{n+1} - S_n$ given initial state x and decision a

18.3 Optimality Equation

Denote by:

- $c(x) = c_0 q + \sum_{j \in J_1(x)} c_j$ the loss rate at the state x.
- $M_1(x) = \sum_{j \in J_1(x)} \mu_j$ the total intensity of service completion in the state x

In particular for the problem of minimizing the mean number of customers in the system, i.e., in the case Eq. (18.2.1), one has

$$c(x) = q + \sum_{1 \le j \le K} d_j(x) = l(x),$$

where $l(x)$ is the number of customers in the system when it is in the state x.

It is well known (see, e.g., [239, 271, 416]) that for the problem under consideration, the *optimality principle* is valid. This means that:

(i) The approximation

$$\inf_\delta \mathrm{E}_x^\delta [Y(t)] \approx gt + v(x) + o(1)$$

takes place, where:

(ii) The *price* of the model (minimal mean service cost per unit of time)

$$g = \inf_{\delta} \lim_{t\to\infty} \frac{1}{t} \mathrm{E}_x^{\delta}[(t)] \tag{18.3.1}$$

jointly with the *value function* of the model

$$v = \{v(x) : \ x \in E\} : \ E \to \mathbb{R}_+$$

exist.

(iii) They satisfy the *optimality equation*

$$v(x) = \min_{a\in A(x)} \frac{c(x) - g + \lambda v(S_a x) + \sum_{1\le j\in J_1(x)} \mu_j v(S_a G_j x)}{\lambda + M_1(x)}. \tag{18.3.2}$$

(iv) Also *optimal strategy* can be chosen as a stationary Markov one, i.e., it is determined by the *optimal policy* $f = \{f(x) : x \in E\}$ with

$$f(x) = \operatorname*{argmin}_{a\in A(x)} \frac{c(x) - g + \lambda v(S_a x) + \sum_{1\le j\in J_1(x)} \mu_j v(S_a G_j x)}{\lambda + M_1(x)}. \tag{18.3.3}$$

Based on the optimality equation (18.3.2), we investigate some qualitative properties of the optimal policy (18.3.3). To do that, we first transform the optimality equations to a more convenient form.

18.4 Transformation of Optimality Equations

Multiply each side of Eq. (18.3.2) by

$$\lambda + M_1(x)$$

and add to each side the term

$$v(x) \sum_{j\in J_0(x)} \mu_j$$

to obtain

$$(\lambda + M)v(x) = \min_{a\in A(x)} \left[c(x) - g + \lambda v(S_a x) + \sum_{1\le j\le K} \mu_j v(S_a G_j x) \right],$$

where

$$M = \sum_{1 \le j \le K} \mu_j$$

and accordingly to Eq. (18.2.4) for $j \in J_0(x)$,

$$v(S_a G_j x) = v(x).$$

Using a scale of time as $(\lambda + M) = 1$ produces the equation

$$v(x) = \min_{a \in A(x)} \left[c(x) - g + \lambda v(S_a x) + \sum_{1 \le j \le K} \mu_j v(S_a G_j x) \right]. \quad (18.4.1)$$

With the help of the operators

$$T_0 v(x) = \min[v(S_k x) : \quad k \in A(x)], \quad (18.4.2)$$

$$T_j v(x) = \begin{cases} T_0 v(G_j x) & \text{for} \ \ j \in J_1(x), \quad q(x) > 0, \\ v(S_j^{-1} x) & \text{for} \ \ j \in J_1(x), \quad q(x) = 0, \\ v(x) & \text{for} \ \ j \in J_0(x) \end{cases} \quad (18.4.3)$$

the latter equation can be represented in the form

$$v(x) = c(x) + \lambda T_0 w(x) + \sum_{l \in J_1(x)} \mu_l T_l v(x) + \sum_{l \in J_0(x)} \mu_l v(x) - g = Bv(x), \quad (18.4.4)$$

where dynamic programming operator $Bv(x)$ is determined with this equality.

The above argumentations can be summarized in the following theorems.

Theorem 18.4.1. *Equations (18.3.2), (18.4.1), and (18.4.4) are equivalent in the sense that their prices, value functions and policies are identical.*

Theorem 18.4.2. *Optimal policy $f : E \to A$ is determined with the model value function $v(x)$ as follows:*

$$f(x) = \operatorname{argmin}\{v(S_k x) : \ k \in A(x)\}. \quad (18.4.5)$$

Remark 18.4.3. The last statement shows that:

- It is enough to study decisions only at the arrival times, because the decisions at the service completion times coincide with appropriate decisions at the arrival times in appropriated shifted states.

- The optimal policy is determined with the function

$$b(x,k) = v(S_k x), \tag{18.4.6}$$

  which we refer to as the *Bellman function* of the model.

It depends on the value function of the model $v(x)$ and is simple enough that allows to investigate the qualitative properties of the model in terms of this function.

18.5 Monotonicity of Optimal Policies

In order to investigate the optimal policy properties, enumerate the servers in the increasing order of their mean full service costs, i.e.,

$$0 \leq \frac{c_1}{\mu_1} \leq \frac{c_2}{\mu_2} \leq \cdots \leq \frac{c_K}{\mu_K}, \tag{18.5.1}$$

and suppose that the following condition holds:

$$0 \leq \frac{1}{\mu_1} \leq \frac{1}{\mu_2} \leq \cdots \leq \frac{1}{\mu_K}. \tag{18.5.2}$$

In the same order arrange the components of the vector $d = (d_1, \ldots, d_K)$. This arrangement determines the usual total order in the decision set A, where $0 \leq 1 \leq \cdots \leq K$.

Define a partial order in $E = \mathbb{N} \times \{0,1\}^K$ with the help of the shift operators S_0 and S_j by the following relations:

$$S_0 x \geq x, \qquad S_j x \geq x \quad \text{for all } j \in J_0(x),$$

$$S_i x \geq S_j x \quad \text{for all } i,j \in J_0(x) \quad \text{with } i \geq j \ (c_i \mu_i^{-1} \geq c_j \mu_j^{-1})$$

One can see that these relations, in fact, determine some partial order on the set E, in which, however, the points $S_0 x$ and $S_j x$, $j \neq 0$ are not comparable.

The main result consists of the following main theorem that allows to produce some qualitative properties of any optimal policy, and for which we need an additional definition.

Definition 18.5.1. Define the state $x \in E$ of the system as a *stable* one with respect to the optimal policy if it does not demand to send the customers to any free server from the queue, formally

$$v(x) = \min_{k \in A_0(x)} v(S_0^{-1} S_k x). \tag{18.5.3}$$

Theorem 18.5.2. *The value function of the model $v : E \to \mathbb{R}_+$ possesses the following properties for all stable states $x \in E$ of the system:*

C1. Non-decreasing property:

$$v(x) + \frac{c_i}{\mu_i} \le v(S_i x), \quad v(S_j x) \le v(S_i x),\ i, j \in J_0(x).$$

C2. Super-modularity property:

$$v(S_0 x) - v(x) \le v(S_0 S_i x) - v(S_i x),\ i \in J_0(x).$$

C3. Super-convexity property:

$$v(S_0 x) - v(S_i x) \le v(S_0^2 x) - v(S_0 S_i x),\ i \in J_0(x).$$

C4. Convexity (with respect to the shift S_0) property:

$$2v(S_0 x) \le v(x) + v(S_0^2 x).$$

The proof of the theorem is based on several lemmas.

Lemma 18.5.3. *The operator T_0, determined by Eq. (18.4.2), preserves the properties (C1—C4) of functions for all stable states of the system.*

Lemma 18.5.4. *The operators T_l, determined by Eq. (18.4.3), preserve the properties (C1–C4) of functions if the l-th server is busy for all stable states of the system contained in the inequalities.*

Lemma 18.5.5. *The operator B, determined by Eq. (18.4.4), preserves the properties (C1–C4) of functions for all stable states of the system.*

The proofs of all these statements are technical and cumbersome; they can be found in Rykov and Efrosinin [405].

Because the function

$$v_0(x) = \sum_{l \in J_1(x)} \frac{c_l}{\mu_l}$$

possesses the properties (C1–C4), then by successive approximation it is possible to prove the following consequence.

Corollary 18.5.6. *An optimal control policy $f(x)$ is a monotone (with respect to the introduced partial order) one and demands:*

- *To switch on some server in the state x only in the case if queue length exceeds the threshold level $q^*(x)$, depending on the system state (structure of busy servers) x.*
- *In this case it is necessary to switch on the server with minimal service cost* $\min\{c_i \mu_i^{-1} : \ i \in J_0(x)\}$.

18.6 Conclusion

The classical SSP has been generalized to the case of mean long-run service cost minimization. The conditions for optimality of threshold policy have been obtained.

The numerical analysis shows that the monotonicity of optimal policy holds without condition (18.5.2)

$$0 \leq \mu_1^{-1} \leq \mu_2^{-1} \leq \cdots \leq \mu_K^{-1}.$$

However, our proof depends on this condition. It would be interesting to establish the optimal policy monotonicity without this condition. The theoretical proof of this fact is still an open problem.

Chapter 19

Dependence Comparison of Multivariate Extremes via Stochastic Tail Orders

Haijun Li

Abstract: A stochastic tail order is introduced to compare right tails of distributions and related closure properties are established. The stochastic tail order is then used to compare the dependence structure of multivariate extreme value distributions in terms of upper tail behaviors of their underlying samples.

19.1 Introduction

Let $\boldsymbol{X}_n = (X_{1,n}, \cdots, X_{d,n})$, $n = 1, 2, \cdots$ be independent and identically distributed (i.i.d.) random vectors with common distribution function (df) F. Define component-wise maxima $M_{i,n} := \vee_{j=1}^{n} X_{i,j}$ and minima $m_{i,n} := \wedge_{j=1}^{n} X_{i,j}$, $1 \le i \le d$. Here and hereafter $\vee$ ($\wedge$) denotes the maximum (minimum). This paper focuses on

H. Li (✉)
Department of Mathematics, Washington State University,
Pullman, WA 99164, USA
e-mail: lih@math.wsu.edu

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9_19,

dependence comparison of the limiting distributions of properly normalized vectors of component-wise maxima $\boldsymbol{M}_n := (M_{1,n}, \ldots, M_{d,n})$ and of component-wise maxima $\boldsymbol{m}_n := (m_{1,n}, \ldots, m_{d,n})$, as $n \to \infty$. The comparison method is based on asymptotic comparisons of upper tails of F of the underlying sample $(\boldsymbol{X}_n, n \geq 1)$.

For any two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$, the sum $\boldsymbol{a}+\boldsymbol{b}$, product $\boldsymbol{ab}$, quotient $\boldsymbol{a}/\boldsymbol{b}$, and vector inequalities such as $\boldsymbol{a} \leq \boldsymbol{b}$ are all operated component-wise. Let G and H be dfs defined on $\mathbb{R}^d$ with nondegenerate margins. A df F is said to be in the domain of attraction of G for the maxima, denoted as $F \in \mathrm{DA}_\vee(G)$, if there exist $\mathbb{R}^d$-valued sequences $\boldsymbol{a}_n = (a_{1,n}, \cdots, a_{d,n})$ with $a_{i,n} > 0$, $1 \leq i \leq d$, and $\boldsymbol{b}_n = (b_{1,n}, \cdots, b_{d,n})$, $n = 1, 2, \cdots$, such that for any $\boldsymbol{x} = (x_1, \cdots, x_d)$, as $n \to \infty$,

$$
\begin{aligned}
& \mathrm{P}\left\{\frac{M_{1,n} - b_{1,n}}{a_{1,n}} \leq x_1, \cdots, \frac{M_{d,n} - b_{d,n}}{a_{d,n}} \leq x_d\right\} \\
= \; & F^n(\boldsymbol{a}_n x + \boldsymbol{b}_n) \to G(\boldsymbol{x}), \qquad (19.1.1)
\end{aligned}
$$

and in this case, G is called a max multivariate extreme value (MEV) distribution. Similar definitions for min MEV distributions and their domain of attraction can be made. For minima, Eq. (19.1.1) is replaced by

$$
\begin{aligned}
& \mathrm{P}\left\{\frac{m_{1,n} - b_{1,n}}{a_{1,n}} > x_1, \cdots, \frac{m_{d,n} - b_{d,n}}{a_{d,n}} > x_d\right\} \\
= \; & \overline{F}^n(\boldsymbol{a}_n x + \boldsymbol{b}_n) \to \overline{H}(\boldsymbol{x}), \qquad (19.1.2)
\end{aligned}
$$

which is denoted by $F \in \mathrm{DA}_\wedge(H)$. Here and hereafter bars on the top of dfs denote (joint) survival functions. A key property of an MEV distribution G is that all positive powers of G are also distributions, and max MEV distributions coincide with the max-stable distributions, which form a subclass of max-infinitely divisible distributions. Similarly min MEV distributions coincide with the min-stable distributions, which form a subclass of min-infinitely divisible distributions. One needs only to study the case of maxima as the theory for minima is similar.

Let $\boldsymbol{X} = (X_1, \ldots, X_d)$ be a generic random vector with distribution F and continuous, univariate margins $F_1, \ldots, F_d$. If $F \in \mathrm{DA}_\vee(G)$, then G is closely related to the upper tail distribution of X, which often possesses the heavy tail property of *regular variation*. Without loss of generality, we may assume that $\boldsymbol{X}$ is nonnegative component-wise.

Consider the standard case in which the survival functions $\overline{F}_i(x) := 1 - F_i(x)$, $1 \le i \le d$ of the margins are right tail equivalent; that is,

$$\frac{\overline{F}_i(x)}{\overline{F}_1(x)} = \frac{1 - F_i(x)}{1 - F_1(x)} \to 1, \text{ as } x \to \infty, \ 1 \le i \le d. \tag{19.1.3}$$

The distribution F or random vector $\boldsymbol{X}$ is said to be *multivariate regularly varying* (MRV) at ∞ with intensity measure ν if there exists a scaling function $b(t) \to \infty$ and a nonzero Radon measure $\nu(\cdot)$ such that as $t \to \infty$,

$$t\,\mathrm{P}\left\{\frac{\boldsymbol{X}}{b(t)} \in B\right\} \to \nu(B), \forall \text{ relatively compact sets } B \subset \overline{\mathbb{R}}_+^d \backslash \{0\}, \quad \text{with } \nu(\partial B) = 0, \tag{19.1.4}$$

where $\overline{\mathbb{R}}_+^d := [0, \infty]^d$. The extremal dependence information of $\boldsymbol{X}$ is encoded in the intensity measure ν that satisfies that $\nu(tB) = t^{-\alpha}\nu(B)$, for all relatively compact subsets B that are bounded away from the origin, where $\alpha > 0$ is known as the *tail index*. Since the set $B_1 = \{\boldsymbol{x} \in \mathbb{R}^d : x_1 > 1\}$ is relatively compact within the cone $\overline{\mathbb{R}}_+^d \backslash \{0\}$ and $\nu(B_1) > 0$ under Eq. (19.1.3), it follows from Eq. (19.1.4) that the scaling function $b(t)$ can be chosen to satisfy that $\overline{F}_1(b(t)) = t^{-1}$, $t > 0$, after appropriately normalizing the intensity measure by $\nu(B_1)$. That is, $b(t)$ can be chosen as $b(t) = \overline{F}^{-1}(t^{-1}) = F_1^{-1}(1 - t^{-1})$ under the condition (19.1.3), and thus, Eq. (19.1.4) can be expressed equivalently as

$$\lim_{t \to \infty} \frac{\mathrm{P}\{\boldsymbol{X} \in tB\}}{\mathrm{P}\{X_1 > t\}} = \nu(B), \ \forall \text{ relatively compact sets } B \subset \overline{\mathbb{R}}_+^d \backslash \{0\}, \tag{19.1.5}$$

satisfying that $\mu(\partial B) = 0$. It follows from Eqs. (19.1.5) and (19.1.3) that for $1 \le i \le d$,

$$\lim_{t \to \infty} \frac{\mathrm{P}\{X_i > ts\}}{\mathrm{P}\{X_i > t\}} = \nu((s, \infty] \times \overline{\mathbb{R}}^{d-1}) = s^{-\alpha}\nu((1, \infty] \times \overline{\mathbb{R}}^{d-1}), \ \forall \ s > 0.$$

That is, univariate margins have regularly varying right tails. In general, a Borel-measurable function $g : \mathbb{R}_+ \to \mathbb{R}_+$ is regularly varying with exponent $\rho \in \mathbb{R}$, denoted as $g \in \mathrm{RV}_\rho$, if and only if

$$g(t) = t^\rho \ell(t), \text{ with } \ell(\cdot) \ge 0 \text{ satisfying that } \lim_{t \to \infty} \frac{\ell(ts)}{\ell(t)} = 1, \text{ for } s > 0. \tag{19.1.6}$$

The function $\ell(\cdot)$ is known as a slowly varying function and denoted as $\ell \in \mathrm{RV}_0$. Since $\overline{F}_1 \in \mathrm{RV}_{-\alpha}$, $1/\overline{F}_1 \in \mathrm{RV}_\alpha$, and thus, by Proposition 2.6 (v) of [387], the scaling function $b \in \mathrm{RV}_{\alpha^{-1}}$.

Since all the margins are tail equivalent as assumed in Eq. (19.1.3), one has

$$\overline{F}_i(t) = t^{-\alpha}\ell_i(t), \text{ where } \ell_i \in \mathrm{RV}_0, \text{ and } \ell_i(t)/\ell_j(t) \to 1 \text{ as } t \to \infty, \text{ for any } i \neq j, \tag{19.1.7}$$

which, together with $\overline{F}_1(b(t)) = t^{-1}$, imply that

$$\lim_{t\to\infty} t\,\mathrm{P}\{X_i > b(t)s\} = \lim_{t\to\infty} \frac{\mathrm{P}\{X_i > b(t)s\}}{\overline{F}_i(b(t))}\frac{\overline{F}_i(b(t))}{\overline{F}_1(b(t))} = s^{-\alpha},\ s > 0,\ 1 \le i \le d. \tag{19.1.8}$$

The detailed discussions on univariate and multivariate regular variations can be found in [61, 387]. The extension of MRV beyond the nonnegative orthant can be done by using the tail probability of $||X||$, where $||\cdot||$ denotes a norm on $\mathbb{R}^d$, in place of the marginal tail probability in Eq. (19.1.5) (see [387], Sect. 6.5.5). The case that the limit in Eq. (19.1.3) is any nonzero constant can be easily converted into the standard tail equivalent case by properly rescaling margins. If the limit in Eq. (19.1.3) is zero or infinity, then some margins have heavier tails than others. One way to overcome this problem is to standardize the margins via marginal monotone transforms (see Theorem 6.5 in [387]) or to use the copula method [281].

Theorem 19.1.1 (Marshall and Olkin [309]). *Assume that Eq. (19.1.3) holds. Then there exist normalization vectors $\boldsymbol{a_n} > \boldsymbol{0}$ and $\boldsymbol{b}_n$ such that, as $n \to \infty$,*

$$\mathrm{P}\Big\{\frac{\boldsymbol{M_n} - \boldsymbol{b_n}}{\boldsymbol{a}_n} \le \boldsymbol{x}\Big\} \to G(\boldsymbol{x}), \qquad \forall\ \boldsymbol{x} \in \mathbb{R}^d_+,$$

where G is a d-dimensional distribution with Fréchet margins $G_i(s) = \exp\{-s^{-\alpha}\}$, $1 \le i \le d$, if and only if F is MRV with intensity measure $\nu([0,\boldsymbol{x}]^c) := -\log G(\boldsymbol{x})$.

In other words, $F \in \mathrm{DA}_\vee(G)$ where G has Fréchet margins with tail index α if and only if F is MRV with intensity measure $\nu([0,\boldsymbol{x}]^c) = -\log G(\boldsymbol{x})$.

Remark 19.1.2.

1. The normalization vectors $\boldsymbol{a}_n > 0$ and $\boldsymbol{b}_n$ in Theorem 19.1.1 can be made precise so that $\boldsymbol{b}_n$=0 and $\boldsymbol{a}_n$=$(\overline{F}_1^{-1}(1/n), \dots, \overline{F}_d^{-1}(1/n))$ that depend only on the margins of F.

2. If Eq. (19.1.3) does not hold, Theorem 19.1.1 can still be established, but the nonstandard global regular variation with different scaling functions among various margins needs to be used in place of Eq. (19.1.5), which uses the same scaling function among different margins.

3. One-dimensional version of Theorem 19.1.1 is due to Gnedenko [184]. Note that the parametric feature enjoyed by univariate extremes is lost in the multivariate context.

4. Let $\mathbb{S}_+^{d-1} = \{\boldsymbol{a} : \boldsymbol{a} = (a_1, \dots, a_d) \in \mathbb{R}_+^d, ||\boldsymbol{a}|| = 1\}$, where $||\cdot||$ is a norm defined in $\mathbb{R}^d$. Using the polar coordinates, G can be expressed as follows:

$$G(\boldsymbol{x}) = \exp\Big\{ - c \int_{\mathbb{S}_+^d} \max_{1\le i\le d} \{(a_i/x_i)^\alpha\} \mathbb{Q}(\mathrm{d}\boldsymbol{a}) \Big\},$$

   where $c > 0$ and $\mathbb{Q}$ is a probability measure defined on $\mathbb{S}_+^{d-1}$. This is known as the Pickands representation [374], and $c\mathbb{Q}(\cdot)$ is known as the spectral or angular measure.

5. Note that the spectral measure is a finite measure that can be approximated by a sequence of discrete measures. Using this idea, Marshall and Olkin [309] showed that the MEV distribution G is positively associated. This implies that as n is sufficiently large, we have asymptotically,

$$\mathrm{E}\Big[f(\boldsymbol{M}_n)g(\boldsymbol{M}_n)\Big] \ge \mathrm{E}\Big[f(\boldsymbol{M}_n)\Big]\mathrm{E}\Big[g(\boldsymbol{M}_n)\Big]$$

   for all nondecreasing functions $f, g : \mathbb{R}^d \mapsto \mathbb{R}$. Observe that the sample vector $\boldsymbol{X}_n$ could have any dependence structure, but the strong positive dependence emerges among multivariate extremes.

6. Since G is max-infinitely divisible, all bivariate margins of G are TP_2, a positive dependence property that is even stronger than the positive association of bivariate margins (see Theorem 2.6 in [211]).

Since the normalization vectors $\boldsymbol{a}_n > 0$ and $\boldsymbol{b}_n$ in Theorem 19.1.1 depend only on the margins, dependence comparison of G can be easily established using the *orthant dependence order* on sample vectors.

Recall that a d-dimensional random vector $\boldsymbol{X} = (X_1, \dots, X_d)$ with df F is said to be smaller than another d-dimensional random vector $\boldsymbol{X}' = (X_1', \dots, X_d')$ with df F' in the upper (lower) orthant order, denoted as $\boldsymbol{X} \le_{\text{uo}} \boldsymbol{X}'$ or $F \le_{\text{uo}} F'$ ($\boldsymbol{X} \le_{\text{lo}} \boldsymbol{X}'$ or $F \le_{\text{lo}} F'$), if

$$\mathrm{P}\{X_1 > x_1, \dots, X_d > x_d\} \le \mathrm{P}\{X_1' > x_1, \dots, X_d' > x_d\}, \qquad (19.1.9)$$

$$\mathrm{P}\{X_1 \le x_1, \dots, X_d \le x_d\} \le \mathrm{P}\{X_1' \le x_1, \dots, X_d' \le x_d\}, \qquad (19.1.10)$$

for all $(x_1, \dots, x_d) \in \mathbb{R}^d$. If, in addition, their corresponding univariate margins are identical, then $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{X}'$ in the upper (lower) orthant dependence order, denoted as $\boldsymbol{X} \le_{\text{uod}} \boldsymbol{X}'$ or $F \le_{\text{uod}} F'$ ($\boldsymbol{X} \le_{\text{lod}} \boldsymbol{X}'$ or $F \le_{\text{lod}} F'$). Clearly $\boldsymbol{X} \le_{\text{uod}} \boldsymbol{X}'$ implies that $\boldsymbol{X} \le_{\text{uo}} \boldsymbol{X}'$, but the order $\le_{\text{uod}}$ focuses on comparing scale-invariant dependence among components. The detailed discussions on these orders can be found in [335, 426]. The following result is immediate due to the fact that the orthant order is closed under weak convergence.

Proposition 19.1.3. *Let* $(\boldsymbol{X}_n, n \ge 1)$ *and* $(\boldsymbol{X}_n', n \ge 1)$ *be two i.i.d. samples with dfs* F *and* F'*, respectively. If* $F \in DA_\vee(G)$ *and* $F' \in DA_\vee(G')$ *with Fréchet margins, then* $\boldsymbol{X}_n \le_{\text{lod}} \boldsymbol{X}_n'$ *implies that* $G \le_{\text{lod}} G'$*.*

Note, however, that the ordering $\boldsymbol{X}_n \le_{\text{lod}} \boldsymbol{X}_n'$ is strongly affected by the behavior at the center and often too strong to be valid. The fact that MRV is a tail property motivates us to focus on comparing only upper tails of $\boldsymbol{X}_n$ and $\boldsymbol{X}_n'$, leading to weaker notions of *stochastic tail orders.* In Sect. 19.2, we introduce a notion of stochastic tail order for random variables and establish related closure properties and discuss its relation with other asymptotic orders that are already available in the literature. In Sect. 19.3, we extend the stochastic tail order to random vectors and show that the stochastic tail order of sample vectors sufficiently implies the orthant dependence order of the corresponding MEV distributions.

19.2 Stochastic Tail Orders

Let X and Y be two $\mathbb{R}_+$-valued random variables. X is said to be smaller than Y in the sense of stochastic tail order, denoted as $X \le_{\text{sto}} Y$, if there exists a threshold constant $t_0 > 0$ (usually large) such that

$$\mathrm{P}\{X > t\} \le \mathrm{P}\{Y > t\}, \qquad \forall\, t > t_0. \qquad (19.2.1)$$

Remark 19.2.1.

1. The stochastic tail order $\leq_{\text{sto}}$ is reflexive and transitive. $\leq_{\text{sto}}$ is antisymmetric if tail identically distributed random variables are considered to be equivalent.

2. If X is smaller than Y in the usual stochastic order (denoted as $X \leq_{\text{st}} Y$; see Sect. 1.A in [426]); that is, $\mathrm{P}\{X > t\} \leq \mathrm{P}\{Y > t\}$ for all t, then $X \leq_{\text{sto}} Y$.

3. $X \leq_{\text{sto}} Y$ if and only if there exists a small open neighborhood of ∞ within which X is stochastically smaller than Y.

4. $X \leq_{\text{sto}} Y$ implies
$$\limsup_{t\to\infty} \frac{\mathrm{P}\{X > t\}}{\mathrm{P}\{Y > t\}} \leq 1. \tag{19.2.2}$$

The stochastic tail orders using limiting inequalities such as Eq. (19.2.2) have been introduced and studied in [43, 242, 243, 389–391] and more recently in [300]. Most of these tail orders, however, are based on limiting approaches rather than stochastic comparison theory.

1. Mainik and Rüschendorf studied in [300] the following weak tail order: A random variable X is said to be smaller than another random variable Y in the asymptotic portfolio loss order, denoted as $X \leq_{\text{apl}} Y$, if the limiting inequality (19.2.2) holds. Observe that $\sup\limits_{s>t} \frac{\mathrm{P}\{X>s\}}{\mathrm{P}\{Y>s\}}$ is decreasing in t, and as such, in the case of $X \leq_{\text{apl}} Y$ with
$$\limsup_{t\to\infty} \frac{\mathrm{P}\{X > t\}}{\mathrm{P}\{Y > t\}} = \lim_{t\to\infty} \left[\sup_{s>t} \frac{\mathrm{P}\{X > s\}}{\mathrm{P}\{Y > s\}}\right] = 1,$$
one can find in any open neighborhood $(c, \infty]$ of ∞ that $\mathrm{P}(X > s) \geq \mathrm{P}(Y > s)$ for some $s > c$. That is, neither X nor Y could dominate the other in any open neighborhood $(c, \infty]$ of ∞, but asymptotically, the right tail of X decays at the rate that is bounded from above by the tail decay rate of Y.

2. Rojo introduced in [391] a stronger version of tail orders: Define $X <_{\text{sq}} Y$ if

$$\limsup_{u \to 1} \frac{F^{-1}(u)}{G^{-1}(u)} < 1$$

where $F^{-1}(\cdot)$ and $G^{-1}(\cdot)$ denote the left-continuous inverses of dfs of X and Y, respectively. Obviously, $X <_{\text{sq}} Y$ implies that $X \leq_{\text{sto}} Y$. Note, however, that $<_{\text{sq}}$ is not a partial ordering.

The stochastic tail orders via limiting inequalities resemble the idea of comparing the asymptotic decay rates that is often employed in theory of large (and small) deviations [284]. In contrast, the notion Eq. (19.2.1) compares stochastically random variables in a small open neighborhood of ∞ within which theory of stochastic orders retains its full power. For example, coupling remains valid in a small open neighborhood of ∞.

Theorem 19.2.2. *Let X and Y be two positive random variables with support $[0, \infty)$. $X \leq_{\text{sto}} Y$ if and only if there exists a random variable Z defined on the probability space $(\Omega, \mathcal{F}, \mathrm{P})$ with support $[a, b]$, and nondecreasing functions ψ_1 and ψ_2 with $\lim_{z \to b} \psi_i(z) = \infty$, $i = 1, 2$, such that $X \stackrel{d}{=} \psi_1(Z)$, $Y \stackrel{d}{=} \psi_2(Z)$ and $\mathrm{P}\{\psi_1(Z) \leq \psi_2(Z) \mid Z \geq z_0\} = 1$ for some $z_0 > 0$.*

Proof: Let X and Y have distributions F and G with support $[0, \infty)$, respectively, and let $F^{-1}(\cdot)$ and $G^{-1}(\cdot)$ denote the corresponding left-continuous inverses. Recall that for any df H on $\mathbb{R}$, the left-continuous inverse of H is defined as

$$H^{-1}(u) := \inf\{s : H(s) \geq u\}, \ 0 \leq u \leq 1.$$

The left-continuous inverse has the following desirable properties:

1. $H(H^{-1}(u)) \geq u$ for all $0 \leq u \leq 1$, and $H^{-1}(H(x)) \leq x$ for all $x \in \mathbb{R}$.

2. $H^{-1}(u) \leq x$ if and only if $u \leq H(x)$.

3. The set $\{s : H(s) \geq u\}$ is closed for each $0 \leq u \leq 1$.

Necessity: Using Properties 1 and 2, $X \leq_{\text{sto}} Y$ implies that $F^{-1}(u) \leq G^{-1}(u)$, $\forall u > u_0$ for some $0 < u_0 < 1$. Let U be a random variable with standard uniform distribution, and thus $\mathrm{P}\{F^{-1}(U) \leq G^{-1}(U) \mid U \geq u_0\} = 1$. Using Property 2, $\mathrm{P}\{F^{-1}(U) \leq x\} = \mathrm{P}\{U \leq F(x)\} = F(x)$. Similarly, $\mathrm{P}\{G^{-1}(U) \leq x\} = G(x)$.

Sufficiency: For all $t \geq c$, a constant with $c > \psi_1(z_0)$,

$$\begin{aligned} \mathrm{P}\{X > t\} &= \mathrm{P}\{Z \geq z_0\}\mathrm{P}\{\psi_1(Z) > t \mid Z \geq z_0\} \\ &\leq \mathrm{P}\{Z \geq z_0\}\mathrm{P}\{\psi_2(Z) > t \mid Z \geq z_0\} \\ &\leq \mathrm{P}\{\psi_2(Z) > t\} \\ &= \mathrm{P}\{Y > t\}. \end{aligned}$$

■

The tail coupling presented in Theorem 19.2.2 enables us to establish desirable closure properties for the stochastic tail order. A Borel measurable function $\psi : \mathbb{R}^d \to \mathbb{R}$ is called a Radon function if ψ is bounded on every compact subset of $\mathbb{R}^d$. Obviously, any nondecreasing function and any continuous function defined on $\mathbb{R}^d$ are Radon functions.

Definition 19.2.3. A Borel measurable function $\psi : \mathbb{R}^d_+ \to \mathbb{R}$ is said to be eventually increasing if there exists a compact subset $S \subset \mathbb{R}^d_+$ such that ψ is component-wise nondecreasing on S^c with $\lim_{x_i \to \infty} \psi(x_1, \ldots, x_i, \ldots, x_d) = \infty$.

Proposition 19.2.4. *Let X and Y be two positive random variables with support $[0, \infty)$.*

1. *$X \leq_{\text{sto}} Y$ implies $g(X) \leq_{\text{sto}} g(Y)$ for any Radon function g that is eventually increasing.*

2. *If X_1, X_2 are independent, and X_1', X_2' are independent, then $X_1 \leq_{\text{sto}} X_1'$ and $X_2 \leq_{\text{sto}} X_2'$ imply $g(X_1, X_2) \leq_{\text{sto}} g(X_1', X_2')$ for any Radon function g that is eventually increasing.*

Proof:

(1) Since g is a Radon function that is eventually increasing, there exists a threshold $x_0 > 0$ such that $g(\cdot)$ is increasing to ∞ on $[x_0, \infty)$. By Theorem 19.2.2, there exists a random variable Z defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and nondecreasing functions ψ_1 and ψ_2 with $\lim_{z \to b} \psi_i(z) = \infty$, $i = 1, 2$, such that $X \stackrel{d}{=} \psi_1(Z)$ and $Y \stackrel{d}{=} \psi_2(Z)$ and

$$\mathrm{P}\{\psi_1(Z) \leq \psi_2(Z) \mid Z \geq z_0\}{=}1 \quad \text{for some } z_0{>}0 \text{ with } \psi_1(z_0){>}x_0.$$

Thus,

$$P\{g(\psi_1(Z)) \le g(\psi_2(Z)) \mid Z \ge z_0\} = 1 \quad \text{for } z_0 > 0.$$

Clearly, $g(X) \stackrel{d}{=} g(\psi_1(Z))$ and $g(Y) \stackrel{d}{=} g(\psi_2(Z))$, and thus,

$$\begin{aligned} P\{g(X) > t\} &= P\{Z \ge z_0\}P\{g(\psi_1(Z)) > t \mid Z \ge z_0\} \\ &\le P\{Z \ge z_0\}P\{g(\psi_2(Z)) > t \mid Z \ge z_0\} \\ &\le P\{g(\psi_2(Z)) > t\} = P\{g(Y) > t\} \end{aligned}$$

for any $t \ge c$ where c is a constant with $c > g(\psi_1(z_0))$.

(2) Without loss of generality, assume that (X_1, X_2) and (X_1', X_2') are independent. We only need to show that $g(X_1, X_2) \le_{\text{sto}} g(X_1', X_2)$. Since $g(\cdot)$ is a Radon function that is eventually increasing, there exists a (x_1, x_2) such that $g(\cdot)$ is bounded on $[0, x_1] \times [0, x_2]$ and increasing on $([0, x_1] \times [0, x_2])^c$.

1. By Theorem 19.2.2, there exists a random variable Z_1 defined on the probability space $(\Omega_1, \mathcal{F}_1, P_1)$ and nondecreasing functions ψ_1 and ψ_1' such that $X_1 \stackrel{d}{=} \psi_1(Z_1)$ and $X_1' \stackrel{d}{=} \psi_1'(Z_1)$ and $P_1\{\psi_1(Z_1) \le \psi_1'(Z_1) \mid Z_1 \ge z_1\} = 1$ for some $z_1 > 0$ with $\psi_1(z_1) > x_1$.
2. Let $(\Omega_2, \mathcal{F}_2, P_2)$ denote the underlying probability space of X_2.

Construct a product probability space $(\Omega, \mathcal{F}, P) = (\Omega_1 \times \Omega_2, \mathcal{F}, P_1 \times P_2)$, where $\mathcal{F}$ is the σ-field generated by $\mathcal{F}_1 \times \mathcal{F}_2$. On this enlarged product probability space, since $P\{\psi_1(Z_1) \le \psi_1'(Z_1) \mid Z_1 \ge z_1\} = 1$, and $g(\cdot)$ is increasing on $([0, x_1] \times [0, x_2])^c$, we have that $P\{g(\psi_1(Z_1), X_2) \le g(\psi_1'(Z_1), X_2) \mid Z_1 \ge z_1 \text{ or } X_2 > x_2\} = 1$. Clearly, $g(X_1, X_2) \stackrel{d}{=} g(\psi_1(Z_1), X_2)$ and $g(X_1', X_2) \stackrel{d}{=} g(\psi_1'(Z_1), X_2)$, and thus

$$\begin{aligned} &P\{g(X_1, X_2) > t\} \\ = \; &P\{Z_1 \ge z_1 \text{ or } X_2 > x_2\}P\{g(\psi_1(Z_1), X_2) > t \mid Z_1 \ge z_1 \text{ or } X_2 > x_2\} \\ \le \; &P\{Z_1 \ge z_1 \text{ or } X_2 > x_2\}P\{g(\psi_1'(Z_1), X_2) > t \mid Z_1 \ge z_1 \text{ or } X_2 > x_2\} \\ \le \; &P\{g(\psi_1'(Z_1), X_2) > t\} \\ = \; &P\{g(X_1', X_2) > t\} \end{aligned}$$

for any $t \ge c$ where c is a constant with $c > g(\psi_1(z_1), x_2)$. That is, $g(X_1, X_2) \le_{\text{sto}} g(X_1', X_2)$. Similarly, $g(X_1', X_2) \le_{\text{sto}} g(X_1', X_2')$. ■

Corollary 19.2.5. *If X_1, X_2 are independent and Y_1, Y_2 are independent, then $X_1 \leq_{\text{sto}} Y_1$ and $X_2 \leq_{\text{sto}} Y_2$ imply*

$$X_1 X_2 \leq_{\text{sto}} Y_1 Y_2, \qquad X_1 + X_2 \leq_{\text{sto}} Y_1 + Y_2,$$
$$X_1 \vee X_2 \leq_{\text{sto}} Y_1 \vee Y_2, \qquad X_1 \wedge X_2 \leq_{\text{sto}} Y_1 \wedge Y_2.$$

In particular, $R_1 \leq_{\text{sto}} R_2$ implies that $R_1 V \leq_{\text{sto}} R_2 V$ for any non-negative random variable V that is independent of R_1, R_2. Mainik and Rüschendorf obtained this inequality in [300] for the random variable V that is bounded using the ordering $\leq_{\text{apl}}$, and their proof is based on the method of mixture.

Proposition 19.2.6. *Let X and Y be two positive random variables with support $[0, \infty)$ and Θ be a random variable with bounded support $[\theta_L, \theta_U]$. Assume that:*

1. *Θ is a random variable with finite masses.*
2. *Θ is a continuous random variable such that $\mathrm{P}\{X > t \mid \Theta = \theta\}$ and $\mathrm{P}\{Y > t \mid \Theta = \theta\}$ are continuous in θ.*

If $[X \mid \Theta = \theta] \leq_{\text{sto}} [Y \mid \Theta = \theta]$ for all θ in the support of Θ, then $X \leq_{\text{sto}} Y$.

Proof: Since $[X \mid \Theta = \theta] \leq_{\text{sto}} [Y \mid \Theta = \theta]$ for all $\theta \in [\theta_L, \theta_U]$, there exists a threshold t_θ that is given by

$$t_\theta := \sup\big\{s : \mathrm{P}\{X > s \mid \Theta = \theta\} > \mathrm{P}\{Y > s \mid \Theta = \theta\}\big\}, \tag{19.2.3}$$

such that

$$\mathrm{P}\{X > t \mid \Theta = \theta\} \leq \mathrm{P}\{Y > t \mid \Theta = \theta\}, \quad \forall\, t > t_\theta.$$

Notice that the threshold t_θ depends on the mixing value θ. Consider the following two cases. Construct the threshold $t_{[\theta_L, \theta_U]}$ as follows:

1. If Θ is discrete with finite masses, then define
$$t_{[\theta_L,\theta_U]} := \max\{t_\theta : \theta \in [\theta_L, \theta_U]\} < \infty.$$
2. If Θ is continuous, then t_θ is continuous in θ due to the assumption that $\mathrm{P}\{X > t \mid \Theta = \theta\}$ and $\mathrm{P}\{Y > t \mid \Theta = \theta\}$ are continuous in θ. Define
$$t_{[\theta_L,\theta_U]} := \sup\{t_\theta : \theta \in [\theta_L, \theta_U]\},$$
which is finite because of the continuity of t_θ and the compactness of $[\theta_0, \theta_n]$.

In any case, for any $\theta \in [\theta_L, \theta_U]$, any $t > t_{[\theta_L,\theta_U]}$,

$$\mathrm{P}\{X > t \mid \Theta = \theta\} \le \mathrm{P}\{Y > t \mid \Theta = \theta\}.$$

Taking the integrations on both sides from θ_L to θ_U, we obtain $\mathrm{P}\{X > t\} \le \mathrm{P}\{Y > t\}$ for any $t > t_{[\theta_L,\theta_U]}$. ∎

Remark 19.2.7. The closure property under mixture when the mixing variable has unbounded support, say $[0, \infty)$, becomes more subtle. This is because the threshold t_θ defined in Eq. (19.2.3) can approach to infinity as θ goes to infinity. Our conjecture is that in the case of unbounded support, $[X \mid \Theta = \theta] \le_{\mathrm{sto}} [Y \mid \Theta = \theta]$ for all θ in the support of Θ implies that $\limsup_{t\to\infty} \frac{\mathrm{P}\{X>t\}}{\mathrm{P}\{Y>t\}} \le 1$; that is, $X \le_{\mathrm{apl}} Y$.

In the examples to be discussed below, all involved random variables fail to satisfy the usual stochastic order.

Example 19.2.8.

1. Let X have the Weibull distribution with unit scale parameter and shape parameter k and Y have the exponential distribution with unit (scale) parameter. If the shape parameter $k > 1$ (i.e., increasing hazard rate), then $X \le_{\mathrm{sto}} Y$. If the shape parameter $k < 1$ (i.e., decreasing hazard rate), then $X \ge_{\mathrm{sto}} Y$. Note that both X and Y have exponentially decayed right tails.

2. Let X have the exponential distribution with unit (scale) parameter and Y have the distribution of Pareto Type II with tail index $\alpha = 2$; that is,

$$\mathrm{P}\{Y > t\} = (1 + t)^{-2},\ t \ge 0. \tag{19.2.4}$$

   Then $X \le_{\mathrm{sto}} Y$. Note that Y has regularly varying right tail as described in Eq. (19.1.6).

3. If X has the Fréchet distribution with tail index $\alpha = 3$ (see Theorem 19.1.1) and Y has the distribution Eq. (19.2.4) of Pareto Type II with tail index $\alpha = 2$, then $X \le_{\mathrm{sto}} Y$. Note that X and Y are regularly varying with respective tail indexes 3 and 2, but Y has a heavier tail than that of X.

4. Let X have the survival function of Pareto Type I as defined as follows:
$$\mathrm{P}\{X > t\} = \left(\frac{t}{0.5}\right)^{-1}, \; t \geq 0.5.$$
Let Y have the survival function of Pareto Type II with tail index 1; that is,
$$\mathrm{P}\{Y > t\} = (1+t)^{-1}, \; t \geq 0.$$
Then $X \leq_{\text{sto}} Y$. Note that both X and Y have regularly varying right tails with same tail index 1.

5. Let R_1 and R_2 have regularly varying distributions with tail indexes α_1 and α_2, respectively. If $\alpha_1 > \alpha_2$, then $R_1 \leq_{\text{sto}} R_2$. That is, the random variable with heavier regularly varying right tail is larger stochastically in the tail.

6. Let R be regularly varying with tail index α. If V_1 and V_2 are random variables with finite moments of any order, independent of R, such that $\mathrm{E}[V_1^\alpha] < \mathrm{E}[V_2^\alpha]$. By Breiman's theorem (see [387], p. 232),
$$\lim_{t\to\infty} \frac{\mathrm{P}\{RV_1 > t\}}{\mathrm{P}\{R > t\}} = \mathrm{E}[V_1^\alpha] < \mathrm{E}[V_2^\alpha] = \lim_{t\to\infty} \frac{\mathrm{P}\{RV_2 > t\}}{\mathrm{P}\{R > t\}}$$
Thus, for $t > t_0$ where t_0 is sufficiently large, $\mathrm{P}\{RV_1 > t\} < \mathrm{P}\{RV_2 > t\}$, implying that $RV_1 \leq_{\text{sto}} RV_2$. ■

A multivariate extension of scale mixtures discussed in Example 19.2.8 (6) includes the multivariate elliptical distribution. A random vector $\boldsymbol{X} \in \mathbb{R}^d$ is called elliptically distributed if $\boldsymbol{X}$ has the representation:
$$\boldsymbol{X} \stackrel{d}{=} \boldsymbol{\mu} + RA\boldsymbol{U} \tag{19.2.5}$$
where $\boldsymbol{\mu} \in \mathbb{R}^d$, $A \in \mathbb{R}^{d\times d}$ and $\boldsymbol{U}$ is uniformly distributed on $\mathbb{S}_2^{d-1} := \{\boldsymbol{x} \in \mathbb{R}^d : ||\boldsymbol{x}||_2 = 1\}$ and $R \geq 0$ is independent of $\boldsymbol{U}$. We denote this by $\boldsymbol{X} \sim \mathcal{E}(\boldsymbol{\mu}, \Sigma, R)$ where $\Sigma = AA^\top$.

Proposition 19.2.9. *Let* $\boldsymbol{X}\sim\mathcal{E}(\boldsymbol{\mu}_1, \Sigma_1, R_1)$ *and* $\boldsymbol{Y}\sim\mathcal{E}(\boldsymbol{\mu}_2, \Sigma_2, R_2)$*. If* $\boldsymbol{\mu}_1\leq\boldsymbol{\mu}_2$, $R_1\leq_{\text{sto}}R_2$, *and* $\boldsymbol{\xi}^\top\Sigma_1\boldsymbol{\xi} \leq \boldsymbol{\xi}^\top\Sigma_2\boldsymbol{\xi}$, *for fixed* $\boldsymbol{\xi}\in\mathbb{S}_1^{d-1}:=\{\boldsymbol{x}\in\mathbb{R}^d : ||\boldsymbol{x}||_1=1\}$, *then* $|\boldsymbol{\xi}^\top\boldsymbol{X}| \leq_{\text{sto}} |\boldsymbol{\xi}^\top\boldsymbol{Y}|$.

Proof: Let $a_i := \boldsymbol{\xi}^\top \Sigma_i \boldsymbol{\xi}$, $i = 1, 2$. Without loss of generality, we can assume that $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \mathbf{0}$ and $a_1 > 0$. Let

$$\boldsymbol{v}_i := \frac{A^\top \boldsymbol{\xi}}{\boldsymbol{\xi}^\top \Sigma_i \boldsymbol{\xi}} = \frac{A^\top \boldsymbol{\xi}}{a_i},\ i = 1, 2.$$

Clearly $\boldsymbol{v_i}^\top \boldsymbol{v_i} = 1$, $i = 1, 2$, and thus, by symmetry, $\boldsymbol{v_1}^\top \boldsymbol{U}$ and $\boldsymbol{v_2}^\top \boldsymbol{U}$ have the same distribution. Let $\Theta := \boldsymbol{v_1}^\top \boldsymbol{U}$, and we have

$$|\boldsymbol{\xi}^\top \boldsymbol{X}| = R_1 a_1 |\boldsymbol{v_1}^\top \boldsymbol{U}| \stackrel{d}{=} R_1 a_1 |\Theta|,\ |\boldsymbol{\xi}^\top \boldsymbol{Y}| = R_2 a_2 |\boldsymbol{v_2}^\top \boldsymbol{U}| \stackrel{d}{=} R_2 a_2 |\Theta|.$$

The inequality then follows from Corollary 19.2.5 immediately. ■

This is our $\le_{\text{sto}}$-version of a similar result that is obtained in [300] using the $\le_{\text{apl}}$ order.

Remark 19.2.10. Anderson in [12], Fefferman, Jodeit, and Perlman in [159] show that if $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$, $R_1 \stackrel{d}{=} R_2$, and

$$\boldsymbol{\xi}^\top \Sigma_1 \boldsymbol{\xi} \le \boldsymbol{\xi}^\top \Sigma_2 \boldsymbol{\xi},\ \forall\ \boldsymbol{\xi} \in \mathbb{R}^d,$$

then $\mathrm{E}(\psi(\boldsymbol{X})) \le \mathrm{E}(\psi(\boldsymbol{Y}))$ for all symmetric and convex functions $\psi : \mathbb{R}^d \mapsto \mathbb{R}$, such that the expectations exist. This is known as *dilatation*, which can be defined on any locally convex topological linear space $\mathbb{V}$ (traced back to Karamata's work in 1932; see [308], pp. 16–17). Dilatation provides various versions of continuous majorization [308].

19.3 Tail Orthant Orders

Let $\boldsymbol{X} = (X_1, \dots, X_d)$ and $\boldsymbol{X'} = (X'_1, \dots, X'_d)$ be nonnegative random vectors with dfs F and F', respectively. Observe that $\boldsymbol{X} \le_{\text{lo}} \boldsymbol{X'}$ is equivalent to that $\max_{1\le i\le d}\{X_i/w_i\} \ge_{\text{st}} \max_{1\le i\le d}\{X'_i/w_i\}$, and $\boldsymbol{X} \le_{\text{uo}} \boldsymbol{X'}$ is equivalent to that $\min_{1\le i\le d}\{X_i/w_i\} \le_{\text{st}} \min_{1\le i\le d}\{X'_i/w_i\}$. In comparing orthant tails of these random vectors, we focus on the cones $\mathbb{R}^d_+$ and $\overline{\mathbb{R}}^d_+ := \{(x_1, \dots, x_d) : x_i > 0, 1 \le i \le d\}$. Note that $\boldsymbol{x} \in \overline{\mathbb{R}}^d_+$ can have some components taking $+\infty$.

Definition 19.3.1.

1. $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{X'}$ in the sense of tail lower orthant order, denoted as $\boldsymbol{X} \le_{\text{tlo}} \boldsymbol{X'}$, if for all $\boldsymbol{w} = (w_1, \dots, w_d) \in \overline{\mathbb{R}}^d_+$, $\max_{1\le i\le d}\{X_i/w_i\} \ge_{\text{sto}} \max_{1\le i\le d}\{X'_i/w_i\}$.

2. $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{X}'$ in the sense of tail upper orthant order, denoted as $\boldsymbol{X} \le_{\rm tuo} \boldsymbol{X}'$, if for all $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d$, $\min\limits_{1 \le i \le d}\{X_i/w_i\} \le_{\rm sto} \min\limits_{1 \le i \le d}\{X_i'/w_i\}$.

It follows from Eq. (19.2.1) that $\boldsymbol{X} \le_{\rm tlo} \boldsymbol{X}'$ is equivalent to for $\boldsymbol{w} = (w_1, \dots, w_d) \in \overline{\mathbb{R}}_+^d$,

$$\mathrm{P}\{X_1 \le tw_1, \dots, X_d \le tw_d\} \le \mathrm{P}\{X_1' \le tw_1, \dots, X_d' \le tw_d\} \quad (19.3.1)$$

for all $t > t_{\boldsymbol{w}}$ for some $t_{\boldsymbol{w}} > 0$ that may depend on $\boldsymbol{w}$. Similarly, $\boldsymbol{X} \le_{\rm tuo} \boldsymbol{X}'$ is equivalent to for $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{R}_+^d$,

$$\mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \le \mathrm{P}\{X_1' > tw_1, \dots, X_d' > tw_d\} \quad (19.3.2)$$

for all $t > t_{\boldsymbol{w}}$ for some $t_{\boldsymbol{w}} > 0$ that may depend on $\boldsymbol{w}$.

In comparing tail dependence, however, we assume that all the margins of F and F' are tail equivalent. Since we need to compare upper interior orthant tails given fixed marginal tails, consider the two smaller cones:

1. $\mathbb{C}_l := \overline{\mathbb{R}}_+^d \backslash \cup_{j=1}^d \{t\boldsymbol{e}_j^{-1}, t \ge 0\}$, where $\boldsymbol{e}_j^{-1}$, $1 \le j \le d$, denotes the vector with the j-th component being 1 and infinity otherwise.
2. $\mathbb{C}_u := \mathbb{R}_+^d \backslash \cup_{j=1}^d \{t\boldsymbol{e}_j, t \ge 0\}$, where $\boldsymbol{e}_j$, $1 \le j \le d$, denotes the vector with the j-th component being 1 and zero otherwise.

That is, $\mathbb{C}_l$ and $\mathbb{C}_u$ are the subsets of $\overline{\mathbb{R}}_+^d$ after eliminating all the axes that correspond to the margins of a distribution. Note that $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{C}_l$ if and only if at least two components of w are finite, and $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{C}_u$ if and only if at least two components of $\boldsymbol{w}$ are positive.

Definition 19.3.2.

1. $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{X}'$ in the sense of tail lower orthant dependence order, denoted as $\boldsymbol{X} \le_{\rm tlod} \boldsymbol{X}'$, if all the margins of F and F' are tail equivalent, and for all $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{C}_l$,
$$\max_{1 \le i \le d}\{X_i/w_i\} \ge_{\rm sto} \max_{1 \le i \le d}\{X_i'/w_i\}.$$
2. $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{X}'$ in the sense of tail upper orthant dependence order, denoted as $\boldsymbol{X} \le_{\rm tuod} \boldsymbol{X}'$, if all the margins of F and F' are tail equivalent, and for all $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{C}_u$,
$$\min_{1 \le i \le d}\{X_i/w_i\} \le_{\rm sto} \min_{1 \le i \le d}\{X_i'/w_i\}.$$

It follows from Eq. (19.2.1) that if all the margins of F and F' are tail equivalent, then $\boldsymbol{X} \le_{\text{tlod}} \boldsymbol{X'}$ is equivalent to that Eq. (19.3.1) holds for $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{C}_l$, and $\boldsymbol{X} \le_{\text{tuod}} \boldsymbol{X'}$ is equivalent to that Eq. (19.3.2) holds for $\boldsymbol{w} = (w_1, \dots, w_d) \in \mathbb{C}_u$.

Remark 19.3.3.

1. If

$$(w_1, \dots, w_d) \in \bigcup_{j=1}^{d} \{t\boldsymbol{e}_j^{-1}, t \ge 0\} \quad \text{or} \quad (w_1, \dots, w_d) \in \bigcup_{j=1}^{d} \{t\boldsymbol{e}_j, t \ge 0\},$$

   then the inequalities in Definition 19.3.2 reduce to the stochastic tail orders of the marginal distributions. Since the margins are assumed to be tail equivalent, which may not satisfy stochastic tail comparison, we need to eliminate the margins from consideration. On the other hand, with given fixed marginal tails, what really matters in dependence comparison is various interior orthant subsets of $\mathbb{C}_l$ or $\mathbb{C}_u$.

2. If some corresponding margins of F and F' are not tail equivalent, one can still define the tail orthant orders $\le_{\text{tlo}}$ and $\le_{\text{tuo}}$ to compare their tail behaviors in orthants. But all corresponding margins of F and F' have to be tail equivalent in order to compare their tail dependence.

3. If some margins of F (or F') are not tail equivalent, then one can still define the tail orthant dependence order, but scaling functions would be different among the components.

4. Another alternative is to convert all the margins of F and F' to standard Pareto margins, resulting in Pareto copulas [245], and then compare their Pareto copulas using the $\le_{\text{tlod}}$ and $\le_{\text{tuod}}$ orders.

The preservation properties under the $\le_{\text{tlod}}$ and $\le_{\text{tuod}}$ orders can be easily established using Definitions 19.3.1 and 19.3.2, and Propositions 19.2.4 and 19.2.6. In particular, we have the following.

Proposition 19.3.4. *Let $\boldsymbol{X} = (X_1, \dots, X_d)$, $\boldsymbol{X'} = (X'_1, \dots, X'_d)$ and $\boldsymbol{Y} = (Y_1, \dots, Y_d)$, $\boldsymbol{Y'} = (Y'_1, \dots, Y'_d)$ be positive random vectors with support $\mathbb{R}^d_+$, and Θ be a random variable with bounded support. Assume that $(\boldsymbol{X}, \boldsymbol{X'})$ and $(\boldsymbol{Y}, \boldsymbol{Y'})$ are independent, and the regularity conditions of Proposition 19.2.6 are satisfied:*

1. *$\boldsymbol{X} \leq_{\mathrm{tlo}} \boldsymbol{X'}$ and $\boldsymbol{Y} \leq_{\mathrm{tlo}} \boldsymbol{Y'}$ imply that $\boldsymbol{X} \vee \boldsymbol{Y} \leq_{\mathrm{tlo}} \boldsymbol{X'} \vee \boldsymbol{Y'}$. $\boldsymbol{X} \leq_{\mathrm{tlod}} \boldsymbol{X'}$ and $\boldsymbol{Y} \leq_{\mathrm{tlod}} \boldsymbol{Y'}$ imply that $\boldsymbol{X} \vee \boldsymbol{Y} \leq_{\mathrm{tlod}} \boldsymbol{X'} \vee \boldsymbol{Y'}$.*

2. *$\boldsymbol{X} \leq_{\mathrm{tuo}} \boldsymbol{X'}$ and $\boldsymbol{Y} \leq_{\mathrm{tuo}} \boldsymbol{Y'}$ imply that $\boldsymbol{X} \wedge \boldsymbol{Y} \leq_{\mathrm{tuo}} \boldsymbol{X'} \wedge \boldsymbol{Y'}$. $\boldsymbol{X} \leq_{\mathrm{tuod}} \boldsymbol{X'}$ and $\boldsymbol{Y} \leq_{\mathrm{tuod}} \boldsymbol{Y'}$ imply that $\boldsymbol{X} \wedge \boldsymbol{Y} \leq_{\mathrm{tuod}} \boldsymbol{X'} \wedge \boldsymbol{Y'}$.*

3. *If $[\boldsymbol{X} \mid \Theta = \theta] \leq_{\mathrm{tlo}} [\boldsymbol{X'} \mid \Theta = \theta]$ for all θ in the bounded support of Θ, then $\boldsymbol{X} \leq_{\mathrm{tlo}} \boldsymbol{X'}$. If $[\boldsymbol{X} \mid \Theta = \theta] \leq_{\mathrm{tlod}} [\boldsymbol{X'} \mid \Theta = \theta]$ for all θ in the bounded support of Θ, then $\boldsymbol{X} \leq_{\mathrm{tlod}} \boldsymbol{X'}$.*

4. *If $[\boldsymbol{X} \mid \Theta = \theta] \leq_{\mathrm{tuo}} [\boldsymbol{X'} \mid \Theta = \theta]$ for all θ in the bounded support of Θ, then $\boldsymbol{X} \leq_{\mathrm{tuo}} \boldsymbol{X'}$. If $[\boldsymbol{X} \mid \Theta = \theta] \leq_{\mathrm{tuod}} [\boldsymbol{X'} \mid \Theta = \theta]$ for all θ in the bounded support of Θ, then $\boldsymbol{X} \leq_{\mathrm{tuod}} \boldsymbol{X'}$.*

Example 19.3.5. Let $\boldsymbol{X} \sim \mathcal{E}(\mathbf{0}, \Sigma_1, R_1)$ and $\boldsymbol{X'} \sim \mathcal{E}(\mathbf{0}, \Sigma_2, R_2)$ [see Eq. (19.2.5)], where $\Sigma_1 = A_1 A_1^\top = (\sigma_{ij})$ and $\Sigma_2 = A_2 A_2^\top = (\lambda_{ij})$. Consider $\boldsymbol{X}_+ = \boldsymbol{X} \vee \mathbf{0}$ and $\boldsymbol{X}'_+ = \boldsymbol{X'} \vee \mathbf{0}$:

1. Suppose that

   $$R_1 \leq_{\mathrm{sto}} R_2,\ \Sigma_1 \leq \Sigma_2 \text{ component-wise with } \sigma_{ii} = \lambda_{ii}, i = 1, \dots, d.$$

   It follows from Example 9.A.8 in [426] that $\boldsymbol{X}_+ \leq_{\mathrm{uo}} R_1(A_2\boldsymbol{U} \vee \mathbf{0})$, which implies that $\boldsymbol{X}_+ \leq_{\mathrm{tuo}} R_1(A_2\boldsymbol{U} \vee \mathbf{0})$. Clearly,

   $$(\underbrace{R_1, \dots, R_1}_{d}) \leq_{\mathrm{tuo}} (\underbrace{R_2, \dots, R_2}_{d}),$$

   which, together with Proposition 19.3.4 (4) and the fact that $A_2\boldsymbol{U}$ has a bounded support, imply that $\boldsymbol{X}_+ \leq_{\mathrm{tuo}} R_1(A_2\boldsymbol{U} \vee \mathbf{0}) \leq_{\mathrm{tuo}} R_2(A_2\boldsymbol{U} \vee \mathbf{0})$. Thus $\boldsymbol{X}_+ \leq_{\mathrm{tuo}} \boldsymbol{X}'_+$.

2. Suppose that

   $$R_1 \geq_{\mathrm{sto}} R_2,\ \Sigma_1 \leq \Sigma_2 \text{ component-wise with } \sigma_{ii} = \lambda_{ii}, i = 1, \dots, d.$$

   It follows from Example 9.A.8 in [426] that $\boldsymbol{X}_+ \leq_{\mathrm{lo}} R_1(A_2\boldsymbol{U} \vee \mathbf{0})$, which implies that $\boldsymbol{X}_+ \leq_{\mathrm{tlo}} R_1(A_2\boldsymbol{U} \vee \mathbf{0})$. Clearly,

   $$(\underbrace{R_1, \dots, R_1}_{d}) \leq_{\mathrm{tlo}} (\underbrace{R_2, \dots, R_2}_{d}),$$

   which, together with Proposition 19.3.4 (3) and the fact that $A_2\boldsymbol{U}$ has a bounded support, imply that $\boldsymbol{X}_+ \leq_{\mathrm{tlo}} R_1(A_2\boldsymbol{U} \vee \mathbf{0}) \leq_{\mathrm{tlo}} R_2(A_2\boldsymbol{U} \vee \mathbf{0})$. Thus $\boldsymbol{X}_+ \leq_{\mathrm{tlo}} \boldsymbol{X}'_+$. ∎

To construct a wide class of examples involving the $\le_{\text{tlod}}$ and $\le_{\text{tuod}}$ orders, we employ the copula approach. A copula C is a multivariate distribution with standard uniformly distributed margins on $[0,1]$. Sklar's theorem (see, e.g., [211], Sect. 1.6) states that every multivariate distribution F with margins $F_1,\dots,F_d$ can be written as $F(x_1,\dots,x_d) = C(F_1(x_1),\dots,F_d(x_d))$ for some d-dimensional copula C. In fact, in the case of continuous margins, C is unique and

$$C(u_1,\dots,u_d) = F(F_1^{-1}(u_1),\dots,F_d^{-1}(u_d))$$

where $F_i^{-1}(u_i)$ are the quantile functions of the i-th margin, $1 \le i \le d$. Let $(U_1,\dots,U_d)$ denote a random vector with $U_i, 1 \le i \le d$, being uniformly distributed on $[0,1]$. The survival copula $\widehat{C}$ is defined as follows:

$$\widehat{C}(u_1,\dots,u_n) = \mathrm{P}\{1-U_1 \le u_1,\dots,1-U_n \le u_n\} = \overline{C}(1-u_1,\dots,1-u_n) \tag{19.3.3}$$

where $\overline{C}$ is the joint survival function of C. The upper exponent and upper tail dependence functions (see [207, 213, 244, 360]) are defined as follows,

$$\begin{aligned} a(\boldsymbol{w};C) &:= \lim_{u\to 0^+} \frac{\mathrm{P}\{\cup_{i=1}^d \{U_i > 1-uw_i\}\}}{u}, \\ &\qquad \forall\ \boldsymbol{w} = (w_1,\dots,w_d) \in \mathbb{R}_+^d \backslash \{0\} \qquad (19.3.4) \\ b(\boldsymbol{w};C) &:= \lim_{u\to 0^+} \frac{\mathrm{P}\{\cap_{i=1}^d \{U_i > 1-uw_i\}\}}{u}, \\ &\qquad \forall\ \boldsymbol{w} = (w_1,\dots,w_d) \in \mathbb{R}_+^d \backslash \{0\} \qquad (19.3.5) \end{aligned}$$

provided that the limits exist. Note that both $a(\boldsymbol{w};C)$ and $b(\boldsymbol{w};C)$ are homogeneous of order 1 in the sense that $a(c\boldsymbol{w};C) = ca(\boldsymbol{w};C)$ and $b(c\boldsymbol{w};C) = cb(\boldsymbol{w};C)$ for any $c > 0$.

Theorem 19.3.6. *Let $\boldsymbol{X} = (X_1,\dots,X_d)$ and $\boldsymbol{X}' = (X_1',\dots,X_d')$ be two d-dimensional random vectors with respective copulas C and C' and their respective continuous margins $F_1,\dots,F_d$ and $F_1',\dots,F_d'$.*

1. *If $C = C'$ and $F_i \le_{\text{sto}} F_i'$, $1 \le i \le d$, then $\boldsymbol{X} \le_{\text{tuo}} \boldsymbol{X}'$.*

2. *Assume that the upper tail dependence functions $b(\cdot;C)$ and $b(\cdot;C')$ exist, and $\overline{F}_i \in RV_{-\alpha_i}$ and $\overline{F}_i' \in RV_{-\alpha_i'}$, $i = 1,\dots,d$. If F_i and F_i', $1 \le i \le d$, are all tail equivalent, and $b(\boldsymbol{w};C) < b(\boldsymbol{w};C')$ for all $\boldsymbol{w} = (w_1,\dots,w_d)$ with $w_i > 0$, $1 \le i \le d$, then $\boldsymbol{X} \le_{\text{tuod}} \boldsymbol{X}'$.*

3. *Assume that the upper tail dependence functions $b(\cdot; C)$ and $b(\cdot; C')$ exist, and $\overline{F}_i \in RV_{-\alpha_i}$ and $\overline{F}'_i \in RV_{-\alpha'_i}$, $i = 1, \dots, d$. If $F_i \leq_{\text{sto}} F'_i$, $1 \leq i \leq d$, and $b(\boldsymbol{w}; C) < b(\boldsymbol{w}; C')$ for all $\boldsymbol{w} = (w_1, \dots, w_d)$ with $w_i > 0$, $1 \leq i \leq d$, then $\boldsymbol{X} \leq_{\text{tuo}} \boldsymbol{X}'$.*

Proof:

(1) Since $F_i(tw_i) \geq F'_i(tw_i)$, $1 \leq i \leq d$, for all $t > t_{w_i}$, we have, for all $t > \max\{t_{w_i}, 1 \leq i \leq d\}$

$$
\begin{aligned}
& \mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \\
= \ & \mathrm{P}\{F_1(X_1) > F_1(tw_1), \dots, F_d(X_d) > F_d(tw_d)\} \\
\leq \ & \mathrm{P}\{F_1(X_1) > F'_1(tw_1), \dots, F_d(X_d) > F'_d(tw_d)\} \\
= \ & \mathrm{P}\{F'_1(X'_1) > F'_1(tw_1), \dots, F'_d(X'_d) > F'_d(tw_d)\} \\
= \ & \mathrm{P}\{X'_1 > tw_1, \dots, X'_d > tw_d\}.
\end{aligned}
$$

(2) Write $\overline{F}_i(t) = L_i(t)\, t^{-\alpha_i}$ and $\overline{F'}_i(t) = L'_i(t)\, t^{-\alpha'_i}$, $1 \leq i \leq d$. Since all the margins are tail equivalent, we have

$$\alpha_i = \alpha'_i =: \alpha, \text{ and } \lim_{t\to\infty} \frac{L_i(t)}{L_1(t)} = \lim_{t\to\infty} \frac{L'_i(t)}{L_1(t)} = 1, \ 1 \leq i \leq d.$$

In addition, since the functions $L_i(\cdot)$s and $L'_i(\cdot)$s are slowly varying, then for all $\boldsymbol{w} = (w_1, \dots, w_d)$ with $w_i > 0$, $1 \leq i \leq d$,

$$\lim_{t\to\infty} \frac{L_i(tw_i)}{L_1(t)} = \lim_{t\to\infty} \frac{L'_i(tw_i)}{L_1(t)} = 1, \ 1 \leq i \leq d.$$

That is, for any $\varepsilon > 0$, there exists t_w that is sufficiently large, such that, for $1 \leq i \leq d$ and all $t > t_w$,

$$(1-\varepsilon)L_1(t) \leq L_i(tw_i) \leq (1+\varepsilon)L_1(t),$$

$$(1-\varepsilon)L_1(t) \leq L'_i(tw_i) \leq (1+\varepsilon)L_1(t).$$

Observe that

$$
\begin{aligned}
& \mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \\
= \ & \mathrm{P}\big\{F_i(X_i) > 1 - L_i(tw_i)t^{-\alpha}w_i^{-\alpha}, 1 \leq i \leq d\big\} \\
\leq \ & \mathrm{P}\big\{F_i(X_i) > 1 - L_1(t)t^{-\alpha}(1+\varepsilon)w_i^{-\alpha}, 1 \leq i \leq d\big\},
\end{aligned}
$$

and thus

$$\begin{aligned}
&\limsup_{t\to\infty} \frac{\mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\}}{\overline{F}_1(t)} \\
\le\ & \lim_{t\to\infty} \frac{\mathrm{P}\big\{F_i(X_i) > 1 - \overline{F}_1(t)(1+\varepsilon)w_i^{-\alpha}, 1 \le i \le d\big\}}{\overline{F}_1(t)} \\
=\ & b\big((1+\varepsilon)\boldsymbol{w}^{-\alpha}; C\big) = (1+\varepsilon) b\big(\boldsymbol{w}^{-\alpha}; C\big).
\end{aligned}$$

Similarly,

$$\liminf_{t\to\infty} \frac{\mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\}}{\overline{F}_1(t)} \ge (1-\varepsilon) b\big(\boldsymbol{w}^{-\alpha}; C\big).$$

Let $\varepsilon \to 0$, we have

$$\lim_{t\to\infty} \frac{\mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\}}{\overline{F}_1(t)} = b\big(\boldsymbol{w}^{-\alpha}; C\big).$$

For $\boldsymbol{X}'$ with copula C', we have

$$\lim_{t\to\infty} \frac{\mathrm{P}\{X_1' > tw_1, \dots, X_d' > tw_d\}}{\overline{F}_1(t)} = b\big(\boldsymbol{w}^{-\alpha}; C'\big).$$

Since $b\big(\boldsymbol{w}^{-\alpha}; C\big) < b\big(\boldsymbol{w}^{-\alpha}; C'\big)$ for each $\boldsymbol{w} = (w_1, \dots, w_d)$ with $w_i > 0$, $1 \le i \le d$, there exists $t_{\boldsymbol{w}}$ such that, for all $t > t_{\boldsymbol{w}}$,

$$\mathrm{P}\{X_1 > tw_1, \dots, X_d > tw_d\} \le \mathrm{P}\{X_1' > tw_1, \dots, X_d' > tw_d\}.$$

(3) The stochastic tail order follows from (1) and (2). ■

For the $\le_{\text{tlo}}$ order, we can establish a similar result.

Theorem 19.3.7. *Let $\boldsymbol{X} = (X_1, \dots, X_d)$ and $\boldsymbol{X}' = (X_1', \dots, X_d')$ be d-dimensional random vectors with respective copulas C and C' and respective continuous margins $F_1, \dots, F_d$ and $F_1', \dots, F_d'$.*

1. *If $C = C'$ and $F_i \ge_{\text{sto}} F_i'$, $1 \le i \le d$, then $\boldsymbol{X} \le_{\text{tlo}} \boldsymbol{X}'$.*
2. *Assume that the exponent functions $a(\cdot; C)$ and $a(\cdot; C')$ exist, and $\overline{F}_i \in RV_{-\alpha_i}$ and $\overline{F}_i' \in RV_{-\alpha_i'}$, $i = 1, \dots, d$. If F_i and F_i', $1 \le i \le d$, are all tail equivalent, and $a(\boldsymbol{w}; C) > a(\boldsymbol{w}; C')$ for all $\boldsymbol{w} = (w_1, \dots, w_d)$ with $w_i > 0$, $1 \le i \le d$, then $\boldsymbol{X} \le_{\text{tlod}} \boldsymbol{X}'$.*
3. *Assume that the upper tail dependence functions $a(\cdot; C)$ and $a(\cdot; C')$ exist, and $\overline{F}_i \in RV_{-\alpha_i}$ and $\overline{F}_i' \in RV_{-\alpha_i'}$, $i = 1, \dots, d$. If $F_i \ge_{\text{sto}} F_i'$, $1 \le i \le d$, and $a(\boldsymbol{w}; C) > a(\boldsymbol{w}; C')$ for all $\boldsymbol{w} = (w_1, \dots, w_d)$ with $w_i > 0$, $1 \le i \le d$, then $\boldsymbol{X} \le_{\text{tlo}} \boldsymbol{X}'$.*

Example 19.3.8. The tail dependence functions of Archimedean copulas were derived in [34, 84] (also see Propositions 2.5 and 3.3 in [213]). Let $C_{\text{Arch}}(\boldsymbol{u};\phi) = \phi(\sum_{i=1}^d \phi^{-1}(u_i))$ be an Archimedean copula with strict generator ϕ^{-1}, where ϕ is regularly varying at ∞ with tail index $\theta > 0$. The upper tail dependence function of the survival copula $\widehat{C}_{\text{Arch}}$ is given by

$$b(\boldsymbol{w};\widehat{C}_{\text{Arch}}) = \Big(\sum_{j=1}^d w_j^{-1/\theta}\Big)^{-\theta}.$$

Observe that $b(\boldsymbol{w};\widehat{C}_{\text{Arch}})$ is strictly increasing in θ. For $\theta < \theta'$, and C and C' be two copulas with df $\widehat{C}_{\text{Arch}}$ having parameters θ and θ' respectively. Thus $b(\boldsymbol{w};C) < b(\boldsymbol{w};C')$ for all $\boldsymbol{w} > \boldsymbol{0}$. For $1 \le i \le d$, let F_i have the Fréchet df with tail index α (i.e., $F_i(s) = \exp\{-s^{-\alpha}\}$) and F_i' have the distribution of Pareto Type II with tail index α (i.e., $F_i'(s) = 1-(1+s)^{-\alpha}$). Clearly, F_i and F_i' are tail equivalent. Let $\boldsymbol{X}$ and $\boldsymbol{X}'$ have the dfs of

$$C(F_1(x_1),\dots,F_d(x_d)),\ \text{and}\ C'(F_1'(x_1),\dots,F_d'(x_d)),$$

respectively, and by Theorem 19.3.6, $\boldsymbol{X} \le_{\text{tuod}} \boldsymbol{X}'$. ∎

Example 19.3.9. The exponent functions of Archimedean copulas were derived in [34, 177] (also see Propositions 2.5 and 3.3 in [213]). Let $C_{\text{Arch}}(\boldsymbol{u};\phi) = \phi(\sum_{i=1}^d \phi^{-1}(u_i))$ be an Archimedean copula, where the generator ϕ^{-1} is regularly varying at 1 with tail index $\beta > 1$. The upper exponent function of C_{Arch} is given by

$$a(\boldsymbol{w};C_{\text{Arch}}) = \big(\sum_{j=1}^d w_j^{\beta}\big)^{1/\beta}.$$

Observe that $a(\boldsymbol{w};C_{\text{Arch}})$ is strictly decreasing in β. For $\beta < \beta'$, and C and C' be two copulas with df C_{Arch} having parameters β and β', respectively. Thus $a(\boldsymbol{w};C) > a(\boldsymbol{w};C')$ for all $w > 0$. For $1 \le i \le d$, let F_i have the Fréchet df with tail index α (i.e., $F_i(s) = \exp\{-s^{-\alpha}\}$) and F_i' have the distribution of Pareto Type II with tail index α (i.e., $F_i'(s) = 1-(1+s)^{-\alpha}$), such that F_i and F_i' are tail equivalent. Let $\boldsymbol{X}$ and $\boldsymbol{X}'$ have the dfs of

$$C(F_1(x_1),\dots,F_d(x_d)) \quad \text{and} \quad C'(F_1'(x_1),\dots,F_d'(x_d)),$$

respectively, and by Theorem 19.3.7, $\boldsymbol{X} \le_{\text{tlod}} \boldsymbol{X}'$. ∎

Remark 19.3.10. Due to the homogeneity property, the conditions on tail dependence and exponent functions used in Theorem 19.3.6 (2) and (3) and in Theorem 19.3.7 (2) and (3) can be simplified. For example, it is sufficient in Theorem 19.3.6 (2) and (3) to verify that $b(\boldsymbol{w}; C) < b(\boldsymbol{w}; C')$ for all $\boldsymbol{w} = (w_1, \dots, w_d)$ with $w_i > 0$, $1 \le i \le d$, and $||\boldsymbol{w}|| = 1$, where $|| \cdot ||$ denotes any norm on $\mathbb{R}^d_+$.

Theorem 19.3.11. *Let $(\boldsymbol{X}_n, n \ge 1)$ and $(\boldsymbol{X}'_n, n \ge 1)$ be two i.i.d. samples with dfs F and F', respectively. Assume that $F \in DA_\vee(G)$ and $F' \in DA_\vee(G')$ with tail equivalent Fréchet margins.*

1. *If $\boldsymbol{X}_n \le_{\text{tlod}} \boldsymbol{X}'_n$, then $G \le_{\text{lod}} G'$.*
2. *$G \le_{\text{lod}} G'$ if and only if $G \le_{\text{tlod}} G'$.*

Proof: Let $\boldsymbol{Y} = (Y_1, \dots, Y_d)$ and $\boldsymbol{Y}' = (Y'_1, \dots, Y'_d)$ denote two random vectors that have the same distributions as these of $\boldsymbol{X}_n$ and $\boldsymbol{X}'_n$, respectively.

(1) It follows from Theorem 19.1.1 and Remark 19.1.2 (1) that

$$[\mathrm{P}\{\boldsymbol{X}_k \le \boldsymbol{a_n x}\}]^n \to G(\boldsymbol{x}), \ \forall\ \boldsymbol{x} = (x_1, \dots, x_d) \in \mathbb{R}^d_+,$$

where $\boldsymbol{a}_n = (a_{1,n}, \dots, a_{d,n}) = (\overline{F}_1^{-1}(1/n), \dots, \overline{F}_d^{-1}(1/n))$. Taking the logarithm on both sides, we have, as $n \to \infty$,

$$n \log \mathrm{P}\{\boldsymbol{X}_k \le \boldsymbol{a_n x}\} \approx -n\mathrm{P}\big\{ \cup_{i=1}^d \{Y_i > a_{i,n}x_i\}\big\} \to \log G(\boldsymbol{x}). \tag{19.3.6}$$

Since the margins are tail equivalent, $a_{i,n}/a_{1,n} \to 1$ as $n \to \infty$. Thus, for any small $\varepsilon > 0$, when n is sufficiently large,

$$(1-\varepsilon)a_{1,n} \le a_{i,n} \le (1+\varepsilon)a_{1,n}, \ i = 1, \dots, n, \tag{19.3.7}$$

which imply that

$$\begin{aligned} -\,n\mathrm{P}\big\{ \cup_{i=1}^d \{Y_i > a_{1,n}x_i\}\big\} &\ge -n\mathrm{P}\big\{ \cup_{i=1}^d \{Y_i > a_{i,n}(1+\varepsilon)^{-1}x_i\}\big\} \\ &\to \log G((1+\varepsilon)^{-1}\boldsymbol{x}). \end{aligned}$$

Observing that $\log G(\boldsymbol{x})$ is homogeneous of order $-\alpha$, we have

$$\liminf_{n\to\infty} \big[-n\mathrm{P}\big\{ \cup_{i=1}^d \{Y_i > a_{1,n}x_i\}\big\}\big] \ge (1+\varepsilon)^\alpha \log G(\boldsymbol{x}).$$

Similarly,

$$\limsup_{n\to\infty}\big[-n\mathrm{P}\big\{\cup_{i=1}^{d}\{Y_i > a_{1,n}x_i\}\big\}\big] \le (1-\varepsilon)^{\alpha}\log G(\boldsymbol{x}).$$

Let $\varepsilon \to 0$, we obtain that

$$\lim_{n\to\infty}\big[-n\mathrm{P}\big\{\cup_{i=1}^{d}\{Y_i > a_{1,n}x_i\}\big\}\big] = \log G(\boldsymbol{x}). \tag{19.3.8}$$

That is, using the tail equivalence Eq. (19.3.7), we can rewrite the limit Eq. (19.3.6) in the form of Eq. (19.3.8), in which the scaling $a_{1,n}$ is the same for all margins. Working on $\boldsymbol{X}'_n$ in the same way, we also obtain that

$$\lim_{n\to\infty}\big[-n\mathrm{P}\big\{\cup_{i=1}^{d}\{Y'_i > a'_{1,n}x_i\}\big\}\big] = \log G'(\boldsymbol{x}), \tag{19.3.9}$$

where $\boldsymbol{a}'_n = (a'_{1,n},\ldots,a'_{d,n}) = (\overline{F'_1}^{-1}(1/n),\ldots,\overline{F'_d}^{-1}(1/n))$. Again, since the margins are tail equivalent, $a'_{1,n}/a_{1,n} \to 1$ as $n \to \infty$. Using the same idea as that of Eq. (19.3.7), the limit Eq. (19.3.9) is equivalent to

$$\lim_{n\to\infty}\big[-n\mathrm{P}\big\{\cup_{i=1}^{d}\{Y'_i > a_{1,n}x_i\}\big\}\big] = \log G'(\boldsymbol{x}). \tag{19.3.10}$$

Since $\boldsymbol{X}_n \le_{\text{tlod}} \boldsymbol{X}'_n$, via Eq. (19.3.1), we have

$$\mathrm{P}\big\{\cup_{i=1}^{d}\{Y_i > a_{1,n}x_i\}\big\} \ge \mathrm{P}\big\{\cup_{i=1}^{d}\{Y'_i > a_{1,n}x_i\}\big\}.$$

It follows from Eqs. (19.3.8) and (19.3.10) that $G(\boldsymbol{x}) \le G'(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^d_+$.

(2) It follows from the Pickands representation [see Remark 19.1.2 (4)] that

$$G(\boldsymbol{x}) = \exp\Big\{-c\int_{\mathbb{S}^d_+}\max_{1\le i\le d}\{(a_i/x_i)^{\alpha}\}\mathbb{Q}(\mathrm{d}\boldsymbol{a})\Big\}, \tag{19.3.11}$$

$$G'(\boldsymbol{x}) = \exp\Big\{-c'\int_{\mathbb{S}^d_+}\max_{1\le i\le d}\{(a_i/x_i)^{\alpha}\}\mathbb{Q}'(\mathrm{d}\boldsymbol{a})\Big\}, \tag{19.3.12}$$

where $c > 0, c' > 0$, and $\mathbb{Q}$ and $\mathbb{Q}'$ are probability measures defined on $\mathbb{S}^{d-1}_+$. Taking the scaling function t for both dfs, we have

$$G(t\boldsymbol{x}) = \exp\Big\{-\frac{c}{t^{\alpha}}\int_{\mathbb{S}^d_+}\max_{1\le i\le d}\{(a_i/x_i)^{\alpha}\}\mathbb{Q}(\mathrm{d}\boldsymbol{a})\Big\},$$

$$G'(t\boldsymbol{x}) \;=\; \exp\Big\{-\frac{c'}{t^\alpha}\int_{\mathbb{S}_+^d}\max_{1\le i\le d}\{(a_i/x_i)^\alpha\}\mathbb{Q}'(\mathrm{d}\boldsymbol{a})\Big\}.$$

For each fixed $\boldsymbol{x}$, when t is sufficiently large,

$$\frac{1-G(t\boldsymbol{x})}{1-G'(t\boldsymbol{x})} \sim \frac{c\int_{\mathbb{S}_+^d}\max_{1\le i\le d}\{(a_i/x_i)^\alpha\}\mathbb{Q}(\mathrm{d}\boldsymbol{a})}{c'\int_{\mathbb{S}_+^d}\max_{1\le i\le d}\{(a_i/x_i)^\alpha\}\mathbb{Q}'(\mathrm{d}\boldsymbol{a})}.$$

If $G \le_{\text{tlod}} G'$, then $1-G(t\boldsymbol{x}) \ge 1-G'(t\boldsymbol{x})$ for $t > t_{\boldsymbol{x}}$, where $t_{\boldsymbol{x}}$ is sufficiently large. That is,

$$c\int_{\mathbb{S}_+^d}\max_{1\le i\le d}\{(a_i/x_i)^\alpha\}\mathbb{Q}(\mathrm{d}\boldsymbol{a}) \ge c'\int_{\mathbb{S}_+^d}\max_{1\le i\le d}\{(a_i/x_i)^\alpha\}\mathbb{Q}'(\mathrm{d}\boldsymbol{a}),$$

which, together with Eqs. (19.3.11) and (19.3.12), imply that $G(\boldsymbol{x}) \le G'(\boldsymbol{x})$ for all $\boldsymbol{x}$.

Conversely, it is trivial that $G \le_{\text{lod}} G'$ implies that $G \le_{\text{tlod}} G'$. ■

Using similar arguments, we can also establish the upper orthant dependence comparisons for MEV distributions.

Theorem 19.3.12. *Let $(\boldsymbol{X}_n, n \ge 1)$ and $(\boldsymbol{X}'_n, n \ge 1)$ be two i.i.d. samples with dfs F and F' respectively. Assume that $F \in DA_\wedge(H)$ and $F' \in DA_\wedge(H')$ with tail equivalent, negative Fréchet margins (i.e., $F_i(x) = F'_i(x) = 1 - \exp\{-(-x)^{-\theta}\}$).*

1. *If $X_n \le_{\text{tuod}} X'_n$, then $H \le_{\text{uod}} H'$.*
2. *$H \le_{\text{uod}} H'$ if and only if $H \le_{\text{tuod}} H'$.*

As illustrated in Theorems 19.3.11 (2) and 19.3.12 (2), upper tail comparisons of MEV dfs trickle down to comparisons of entire distributions due to the scalable property of homogeneity.

We conclude this paper with an example to illustrate an idea for obtaining asymptotic Fréchet bounds.

Example 19.3.13. Let (T_1, T_2) denote a random vector with a Marshall-Olkin distribution [307] as defined as follows:

$$T_1 = E_1 \wedge E_{12},\; T_2 = E_2 \wedge E_{12},$$

where E_1, E_2, E_{12} are i.i.d. exponentially distributed with unit mean. Clearly, for $t_1 > 0$, or $t_2 > 0$,

$$\mathrm{P}\{T_1 > t_1, T_2 > t_2\} = e^{-(t_1+t_2)-t_1\vee t_2} < e^{-t_1\vee t_2} = \mathrm{P}\{T_1 > t_1, T_1 > t_2\}.$$

Hence $(T_1, T_2) \leq_{\mathrm{uod}} (T_1, T_1)$, which is known as the Fréchet upper bound for the class of dfs with fixed exponential margins.

Let R_1 and R_2 denote two nonnegative random variables that are independent of (T_1, T_2). Assume that the survival functions of R_1^{-1} and R_2^{-1} are tail equivalent, regularly varying with tail index $-\alpha$ [see Eq. (19.1.6)]. Since R_1^{-1} and R_2^{-1} are tail equivalent, the margins of $(R_2^{-1}T_1, R_2^{-1}T_2)$ and $(R_1^{-1}T_1, R_1^{-1}T_1)$ are all tail equivalent. It follows from Theorem 3.2 of [280] that the upper tail dependence functions of $(R_2^{-1}T_1, R_2^{-1}T_2)$ and $(R_1^{-1}T_1, R_1^{-1}T_1)$ are given by

$$\begin{aligned} b(w_1, w_2) &= 2^{\alpha}(w_1 + w_2 + w_1 \vee w_2)^{-\alpha} \\ b'(w_1, w_2) &= 2^{\alpha}(w_1 \vee w_2)^{-\alpha}. \end{aligned}$$

Clearly, $b(w_1, w_2) < b'(w_1, w_2)$, and thus by Theorem 19.3.6 we have $(R_2^{-1}T_1, R_2^{-1}T_2) \leq_{\mathrm{tuod}} (R_1^{-1}T_1, R_1^{-1}T_1)$. Note that the df of $(R_1^{-1}T_1, R_1^{-1}T_1)$ is viewed as an asymptotic Fréchet upper bound in the sense of tail upper orthant order, because the respective margins of $(R_2^{-1}T_1, R_2^{-1}T_2)$ and $(R_1^{-1}T_1, R_1^{-1}T_1)$ are only tail equivalent, rather than being identical as required in the case of Fréchet bounds. ■

Acknowledgments: The author would like to thank Alfred Müller for his comments on this paper during IWAP 2012 in Jerusalem, Israel, and especially for his comment on several references that are closely related to this work. Haijun Li is supported by NSF grants CMMI 0825960 and DMS 1007556.

Bibliography

[1] Abbasnejad, M.: Some characterization results based on dynamic survival and failure entropies. Communications of the Korean Statistical Society, **18**, 1–12 (2011)

[2] Abramowitz, M. and Stegun, I. A.: Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover publications (1964)

[3] Acerbi, C.: Spectral measures of risk: a coherent representation of subjective risk aversion. Journal of Banking and Finance, **26** (7), 1505–1518 (2002)

[4] Acerbi, C. and Tasche, D.: On the coherence of expected shortfall. Journal of Banking & Finance, **26** (7), 1487–1503 (2002)

[5] Ahmad, I. A. and Kayid, M.: Characterizations of the RHR and MIT orderings and the DRHR and IMIT classes of life distributions. Probability in the Engineering and Informational Sciences, **19**, 447–461 (2005)

[6] Ahmad, I. A., Kayid, M. and Pellerey, F.: Further results involving the MIT order and IMIT class. Probability in The Engineering and Informational Sciences, **19**, 377–395 (2005)

[7] Ali, S. M. and Silvey, S. D.: Association between random variables and the dispersion of a Radon-Nikodym derivative. Journal of the Royal Statistical Society, Series B, **27**, 100–107 (1965)

[8] Ali, S. M. and Silvey, S. D.: A further result about the relevance of the dispersion of a Radon-Nikodym derivative to the problem of measuring association. Journal of the Royal Statistical Society, Series B, **27**, 108–110 (1965)

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9,

[9] Aly, E.-E.A.A.: On some confidence bands for percentile residual life functions. Journal of Nonparametric Statistics, **2**, 59–70 (1992)

[10] American Society for Testing and Materials.: ASTM Manual on Presentation of Data and Control Chart Analysis. ASTM Special Technical Publication 15D. ASTM, Philadelphia, Pennsylvania, USA (1976)

[11] Amiri, L., Khaledi, B.-E. and Samaniego, F. J.: On skewness and dispersion among convolutions of independent gamma random variables. Probability in the Engineering and Informational Sciences, **25**, 55–69 (2011)

[12] Anderson, T. W.: The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proceedings of the American Mathematical Society, **6**, 170–176 (1955)

[13] Arjas, E.: A stochastic process approach to multivariate reliability systems: notions based on conditional stochastic order. Mathematics of Operations Research, **6**, 263–276 (1981)

[14] Arjas, E. and Norros, I.: Life lengths and association: a dynamic approach. Mathematical Operation Research, **9**, 151–158 (1984)

[15] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N.: Records. Wiley Series in Probability and Statistics (1998)

[16] Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N.: A First Course in Order Statistics. SIAM, Philadelphia (2008)

[17] Asadi, M.: On the mean past lifetime of the components of a parallel system. Journal of Statistical Planning and Inference, **136**, 1197–1206 (2006)

[18] Asadi, M. and Bairamov, I.: A note on the mean residual life function of a parallel system. Communications in Statistics – Theory and Methods, **34**, 475–484 (2005)

[19] Asadi, M. and Ebrahimi, N.: Residual entropy and its characterizations in terms of hazard function and mean residual life function. Statistics & Probability Letters, **49**, 263–269 (2000)

[20] Asadi, M. and Zohrevand, Y.: On the dynamic cumulative residual entropy. Journal of Statistical Planning and Inference, **137**, 1931–1941 (2007)

[21] Asmussen, S.: Applied probability and queues. Springer-Verlag, New York (2003)

[22] Asmussen, S. and Albrecher, H.: Ruin Probabilities (second edition). World Scientific Publishing Company (2010)

[23] Avérous, J., Genest, C., and Kochar, S. C.: On the dependence structure of order statistics. Journal of Multivariate Analysis, **94**, 159–171 (2005)

[24] Bagai, I. and Kochar, S. C.: On tail ordering and comparison of failure rates. Communications in Statistics-Theory and Methods, **15**, 1377–1388 (1986)

[25] Balakrishnan, N.: Permanents, order statistics, outliers and robustness. Revista Matematica Complutense, **20**, 7–107 (2007)

[26] Balakrishnan, N. and Basu, A. P.: The Exponential Distribution: Theory, Methods and Applications. Newark, New Jersey: Gordon and Breach Publishers (1995)

[27] Balakrishnan, N., Belzunce, F., Hami, N. and Khaledi, B.-E.: Univariate and multivariate likelihood ratio ordering of generalized order statistics and associated conditional variables. Probability in the Engineering and Informational Sciences, **24**, 441–455 (2010)

[28] Balakrishnana, N., Belzunce, F., Sordo, M. A. and Suárez-Llorensc, A.: Increasing directionally convex orderings of random vectors having the same copula, and their use in comparing ordered data. Journal of Multivariate Analysis, **105**, 45–54 (2012)

[29] Balakrishnan, N. and Rao, C. R.: Order Statistics: Theory and Methods. Amsterdam: Elsevier (1998)

[30] Balakrishnan, N. and Rao, C. R.: Order Statistics: Applications. Amsterdam: Elsevier (1998)

[31] Bapat, R. B. and Beg, M. I.: Order statistics for nonidentically distributed variables and permanents, Sankhyá (Ser. A), **51**, 78–93 (1989)

[32] Bapat, R. B. and Kochar, S. C.: On likelihood-ratio ordering of order statistics. Linear Algebra and its Applications, **199**, 281–291 (1994)

[33] Barabás, B., Csörgö, M., Horváth, L. and Yandell, B. S.: Bootstrapped confidence bands for percentile lifetime. Annals of the Institute of Statistical Mathematics, **38**, 429–438 (1986)

[34] Barbe, P., Fougères, A. L. and Genest, C.: On the tail behavior of sums of dependent risks. ASTIN Bulletin, **36**, 361–374 (2006)

[35] Barlow, R. E. and Campo, R.: Total time on test processes and applications to failure data analysis. Reliability and Fault Tree Analysis (edited by R. E. Barlow, J. B. Fussell, and N. D. Singpurwalla), SIAM, Philadephia, 451–481 (1975)

[36] Barlow, R. E., Bartholomew, D. J., Bremner, J. M. and Brunk, H. D.: Statistical Inference Under Order Restrictions. John Wiley, New York (1972)

[37] Barlow, R. E. and Proschan, F.: Inequalities for linear combinations of order statistics from restricted families. Annals of Mathematical Statistics, **38**, 1574–1592 (1966)

[38] Barlow, R. E. and Proschan, F.: Statistical Theory of Reliability and Life Testing: Probability Models. Silver Spring, Maryland: To Begin With (1975)

[39] Barlow, R. E. and Proschan, F.: Statistical Theory of Reliability and Life Testing. To Begin With, Silver Spring, Maryland (1981)

[40] Barnett, V. and Lewis, T.: Outliers in statistical data. John Wiley and Sons Ltd., Chichester (1994)

[41] Barrett, G. and Donald, S.: Consistent Tests for Stochastic Dominance. Econometrica, **71**, 71–104 (2003)

[42] Bartosewicz, J.: Dispersive ordering and the total time on test transform. Statistics and Probability Letters, **4**, 285–288 (1986)

[43] Bartoszewicz, J.: Tail orderings and the total time on test transforms. Applicationes Mathematicae **24**, 77–86 (1996)

[44] Bassan, B. and Spizzichino, F.: Relations among univariate aging, bivariate aging and dependence for exchangeable lifetimes. Journal of Multivariate Analysis, **93**, 313–339 (2005)

[45] Batson, R. G., Jeong, Y., Fonseca, D. J. and Ray, P. S.: Control Charts for Monitoring Field Data.: Quality and Reliability Engineering International. **22**, 733–755 (2006)

[46] Bäuerle, N.: Inequalities for stochastic models via supermodular orderings. Communications in Statistics-Stochastic Models, **13**, 181–201 (1997)

[47] Bäuerle, N. and Müller, A.: Modelling and comparing dependencies in multivariate risk portfolios. ASTIN Bulletin, **28**, 59–76 (1998)

[48] Belzunce, F., Franco, M. and Ruiz, J. M.: On aging properties based on the residual life of k-out-of-n systems. Probability in the Engineering and Informational Sciences, **13**, 193–199 (1999)

[49] Belzunce, F., Gao, X., Hu, T. and Pellerey, F.: Characterizations of the hazard rate order and IFR aging notion. Statistics and Probability Letters, **70**, 235–242 (2004)

[50] Belzunce, F., Gurler, S. and Ruiz, J. M.: Revisiting multivariate likelihood ratio ordering results for order statistics. Probability in the Engineering and Informational Sciences, **25**, 355–368 (2011)

[51] Belzunce, F., Lillo, R. E., Ruiz, J. M. and Shaked, M.: Stochastic comparisons of nonhomogeneous processes. Probability in the Engineering and Informational Sciences, **15**, 199–224 (2001)

[52] Belzunce, F., Lillo, R., Ruiz, J. M. and Shaked, M.: Stochastic ordering of record and inter-record values. Recent developments in ordered random variables, 119–137, Nova Science Publishers, New York (2007)

[53] Belzunce, F., Martínez-Puertas, H. and Ruiz, J. M.: On optimal allocation of redundant components for series and parallel systems of two dependent components. Journal of Statistical Planning and Inference, **141**, 3094–3104 (2011)

[54] Belzunce, F., Mercader, J. A. and Ruiz, J. M.: Stochastic comparisons of generalized order statistics. Probability in the Engineering and Informational Sciences, **19**, 99–120 (2005)

[55] Belzunce, F., Pinar, J. F. and Ruiz, J. M.: A family of tests for right spread order. Statistics and Probability Letters, **54**, 79–92 (2001)

[56] Belzunce, F., Ruiz, J. M. and Suárez-Llorens, A.: On multivariate dispersion orderings based on the standard construction. Statistics and Probability Letters, **78**, 271–281 (2008)

[57] Belzunce, F. and Shaked, M.: Stochastic comparisons of mixtures of convexly ordered distributions with applications in reliability theory. Statistics and Probability Letters, **53**, 363–372 (2001)

[58] Benneyan, J.: Misuse of XmR quality control charts for common single parameter probability distributions. International Conference on Computers & Industrial Engineering, University of Technology of Troyes, France, July 6–8, 2009 1291–1296 (2009)

[59] Bergman, B.: On age replacement and the total time on test concept. Scandinavian Journal of Statistics, **6**, 161–168 (1979)

[60] Berrendero, J. R. and Cárcamo, J.: Tests for the second order stochastic dominance based on l-statistics. Journal of Business and Economic Statistics, **29**, 260–270 (2011)

[61] Bingham, N. H., Goldie, C. M. and Teugels, J. L.: Regular Variation. Cambridge University Press, Cambridge (1987)

[62] Block, H. W., Sampson, A. R. and Savits, T. H.: Topics in Statistical Dependence. Lecture Notes-Manograph Series, IMS, Hayward, CA (1990)

[63] Block, H. W., Savits, T. H. and Singh, H.: The reversed hazard rate function. Probability in the Engineering and Informational Sciences, **12**, 69–90 (1998)

[64] Bock, M. E., Diaconis, P., Huffer, F. W. and Perlman, M. D.: Inequalities for linear combinations of Gamma random variables. The Canadian Journal of Statistics, **15**, 387–395 (1987)

[65] Boland, P. J.: Signatures of indirect majority systems. Journal of Applied Probability, **38**, 597–603 (2001)

[66] Boland, P. J. and El-Newehi, E.: Component redundancy vs. system redundancy in the hazard rate ordering. IEEE Transactions on Reliability, **44**, 614–619 (1995)

[67] Boland, P. J., EI-Newehi, E. and Proschan, F.: Optimal allocation of components in parallel-seires and series-parallel systems. Journal of Applied Probability, **23**, 770–777 (1986)

[68] Boland, P. J., EI-Newehi, E. and Proschan, F.: Active redundancy allocation in coherent systems. Probability in the Engineering and Informational Science, **2**, 343–353 (1988)

[69] Boland, P. J., EI-Newehi, E. and Proschan, F.: Redundancy importance and allocation of spares in coherent systems. Journal of Statistical Planning and Inference, **29**, 55–66 (1991)

[70] Boland, P. J., EI-Newehi, E. and Proschan, F.: Stochastic order for redundancy allocations in series and parallel systems. Advancesin Applied Probability, **24**, 161–171 (1992)

[71] Boland, P. J., EI-Newehi, E. and Proschan, F.: Schur properties of convolutions of exponential and geometric random variables. Journal of Multivariate Analysis, **48**, 157–167 (1994)

[72] Boland, P. J., EI-Newehi, E. and Proschan, F.: Applications of hazard rate ordering in reliability and order statistics. Journal of Applied Probability, **31**, 180–192 (1994)

[73] Boland, P. J., Hu, T., Shaked, M. and Shanthikumar, J. G.: Stochastic ordering of order statistics II. Modelling Uncertainty: An Examination of Stochastic Theory, Methods and Applications (Eds. M. Dror, P. L'Ecuyer and F. Szidarovszky, Kluwer, Boston), 607–623 (2002)

[74] Boland, P. J., Shaked, M. and Shanthikumar, J. G.: Stochastic ordering of order statistics. Handbook of Statistics (Eds. N. Balakrishnan and C.R. Rao), **16**, 89–103 (1998)

[75] Bon, J. L. and Păltănea, E.: Ordering properties of convolutions of exponential random variables. Lifetime Data Analysis, **5**, 185–192 (1999)

[76] Border, K. C.: Functional analytic tools for expected utility theory. In Positive Operators, Riesz Spaces, and Economics (Pasadena, CA, 1990), 69–88. Springer, Berlin (1991)

[77] Borgan, Ø.: Nelson-Aalen estimator. Online Encyclopedia of Biostatistics, John Wiley & Sons, Ltd (2005)

[78] Bowden, R. J.: Information, measure shifts and distribution metrics. Statistics, **46**, 249–262 (2010)

[79] Brown, M. and Chaganty, N. R.: On the first passage time distribution for a class of Markov chains. Annals of Probability, **8**, 1000–1008 (1983)

[80] Bryson, M. C. and Siddiqui, M. M.: Some criteria for aging. Journal of the American Statistical Association, **64**, 1472–1483 (1969)

[81] Bühlmann, H.: Mathematical methods in risk theory. Springer-Verlag, New York (1970)

[82] Cha, J. H., Mi, J. and Yun, W. Y.: Modelling a general standby system and evaluation of its performance. Applied Stochastic Models in Business and Industry, **24**, 159–169 (2008)

[83] Chandler, K. N.: The distribution and frequency of record values. Journal of the Royal Statistical Society, Series B, **14**, 220–228 (1952)

[84] Charpentier, A. and Segers, J.: Tails of multivariate Archimedean copulas. Journal of Multivariate Analysis, **100**, 1521–1537 (2009)

[85] Chen, H. and Cheng, Y.: Non-normality effects on the economic-statistical design of Xbar charts with Weibull in-control time.: European Journal of Operational Research, **176**, 986–998 (2007)

[86] Chen, J. and Hu, T.: Multivariate dispersive ordering of generalized order statistics. Journal of Iranian Statistical Society, **6**, 61–75 (2007)

[87] Chen, H. and Hu, T.: Multivariate likelihood ratio orderings between spacings of heterogeneous exponential random variables. Metrika, **68**, 17–29 (2008)

[88] Cheng, S. and Jiang, C.: The Edgeworth expansion for distributions of extreme values. Science in China, **44**, 427–437 (2001)

[89] Choquet, G.: Lectures on Analysis. Vol. II: Representation Theory. Benjamin, W A Inc., New York-Amsterdam (1969)

[90] Chung, C. F.: Confidence bands for percentile residual lifetime under random censorship model. Journal of Multivariate Analysis, **29**, 94–126 (1989)

[91] Colangelo, A., Hu, T. and Shaked, M.: Conditional orderings and positive dependence. Journal of Multivariate Analysis, **99**, 358–371 (2008)

[92] Colangelo, A., Scarsini, M. and Shaked, M.: Some positive dependence stochastic orders. Journal of Multivariate Analysis, **97**, 46–78 (2006)

[93] Coles, S., Heffernan, J. and Tawn, J.: Dependence Measures for Extreme Value Analyses. Extremes, **2**, 339–365 (1999)

[94] Cossette, H., Mailhot, M. and Marceau, E.: TVaR-based capital allocation for multivariate compound distributions. Insurance: Mathematics and Economics, **50** (2), 247–256 (2012)

[95] Cossette, H. and Marceau, E.: The discrete time risk model with correlated classes of business. Insurance: Mathematics and Economics, **26**, 133–149 (2000)

[96] Cover, T. M. and Thomas, J. A.: Elements of Information Theory. Wiley, New York (1991)

[97] Crowder, S. V.: A Simple Method for Studying Run-Length Distributions of Exponentially Weighted Moving Average Charts. Technometrics, **29**, 401–407 (1987)

[98] Csörgö, M. and Csörgö, S.: Estimation of percentile residual life. Operations Research, **35**, 598–606 (1987)

[99] Csörgö, S. and Mason, D. M.: Bootstrapping empirical functions. Annals of Statistics, **17**, 1447–1471 (1989)

[100] Csörgö, S. and Viharos, L.: Confidence bands for percentile residual lifetimes. Journal of Statistical Planning and Inference, **30**, 327–337 (1992)

[101] Dabrowska, D.: Regression-based orderings and measures of stochastic dependence. Statistics, **12**, 317–325 (1981)

[102] Da, G., Xia, L. and Hu, T.: On computing signatures of k-out-of-n systems consisting of modules. Methodology and Computing in Applied Probability, DOI:10.1007/s11009-012-9308-5 (2012)

[103] Da, G., Zheng, B. and Hu, T.: On computing signatures of coherent systems. Journal of Multivariate Analysis, **103**, 142–150 (2012)

[104] Das, B. and Ghosh, S.: Weak limits of exploratory plots in Extreme Value Analysis. Bernoulli, to appear (2012)

[105] Davidov, O. and Herman, A.: New tests for stochastic order with application to case control studies. Journal of Statistical Planning and Inference, **139**, 2614–2623 (2009)

[106] Davidov, O. and Herman, A.: Multivariate stochastic orders induced by case-control sampling. Methodology and Computing in Applied Probability, **13**, 139–154 (2011)

[107] David, H. A. and Nagaraja, H. N.: Order statistics (3rd ed). Wiley, New York (2003)

[108] Davison, A. C. and Smith, R. L.: Models for exceedances over high thresholds (with discussion). Journal of the Royal Statistical Society, Series B, **52**, 393–442 (1990)

[109] Decancq, K.: Elementary multivariate rearrangements and stochastic dominance on a Fréchet class. Journal of Economic Theory, **147**, 1450–1459 (2012)

[110] De Finetti, B.: Su un'impostazione alternativa della teoria collettiva del rischio. Transactions of the XVth International Congress of Actuaries, **2**, 433–443 (1957)

[111] de Haan, L.: On Regular Variation and Its Application to The Weak Convergence of Sample Extremes. Mathematisch Centrum (1970)

[112] de Haan, L.: Equivalence classes of regularly varying functions. Stochastic Processes and their Applications, **2**, 243–259 (1974)

[113] de Haan, L. and Ferreira, A.: Extreme Value Theory: An introduction. Springer, New York (2006)

[114] de Haan, L. and Stadtmüller, U.: Generalized regular variation of second order. Journal of the Australian Mathematical Society, **61**, 381–395 (1996)

[115] Deheuvels, P.: The characterization of distributions by order statistics and record values - a unified approach. Journal of Applied Probability, **21**, 326–334 (1984)

[116] Denneberg, D.: Non-Additive Measure and Integral. Kluwer Academic Publishers, Dordrecht, The Netherlands (1994)

[117] Denuit, M., Dhaene, J., Goovaerts, M. J. and Kaas, R.: Actuarial Theory for Dependent Risks – Measures, Orders and Models. Wiley, New York (2005)

[118] Denuit, M., Genest, C. and Marceau, E.: Criteria for the stochastic ordering of random sums, with actuarial applications. Scandinavian Actuarial Journal, 3–16 (2002)

[119] Denuit, M., Goderniaux, A.-C. and Scaillet, O.: A Kolmogorov-Smirnov type test for shortfall dominance against parametric alternatives. Technometrics, **49**, 88–99 (2007)

[120] Denuit, M. and Müller, A.: Smooth generators of integral stochastic orders. Annals of Applied Probability, **12**, 1174–1184 (2002)

[121] Deshpande, J. V., Kochar, S. C. and Singh, H.: Aspects of positive ageing. Journal of Applied Probability, **23**, 748–758 (1986)

[122] Dette, H., Siburg, K. F. and Stoimenov, P. A.: A copula-based non-parametric measure of regression dependence. Scandinavian Journal of Statistics, to appear (2012)

[123] Diaconis, P. and Perlman, M. D.: Bounds for tail probabilities of linear combinations of independent gamma random variables. The Symposium on dependence in Statistics and Probability, Hidden Valley, Pennsylvania (1987)

[124] Dickson, D. C. M.: Insurance Risk and Ruin. Cambridge University Press, Cambridge (2005)

[125] Di Crescenzo, A.: A probabilistic analogue of the mean value theorem and its applications to reliability theory. Journal of Applied Probability, **36**, 706–719 (1999)

[126] Di Crescenzo, A. and Longobardi, M.: Entropy-based measure of uncertainty in past lifetime distributions. Journal of Applied Probability, **39**, 434–440 (2002)

[127] Di Crescenzo, A. and Longobardi, M.: On weighted residual and past entropies. Scientiae Mathematicae Japonicae, **64**, 255–266 (2006)

[128] Di Crescenzo, A. and Longobardi, M.: On cumulative entropies. Journal of Statistical Planning and Inference, **139**, 4072–4087 (2009)

[129] Di Crescenzo, A. and Longobardi, M.: On cumulative entropies and lifetime estimations. Methods and Models in Artificial and Natural Computation, Lecture Notes in Computer Science, Vol. 5601 (Mira J., Ferrandez J. M., Alvarez Sanchez J. R., Paz F., Toledo J. eds.), 132–141. Springer-Verlag, Berlin (2009)

[130] Di Crescenzo, A. and Longobardi, M.: More on cumulative entropy. Cybernetics and Systems 2010, (Trappl R. ed.), 181–186. Austrian Society for Cybernetic Studies, Vienna (2010)

[131] Di Crescenzo, A. and Longobardi, M.: Neuronal data analysis based on the empirical cumulative entropy. Computer Aided Systems Theory, EUROCAST 2011, Part I, Lecture Notes in Computer Science, Vol. 6927 (Moreno-Diaz R., Pichler F., Quesada-Arencibia A. eds.), 72–79. Springer-Verlag, Berlin (2012)

[132] Di Crescenzo, A. and Shaked, M.: Some applications of the Laplace transform ratio order. Arab Journal of Mathematical Sciences, **2**, 121–128 (1996)

[133] Ding, W. and Li, X.: Optimal allocation of active redundancies to k-out-of-n systems with respect to the hazard rate order. Journal of Statistical Planning and Inference, **142**, 1878–1887 (2012)

[134] Dolati, A., Genest, C. and Kochar, S. C.: On the dependence between the extreme order statistics in the proportional hazards model. Journal of Multivariate Analysis, **99**, 777–786 (2008)

[135] Drees, H.: On smooth statistical tail functionals. Scandinavian Journal of Statistics, **25**, 187–210 (1998)

[136] Duncan, A. J.: Quality Control and Industrial Statistics, Irwin, Homewood, Illinois, USA (1974)

[137] Durante, F., Foschi, R. and Spizzichino, F.: Threshold copulas and positive dependence. Statistics and Probability Letters, **17**, 2902–2909 (2008)

[138] Dyckerhoff, R. and Mosler, K.: Stochastic dominace with non-additive probabilities. Methods and Models of Operations Research, **37**, 231–256 (1993)

[139] Ebrahimi, N.: How to measure uncertainty in the residual life time distribution. Sankhyä, Series A, **58**, 48–56 (1996)

[140] Ebrahimi, N. and Kirmani, S. N. U. A.: Some results on ordering of survival functions through uncertainty. Statistics & Probability Letters, **29**, 167–176 (1996)

[141] Ebrahimi, N. and Pellerey, F.: New partial ordering of survival functions based on the notion of uncertainty. Journal of Applied Probability, **32**, 202–211 (1995)

[142] Ebrahimi, N., Soofi, E. S. and Soyer, R.: Information measures in perspective. International Statistical Review, **78**, 383–412 (2010)

[143] Efrosinin, D.: Controlled queueing systems with heterogeneous servers: Dynamic optimization and monotonicity properties. Saarbrücken: VDM Verlag (2008)

[144] Efrosinin, D. and Breuer, L.: Threshold policies for controlled retrial queues with heterogeneous servers. Annals of Operations Research, **141**, 139–162 (2006)

[145] EI-Newehi, E., Proschan, F. and Sethuraman, J.: (1978). Multistate coherent system. Journal of Applied Probability, **15**, 675–688 (1978)

[146] EI-Newehi, E., Proschan, F. and Sethuraman, J.: Optimal allocation of components in parallel-series and series-parallel systems. Journal of Applied Probability, **23**, 770–777 (1986)

[147] EI-Newehi, E. and Sethuraman, J.: Optimal allocation under partial ordering of lifetimes of components. Advances in Applied Probability, **25**, 914–925 (1993)

[148] Elton, J. and Hill, T. P.: Fusions of a probability distribution. Annals of Probability, **20**, 421–454 (1992)

[149] Elton, J. and Hill, T. P.: On the basic representation theorem for convex domination of measures. Journal of Mathematical Analysis and Applications, **228**, 449–466 (1998)

[150] Embrechts, P., Kluppelberg, C. and Mikosch, T.: Modelling Extreme Events for Insurance and Finance. Springer-Verlag, Berlin (1997)

[151] Eryilmaz, S. and Zuo, M.: Computing and applying the signature of a system with two common failure critera. IEEE Transaction on Reliability, **59**, 576–580 (2010)

[152] Esary, J. D. and Marshall, A. W.: Coherent life functions. SIAM Journal of Applied Mathematics, **18**, 810–814 (1970)

[153] Esary, J. D., Marshall, A. W. and Proschan, F.: Some reliability applications of the hazard transform. SIAM Journal of Applied Mathematics, **18**, 849–860 (1970)

[154] Fagiuoli, E., Pellerey, F. and Shaked, M.: A characterization of the dilation order and its applications. Statistical Papers, **40**, 393–406 (1999)

[155] Fahmi, S., Pereira, C. A. de B., Proschan, F. and Shaked, M.: The influence of the sample on the posterior distribution. Communications in Statistics - Theory and Methods, **11**, 1757–1768 (1982)

[156] Fang, Z., Hu, T., Wu, Y. and Zhuang, W.: Multivariate stochastic orderings of spacings of generalized order statistics. Chinese Journal fo Applied Probability and Statistics, **22**, 295–303 (2006)

[157] Fang, Z. and Joe, H.: Further developments on some dependence orderings for continuous bivariate distributions. Annals of the Institute of Statistical Mathematics, **44**, 501–517 (1992)

[158] Fang, K. T., Kotz, S. and Ng, K. W.: Symmetric Multivariate and Related Distributions, Monographs on Statistics and Applied Probability, 36. Chapman & Hall, London (1990)

[159] Fefferman, C., Jodeit, M. and Perlman, M. D.: A spherical surface measure inequality for convex sets. Proceedings of the American Mathematical Society, **33**, 114–119 (1972).

[160] Fernández-Ponce, J. M., Kochar, S. C. and Muñoz-Perez, J.: Partial orderings of distributions based on right-spread functions. Journal of Applied Probability, **35**, 221–228 (1998)

[161] Fernandez-Ponce, J. M. and Suárez-Llorens, A.: A multivariate dispersion ordering based on quantiles more widely separated. Journal of Multivariate Analisys, **85**, 40–53 (2003)

[162] Fleming, T. R. and Harrington, D. P.: Counting Processes and Survival Analysis. Wiley, New York (1991)

[163] Fraga Alves, M. I., Gomes, M. I., de Haan, L. and Neves, C.: A note on second order conditions in extreme value theory: Linking general and heavy tails conditions. Revstat, **5**, 285–304 (2007)

[164] Franco-Pereira, A. M. and de Uña-Álvarez, J.: Estimation of a monotone percentile residual life function under random censorship. Technical Report, Department of Statistics and OR, University of Vigo (2012)

[165] Franco-Pereira, A. M., Lillo, R. E. and Romo, J.: Characterization of bathtub distributions via percentile residual life functions. UC3M Working papers. Statistics and Econometrics, **10–12** (2010)

[166] Franco-Pereira, A. M., Lillo, R. E., Romo, J. and Shaked, M.: The percentile residual life orders. Applied Stochastic Models in Business and Industry, **27**, 235–252 (2010)

[167] Franco-Pereira, A. M., Lillo, R. E. and Shaked, M.: The decreasing percentile residual life aging notion. Statistics, to appear (2011)

[168] Franco, M., Ruiz, J. M. and Ruiz, M. C.: Stochastic orderings between spacings of generalized order statistics. Probability in the Engineering and Informational Sciences, **16**, 471–484 (2001)

[169] Franco, M., Ruiz, J. M. and Ruiz, M. C.: On the closure of the IFR(2) and NBU(2) classes. Journal of Applied Probability, **38**, 235–241 (2001)

[170] Furman, E.: On a multivariate gamma distribution. Statistics & Probability Letters, **15**, 2353–2360 (2008)

[171] Galambos, J.: The Asymptotic Theory of Extreme Order Statistics. John Wiley & Sons, Inc., New York (1981).

[172] Ganuza, J.-J. and Penalva, J. S.: Signal orderings based on dispersion and the supply of private information in auctions. Econometrica, **78**, 1007–1030 (2010)

[173] Geluk, J. and de Haan, L.: Regular Variation, Extensions and Tauberian Theorems. CWI Tract 40, Center for Mathematics and Computer Science, Amsterdam, The Netherlands (1987)

[174] Genest, C., Kochar, S. C. and Xu, M.: On the range of heterogeneous samples. Journal of Multivariate Analysis, **100**, 1587–1592 (2009)

[175] Genest, C. and Nešlehovà, J.: A primer on copulas for count data. ASTIN Bulletin, **37** (2), 475–515 (2007)

[176] Genest, C., Nešlehová, J. and Ziegel, J.: Inference in multivariate Archimedean copula models. TEST, **20**, 223–256 (2011)

[177] Genest, C. and Rivest, L.-P.: A characterization of Gumbel's family of extreme value distributions. Statistics and Probability Letters, **8**, 207–211 (1989)

[178] Gerber, H. U.: An Introduction to Mathematical Risk Theory. S.S. Huebner Foundation Monographs, University of Pennsylvania (1979)

[179] Gertsbakh, I. and Shpungin, Y.: Multidimensional Spectra of Multistate Binary Systems. Chapter 4 in Recent Advances in System Reliability, Signatures, Multistate Systems and Statistical Inference. Lisnianski A. and Frenkel I. (Eds). Springer Series in Reliability Engineering. Springer (2011)

[180] Gertsbakh, I. and Shpungin, Y.: Network reliabilty and Resilience. Springer Briefs in Electrical and Computer Engineering, Springer (2011)

[181] Gertsbakh, I., Shpungin, Y. and Spizzichino, F.: On the signature of coherent systems obtained by connecting independent modules. Journal of Applied Probability, **48**, 843–855 (2011)

[182] Gertsbakh, I., Shpungin, Y. and Spizzichino, F.: On two-dimensional signatures. Journal of Applied Probability, **49**, 416–429 (2012)

[183] Ghosh, S. and Resnick, S. I.: A discussion on mean excess plots. Stochastic Processes and their Applications, **120** 1492–1517 (2010)

[184] Gnedenko, B. V.: Sur la distribution limite du terme maximum d'une série aléatoire. Annals of Mathematics, **44**, 423–453 (1943)

[185] Gupta, R. C. and Kirmani, S. N. U. A.: Closure and monotonicity properties of nonhomogeneous Poisson processes and record values. Probability in the Engineering and Informational Sciences, **2**, 475–484 (1988)

[186] Haines, A. L. and Singpurwalla, N. D.: Some contributions to the stochastic characterization of wear. Reliability and Biometry, Statistical Analysis of Lifelength (edited by F. Proschan and R.J. Serfling), SIAM, Philadelphia, 47–80 (1974)

[187] Hajek, B.: Optimal control of two interacting service stations. IEEE Transactions on Automatic Control, **29**, 491–499 (1984)

[188] Hashorva, E.: Asymptotic properties of type I elliptical random vectors. Extremes, **10**, 175–206 (2007)

[189] Hashorva, E.: Tail asymptotic results for elliptical distributions. Insurance: Mathematics and Economics, **43**, 158–164 (2008)

[190] Hashorva, E.: On the residual dependence index of elliptical distributions. Statistics & Probability Letters, **80**, 1070–1078 (2010)

[191] Heffernan, J. E.: A directory of coefficients of tail dependence. Extremes, **3**, 279–290 (2001)

[192] Hua, L.: Multivariate Extremal Dependence and Risk Measures. University of British Columbia, PhD Thesis (2012)

[193] Hua, L. and Joe, H.: Tail order and intermediate tail dependence of multivariate copulas. Journal of Multivariate Analysis, **102**, 1454–1471 (2011)

[194] Hua, L. and Joe, H.: Second order regular variation and conditional tail expectation of multiple risks. Insurance: Mathematics and Economics **49**, 537–546 (2011)

[195] Huang, X. and Pascual, F.: Monitoring Weibull Quantiles with Control Charts Based on the First Order Statistic. Journal of Statistical Computation and Simulation, **81**, 1677–1696 (2011)

[196] Huang, W. J. and Su, J. C.: On certain problems involving order statistics: a unified approach through order statistics property of point processes. Sankhyā: the Indian Journal of Statistics, **61**, 36–49 (1999)

[197] Hu, T., Chen, J. and Yao, J.: Preservation of the location independent risk order under convolution. Insurance: Mathematics and Economics, **38**, 406–412 (2006)

[198] Hu, T., Jin, W. and Khaledi, B.-E.: Ordering conditional distributions of generalized order statistics. Probability in the Engineering and Informational Sciences, **21**, 401–417 (2007)

[199] Hu, T. and Li, Y.: Increasing failure rate and decreasing reversed hazard rate properties of the minimum and the maximum of multivariate distributions with log-concave densities. Metrika, **65**, 325–330 (2007)

[200] Hu, T., Lu, Q. and Wen, S.: Some new results on likelihood ratio orderings for spacings of heterogeneous exponential random variables. Communications in Statistics: Theory and Methods, **37**, 2506–2515 (2008)

[201] Hu, T. and Wang, Y.: Optimal allocation of active redundancies in r-out-of-n system. Journal of Statistical Planning and Inference, **139**, 3733–3737 (2009)

[202] Hu, T., Wang, F. and Zhu, Z.: Stochastic comparisons and dependence of spacings from two samples of exponential random variables. Communications in Statistics, **35**, 979–988 (2006)

[203] Hu, T. and Wei, Y.: Stochastic comparisons of spacings from restricted families of distributions. Statistics and Probability Letters, **53**, 91–99 (2001)

[204] Hu, T. and Zhuang, W.: Stochastic properties of p-spacings of generalized order statistics. Probability in the Engineering and Informational Sciences, **19**, 257–276 (2005)

[205] Hu, T. and Zhuang, W.: A note on stochastic comparisons of generalized order statistics. Statistics and Probability Letters, **72**, 163–170 (2005)

[206] Jakeman, E. and Pusey, P. N.: Significance of K distributions in scattering experiments. Physical Review Letters, **40**, 546 (1978)

[207] Jaworski, P.: On uniform tail expansions of multivariate copulas and wide convergence of measures. Applicationes Mathematicae, **33**(2), 159–184 (2006)

[208] Jaynes, E. T.: Information theory and statistical mechanics. Statistical Physics, 1962 Brandeis Lectures in Theoretical Physics, Vol. 3. (Ford K. ed.), 181–218. Benjamin, New York (1963)

[209] Joe, H.: An ordering of dependence for contingency tables. Linear Algebra and its Applications, **70**, 89–103 (1985)

[210] Joe, H.: Majorization, randomness and dependence for multivariate distributions. Annals of Probability, **15**, 1217–1225 (1987)

[211] Joe, H.: Multivariate Models and Dependence Concepts. Chapman and Hall, London (1997)

[212] Joe, H. and Hu, T.: Multivariate distributions from mixtures of max-infinitely divisible distributions. Journal of Multivariate Analysis, **57**, 240–265 (1996)

[213] Joe, H., Li, H. and Nikoloulopoulos, A. K.: Tail dependence functions and vine copulas. Journal of Multivariate Analysis, **101**, 252–270 (2010)

[214] Joe, H. and Ma, C.: Multivariate survival functions with a min-stable property. Journal of Multivariate Analysis, **75**, 13–35 (2000)

[215] Joe, H. and Proschan, F.: Percentile residual life functions. Operations Research, **32**, 668–678 (1984)

[216] Joe, H. and Proschan, F.: Comparison of two life distributions on the basis of their percentile residual life functions. The Canadian Journal of Statistics, **12**, 91–97 (1984)

[217] Jogdeo, K.: On a probability bound of Marshall and Olkin. Annals of Statistics, **6**, 232–234 (1978)

[218] Johnson, N. L., Kotz, S. and Balakrishnan, N.: Continuous Univariate Distributions C Vol. 1. John Wiley & Sons, Inc., New York (1994)

[219] Kaas, R., Goovaerts, M., Dhaene, J. and Denuit, M.: Modern Actuarial Risk Theory. Kluwer Academic Publishers (2001)

[220] Kamae, T, Krengel, U. and O'brien, G. L.: Stochastic inequalities on partially ordered spaces. Annals of Probability, **5** 899–912 (1977)

[221] Kamps, U.: A concept of generalized order statistics. Journal of Statistical Planning and Inference, **48**, 1–23 (1995)

[222] Kamps, U.: A Concept of Generalized Order Statistics. B.G. Taubner, Stuttgart (1995)

[223] Karamata, J.: Sur une inégalité rélative aux fonctions convexes. Publications de l'Institut Mathématique of University of Belgrade, **1**, 145–148 (1932)

[224] Karlin, S.: Total positivity. Vol. I, Stanford University Press, Stanford, CA (1968)

[225] Kerridge, D. F.: Inaccuracy and Inference. Journal of the Royal Statistical Society, Series B, **23**, 184–194 (1961)

[226] Khaledi, B.-E.: Stochastic Comparisons of Order Statistics; Spacing and Concomitants of Order Statistics. PhD Thesis, Stat. Math. Unit, Indian Statistical Institute, New Delhi, India (2000)

[227] Khaledi, B.-E.: Some new results on stochastic comparisons between generalized order statistics. Journal of the Iranian Statistical Society, **4**, 35–49 (2005)

[228] Khaledi, B.-E., Amiripour, F., Hu, T. and Shojaei, R.: Some new results on stochastic comparisons of record values. Communications in Statistics: Theory and Methods, **38**, 2056–2066 (2009)

[229] Khaledi, B.-E. and Kochar, S. C.: Stochastic orderings between distributions and their sample spacings - II. Statistics and Probability Letters, **44**, 161–166 (1999)

[230] Khaledi, B.-E. and Kochar, S. C.: Dependence properties of multivariate mixture distributions and their applications Annals of the Institute of Statistical Mathematics, **53**, 620–630 (2001)

[231] Khaledi, B.-E. and Kochar, S. C.: Stochastic properties of spacings in a single-outlier exponential model. Probability in Engineering and Information Sciences, **15**, 401–408 (2001)

[232] Khaledi, B.-E. and Kochar, S. C.: Stochastic orderings among order statistics and sample spacings. Uncertainty and Optimality: probability, statistics and operations research, 167–203. World Scientific Publications, Singapore (2002)

[233] Khaledi, B.-E. and Kochar, S. C.: Ordering convolutions of gamma random variables. Sankhyā, **66**, 466–473 (2004)

[234] Khaledi, B.-E. and Kochar, S. C.: Dependence orderings of generalized order statistics. Statistics and Probability Letters, **73**, 357–367 (2005)

[235] Khaledi, B.-E. and Shaked, M.: Ordering conditional lifetimes of coherent systems. Journal of Statistical Planning and Inference, **137**, 1173–1184 (2007)

[236] Khaledi, B.-E. and Shojaei, R.: On stochastic orderings between residual record values. Statistics and Probability Letters, **77**, 1467–1472 (2007)

[237] Kijima, M., Li, H. and Shaked, M.: Stochastic Processes in Reliability, Handbook of Statistics, **18**, Eds. C. R. Rao and D. N. Shanbhag, 471–510 (1999)

[238] Kimeldorf, G. and Sampson, A. R.: Positive dependence orderings. Annals of the Institute of Statistical Mathematics, **39**, 113–128 (1987)

[239] Kitaev, M. Y. and Rykov, V. V.: Controlled queueing systems. CRC Press, New York (1995)

[240] Klefsjö, B.: TTT-plotting — a tool for both theoretical and practical problems. Journal of Statistical Planning and Inference, **29**, 111–124 (1991)

[241] Kleiber, C. and Kotz, S.: Statistical Size Distributions in Economics and Actuarial Sciences. Wiley-Interscience (2003)

[242] Klüppelberg, C.: Asymptotic ordering of distribution functions and convolution semigroups. Semigroup Forum, **40**(1), 77–92 (1990)

[243] Klüppelberg, C.: Asymptotic ordering of risks and ruin probabilities. Insurance: Mathematics and Economics, **12**, 259–264 (1993)

[244] Klüppelberg, C., Kuhn, G. and Peng, L.: Semi-parametric models for the multivariate tail dependence function – the asymptotically dependent. Scandinavian Journal of Statistics, **35**(4), 701–718 (2008)

[245] Klüppelberg, C. and Resnick, S. I.: The Pareto copula, aggregation of risks, and the emperor's socks. Journal of Applied Probability, **45**(1), 67–84 (2008)

[246] Kochar, S. C.: On extensions of DMRL and related partial orderings of life distributions. Communications in Statistics—Stochastic Models, **5**, 235–245 (1989)

[247] Kochar, S. C.: On stochastic ordering between distributions and their sample spacings. Statistics and Probability Letters, **42**, 345–352 (1999)

[248] Kochar, S. C. and Kirmani, S. N. U. A.: Some results on normalized spacings from restricted families of distributions. Journal of Statistical Planning and Inference, **46**, 47–57 (1995)

[249] Kochar, S. C. and Korwar, R.: Stochastic orders for spacings of heterogeneous exponential random variables. Journal of Multivariate Analysis, **57**, 69–83 (1996)

[250] Kochar, S. C., Li, X. and Shaked, M.: The total time on test transform and the excess wealth stochastic order of distributions. Advances in Applied Probability, **34**, 826–845 (2002)

[251] Kochar, S. C., Li, X. and Xu, M.: Excess wealth order and sample spacings. Statistical Methodology, **4**, 385–392 (2007)

[252] Kochar, S. C. and Ma, C.: Dispersive ordering of convolutions of exponential random variables. Statistics & Probability Letters, **43** 321–324 (1999)

[253] Kochar, S. C., Mukerjee, H. and Samaniego, F. J.: The "signature" of a coherent system and its application to comparison among systems. Naval Research Logistics, **46**, 507–523 (1999)

[254] Kochar, S. C., Mukerjee, H. and Samaniego, F. J.: Estimation of a monotone mean residual life. Annals of Statistics, **28**, 905–921 (2000)

[255] Kochar, S. C. and Rojo, J.: Some new results on stochastic comparisons of spacings from heterogeneous exponential distributions. Journal of Multivariate Analysis, **59**, 272–281 (1996)

[256] Kochar, S. C. and Xu, M.: Stochastic comparisons of parallel systems when components have proportional hazard rates. Probability in the Engineering and Informational Sciences, **21**, 597–609 (2007)

[257] Kochar, S. C. and Xu, M.: A new dependence ordering with applications. Journal of Multivariate Analysis, **99**, 2172–2184 (2008)

[258] Kochar, S. C. and Xu, M.: Comparisons of parallel systems according to the convex transform order. Journal of Applied Probability, **46**, 342–352 (2009)

[259] Kochar, S. C. and Xu, M.: On the right spread order of convolutions of heterogeneous exponential random variables. Journal of Multivariate Analysis, **101**, 165–176 (2010)

[260] Kochar, S. C. and Xu, M.: On residual lifetime of k-out-of-n systems with nonidentical components. Probability in the Engineering and Informational Sciences, **24**, 109–127 (2010)

[261] Kochar, S. C. and Xu, M.: Stochastic comparisons of spacings from heterogeneous samples. (Martin Wells and Ashis Sengupta edited) Festschrift Volume for Sreenivasa Rao Jammalamadaka, 113–129, Springer (2011)

[262] Kochar, S. C. and Xu, M.: On the skewness of order statistics in the multiple-outlier models. Journal of Applied Probability, **48**, 271–284 (2011)

[263] Kochar, S. C. and Xu, M.: The tail behavior of the convolutions of Gamma random variables. Journal of Statistical Planning and Inference, **141**, 418–428 (2011)

[264] Kochar, S. C. and Xu, M.: Some unified results on comparing linear combinations of independent gamma random variables. Probability in Engineering and Information Sciences, **26**, 393–404 (2012)

[265] Koole, G.: A simple proof of the optimality of a threshold policy in two-server queueing system. Systems & Control Letters, **26**, 301–303 (1995)

[266] Korwar, R. M.: On stochastic orders for sums of independent random variables. Journal of Multivatiate Analysis, **80**, 344–357 (2002)

[267] Krishna, V.: Auction theory. Academic Press, New York (2010)

[268] Krishnamoorthy, B.: On Poisson queue with two heterogeneous servers. Operational Research, **11**, 321–330 (1963)

[269] Lai, C. D. and Xie, M.: Stochastic Ageing and Dependence for Reliability. Springer, New York (2006)

[270] Langberg, N., Leon, R. V. and Proschan, F.: Characterization of nonparametric classes of life distributions. Annals of Probability, **8**, 1163–1170 (1980)

[271] Langrock, P. and Rykov V. V.: Methoden und Modelle zur Steurung von Bedienungssystemen. Handbuch der Bedienungs theorie. Berlin, Akademie-Verlag, **B. 2**, 422–486 (1984)

[272] Larsson, M. and Nešlehová, J.: Extremal behavior of Archimedean copulas. Advances in Applied Probability, **43**, 195–216 (2011)

[273] Launer, R. L.: Graphical techniques for analyzing failure data with the percentile residual-life function. IEEE Transactions on Reliability, **42**, 71–75, 80 (1993)

[274] Lawless, J. F.: Statistical Models and Methods for Lifetime Data (Wiley Series in Probability and Statistics), John Wiley & Sons, Inc., Hoboken, New Jersey (2002)

[275] Ledford, A. W. and Tawn, J. A.: Modelling dependence within joint tail regions. Journal of the Royal Statistical Society. Series B. Methodological, **59**, 475–499 (1997)

[276] Ledford, A. W. and Tawn, J. A.: Statistics for near independence in multivariate extreme values. Biometrika, **83**, 169–187 (1996)

[277] Lee, M.-L. T.: Dependence by total positivity. Annals of Probability, **13**, 572–582 (1985)

[278] Lehmann, E. L.: Some concepts of dependence. Annals of Mathematical Statistics, **37**, 1137–1153 (1966)

[279] Levitin, G., Gertsbakh, I. and Shpungin, Y.: Evaluating the damage associate with intentional network disintegration. Reliability Engineering and System Safety, **96**, 433–439 (2011)

[280] Li, H.: Orthant tail dependence of multivariate extreme value distributions. Journal of Multivariate Analysis **100**, 243–256 (2009)

[281] Li, H. and Sun, Y.: Tail dependence for heavy-tailed scale mixtures of multivariate distributions. Journal of Applied Probability **46**(4), 925–937 (2009)

[282] Lin, W. and Kumar, P. R.: Optimal control of a queueing system with two heterogeneous servers. IEEE Transactions on Automatic Control, **29**, 696–703 (1984)

[283] Lisnianski, A. and Levitin, G.: Multistate system reliability, **6** in Quality, Reliability and Engineering Statistics, World Scientific (2003)

[284] Li, W. V. and Shao, Q.-M.: Gaussian processes: Inequalities, small ball probabilities and applications. Stochastic Processes: Theory and Methods, edited by C.R. Rao and D. Shanbhag. Handbook of Statistics, Elsevier, New York, **19**, 533–598 (2001)

[285] Li, X.: A note on the expected rent in auction theory. Operations Research Letters, **33**, 531–534 (2005)

[286] Li, X. and Chen, J.: Aging properties of the residual life of k-out-of-n systems with independent but non-identical components. Applied Stochastic Models in Business and Industry, **20**, 143–153 (2004)

[287] Li, X. and Ding, W.: Optimal allocation of active redundancies to k-out-of-n systems with heterogenous components. Journal of Applied Probability, **47**, 254–263 (2010)

[288] Li, X. and Hu, X.: Some new stochastic comparisons for redundancy allocations in series and parallel systems. Statistics and Probability Letters, **78**, 3388–3394 (2008)

[289] Li, X. and Kochar, S. C.: Some new results of NBU(2) class of life distributions. Journal of Applied Probability, **38**, 242–237 (2001)

[290] Li, X., Li, Z. and Jing, B-Y.: Some results about the NBUC class of life distributions. Statistics and Probability Letters, **46**, 229–237 (2000)

[291] Li, X. and Lu, J.: Stochastic comparisons on residual life and inactivity time of series and parallel systems. Probability in Engineering and Informational Sciences, **17**, 267–275 (2003)

[292] Li, X. and Pellerey, F.: Generalized Marshall-Olkin distributions, and related bivariate aging properties. Journal of Multivariate Analysis, **102**, 1399–1409 (2011)

[293] Li, X., Wu, Y. and Zhang, Z.: On allocation of general standby redundancy to series and parallel systems. Communications in Statistics-Theory and Methods. (in press) (2012)

[294] Li, X., Yan, R. and Hu, X.: On the allocation of redundancies in series and parallel systems. Communications in Statistics-Theory and Methods, **40**, 959–968 (2011)

[295] Li, X., Zhang, Z. and Wu, Y.: Some new results involving general standby systems. Applied Stochastic Models in Business and Industry, **25**, 632–642 (2009)

[296] Li, X. and Zhao, P.: Some aging properties of the residual life of k-out-of-n systems. IEEE Transactions on Reliability, **55**, 535–541 (2006)

[297] Li, X. and Zhao, P.: Stochastic comparison on general inactivity time and general residual life of k-out-of-n systems. Communications in Statistics – Simulation and Computation, **37**, 1005–1019 (2008)

[298] Li, X. and Zuo, M. J.: On the behaviour of some new aging properties based upon the residual life of k-out-of-n systems. Journal of Applied Probability, **39**, 426–433 (2002)

[299] Lv, W., Mao, T. and Hu, T.: Properties of second-order regular variation and expansions for risk concentration. Probability in the Engineering and Informational Sciences, To appear (2012)

[300] Mainik, G. and Rüschendorf, L.: Ordering of multivariate risk models with respect to extreme portfolio losses. Statistics & Risk Modeling, **29**(1), 73–106 (2012)

[301] Malov, S. V.: On finite-dimensional Archimedean copulas. Asymptotic Methods in Probability and Statistics with Applications, 19–35 (2001)

[302] Mao, T. and Hu, T.: Equivalent characterizations on orderings of order statistics and sample ranges. Probability in the Engineering and Informational Sciences, **24**, 245–262 (2010)

[303] Marichal, J-L. and Mathonet, P.: Extensions of system signatures to dependent lifetimes: Explicit expressions and interpretations. Journal of Multivariate Analysis, **102**, 931–936 (2011)

[304] Marichal, J-L., Mathonet, P. and Waldhauser, T.: On signature-based expressions of system reliability. Journal of Multivariate Analysis, **102**, 1410–1416 (2011)

[305] Marinacci, M. and Montrucchio, L.: Ultramodular functions. Mathematics of Operations Research, **30**, 311–332 (2005)

[306] Marshall, A. W.: Multivariate stochastic orderings and generating cones of functions. In Mosler, K. and Scarsini, M. (eds.), Stochastic orders and decision under risk (Hamburg, 1989). IMS Lecture Notes Vol. 19, (pp. 231–247), Hayward, CA (1991)

[307] Marshall, A. W. and Olkin, I.: A multivariate exponential distribution. Journal of the American Statistical Association, **2**, 84–98 (1967)

[308] Marshall, A. W. and Olkin, I.: Inequalities: theory of majorization and its applications. Academic Press, Inc., New York (1979)

[309] Marshall, A. W. and Olkin, I.: Domains of attraction of multivariate extreme value distributions. Annuals of Probability, **11**, 168–177 (1983)

[310] Marshall, A. W. and Olkin, I.: Families of multivariate distributions. Journal of the American Statistical Association, **83**, 834–841 (1988)

[311] Marshall, A. W. and Olkin, I.: Life Distributions. Springer, New York (2007)

[312] Marshall, A. W., Olkin, I. and Arnold, B. C.: Inequalities: Theory of Majorization and Its Applications. Springer, New York (2011)

[313] Maynar, P. and Trizac, E.: Entropy of continuous mixtures and the measure problem. Physical Review Letters, **106**, 160603 [4 pages] (2011)

[314] McNeil, A., Frey, R. and Embretchs, P.: Quantitative Risk Management. Princeton Press, Princeton (2005)

[315] McNeil, A. J. and Nešlehová, J.: Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions. Annals of Statistics, **37**, 3059–3097 (2009)

[316] McNeil, A. J. and Nešlehová, J.: From Archimedean to Liouville copulas. Journal of Multivariate Analysis, **101**, 1772–1790 (2010)

[317] Meeker, W. Q. and Escobar, L.: Statistical Methods for Reliability Data (Wiley Series in Probability and Statistics), John Wiley & Sons, Inc., New York (1998)

[318] Meyer, M. and Strulovici, B.: The supermodular stochastic ordering. unpublished manuscript (2011)

[319] Mi, J.: Optimal active redundancy allocation in k-out-of-n system. Journal of Applied Probability, **36**, 927–933 (1999)

[320] Mi, J., Shi, W. and Zhou, Y. Y.: Some properties of convolutions of Pascal and Erlang random variables. Statistics & Probability Letters, **78**, 2378–2387 (2008)

[321] Misagh, F., Panahi, Y., Yari, G. H. and Shahi, R.: Weighted cumulative entropy and its estimation. IEEE International Conference on Quality and Reliability, 477–480 (2011)

[322] Misra, N., Dhariyal, I. and Gupta, N.: Optimal allocation of active spares in series systems and comparison of component and system redundancies. Journal of Applied Probability, **46**, 19–34 (2009)

[323] Misra, N., Gupta, N. and Dhariyal, I. D.: Stochastic properties of residual life and inactivity time at a random time. Stochastic Models, **24**, 89–102 (2008)

[324] Misra, N., Misra, A. K. and Dhariyal, I.: Active redundancy allocations in series systems. Probability in the Engineering and Informational Sciences, **25**, 219–235 (2011)

[325] Misra, N., Misra, A. K. and Dhariyal, I.: Standby redundancies allocation in series and parallel systems. Journal of Applied Probability, **48**, 1–13 (2011)

[326] Misra, N. and van der Meulen, E. C.: On stochastic properties of m-spacings. Journal of Statistical Planning and Inference, **115**, 683–697 (2003)

[327] Mitrinovic, D. S.: Analytic Inequalities. Springer-Verlag, Berlin (1970)

[328] Mizuno, T.: A relation between positive dependence of signal and the variability of conditional expectation given signal. Journal of Applied Probability, **43**, 1181–1185 (2006)

[329] Mosler, K. C.: Entscheidungsregeln bei Risiko Multivariate stochastische Dominanz. Springer-Verlag, New York (1982)

[330] Mosler, K. and Scarsini, M.: Stochastic Orders and Decision under Risk. Lecture Notes-Manograph Series, IMS, Hayward, CA (1991)

[331] Müller, A.: Stochastic orders generated by integrals: a unified study. Advances in Applied Probability, **29**, 414–428 (1997)

[332] Müller, A.: On the waiting times in queues with dependency between interarrival and service times. Operations Research Letters, **26**, 43–47 (2000)

[333] Müller, A. and Scarsini, M.: Archimedean copulae and positive dependence. Journal of Multivariate Analysis, **93**, 434–445 (2005)

[334] Müller, A. and Scarsini, M.: Fear of loss, inframodularity, and transfers. Journal of Economic Theory, **147**, 1490–1500 (2012)

[335] Müller, A. and Stoyan, D.: Comparison Methods for Stochastic Models and Risks. Wiley, Chichester (2002)

[336] Muth, E. J.: Reliability models with positive memory derived from the mean residual life function. Theory and Applications of Reliability (edited by C.P. Tsokos and I.N. Shimi), Academic Press, New York (1977)

[337] Nair, N. U. and Sankaran, P. G.: Properties of a mean residual life function arising from renewal theory. Naval Research Logistics, **57**, 373–379 (2010)

[338] Navarro, J.: Likelihood ratio ordering of order statistics, mixtures and systems. Journal of Statistical Planning and Inference, **138**, 1242–1257 (2008)

[339] Navarro, J., Balakrishnan, N. and Samaniego, F. J.: Mixture representations of residual lifetimes of used systems. Journal of Applied Probability, **45**, 1097–1112 (2008)

[340] Navarro, J., del Aguila, Y. and Asad M.: Some new results on the cumulative residual entropy. Journal of Statistical Planning and Inference, **140**, 310–322 (2010)

[341] Navarro, J. and Eryilmaz, S.: Mean residual lifetimes of consecutive k-out-of-n systems. Journal of Applied Probability, **44**, 82–98 (2007)

[342] Navarro, J. and Hernandez P. J.: Mean residual life functions of finite mixtures, order statistics and coherent systems. Metrika, **67**, 277–298 (2008)

[343] Navarro, J., Ruiz, J. M. and Sandoval, C. J.: A note on comparisons among coherent systems with dependent components using signatures. Statistics and Probability Letters, **72**, 179–185 (2005)

[344] Navarro, J., Ruiz, J. M. and Sandoval, C. J.: Properties of coherent systems with dependent components. Communications in Statistics — Theory and Methods, **36**, 175–191 (2007)

[345] Navarro, J. and Rychlik, T.: Reliability and expectation bounds for coherent systems with exchangeable components. Journal of Multivariate Analysis, **98**, 102–113 (2007)

[346] Navarro, J. and Rychlik, T.: Comparisons and bounds for expected lifetimes of reliability systems. European Journal of Operational Research, **207**, 309–317 (2010)

[347] Navarro, J., Rychlik, T. and Shaked, M.: Are the order statistics ordered? A survey of recent results. Communications in Statistics - Theory and Methods, **36**, 1273C1290 (2007)

[348] Navarro, J., Samaniego, F. J. and Balakrishnan, N.: The joint signature of coherent systems with shared components. Journal of Applied Probability, **47**, 235–253 (2010)

[349] Navarro, J., Samaniego, F. J. and Balakrishnan, N.: Signature-based representations for the reliability of systems with heterogeneous components. Journal of Applied Probability, **48**, 856–867 (2011)

[350] Navarro, J., Samaniego, F. J., Balakrishnan, N. and Bhattacharya, D.: On the application and extension of system signatures in engineering reliability. Naval Research Logistics, **55**, 313–327 (2008)

[351] Navarro, J. and Shaked M.: Some properties of the minimum and the maximum of random variables with joint logconcave distributions. Metrika, **71**, 313–317 (2010)

[352] Navarro, J. and Shaked, M.: Hazard rate ordering of order statistics and systems. Journal of Applied Probability, **43**, 391–408 (2006)

[353] Nelson, L. S.: A Control Chart for Parts-Per-Million Nonconforming Bibitems. Journal of Quality Technology, **26**, 239–240 (1994)

[354] Nelson, P. R.: Control Charts for Weibull Processes with Standards Given. IEEE Transactions on Reliability, **28**, 283–287 (1979)

[355] Nelsen, R. B.: An Introduction to Copulas (2nd Edition). Springer (2006)

[356] Nelson, W.: Applied Life Data Analysis (Wiley Series in Probability and Statistics). John Wiley & Sons, Inc., Hoboken, New Jersey (2003)

[357] Nelson, W.: Accelerated Testing: Statistical Models, Test Plans, and Data Analysis (Wiley Series in Probability and Statistics). John Wiley & Sons, Inc., Hoboken, New Jersey (2004)

[358] Neves, C.: From extended regular variation to regular variation with application in extreme value statistics. Journal of Mathematical Analysis and Applications, **355**, 216–230 (2009)

[359] Nichols, M. D. and Padgett, W. J.: A Bootstrap Control Chart for Weibull Percentiles.: Quality and Reliability Engineering International, **22**, 141–151 (2006)

[360] Nikoloulopoulos, A. K., Joe, H. and Li, H.: Extreme value properties of multivariate t copulas. Extremes, **12**, 129–148 (2009)

[361] Norros, I.: System weakened by failures. Stochastic Processes and their Applications, **20**, 181–196 (1985)

[362] Oakes, D.: Bivariate survival models induced by frailties. Journal of the American Statistical Association, 487–493 (1989)

[363] Padgett, W. J. and Spurrier, J. D.: Shewhart-type charts for percentiles of strength distributions.: Journal of Quality Technology, **22**, 283–288 (1990)

[364] Page, E. S.: Continuous Inspection Schemes. Biometrika, **41**, 100–115 (1954)

[365] Păltănea, E.: On the comparison in hazard rate ordering of fail-safe systems. Journal of Statistical Planning and Inference, **138**, 1993–1997 (2008)

[366] Pascual, F.: EWMA Charts for the Weibull Shape Parameter. Journal of Quality Technology, **42**, 400–416 (2010)

[367] Pascual, F. and Nguyen, D.: Moving Range Charts for Monitoring the Weibull Shape Parameter with Single-Observation Samples. Quality and Reliability Engineering International, **27**, 905–919 (2011)

[368] Pascual, F. and Zhang, H.: Monitoring the Weibull Shape Parameter by Control Charts for the Sample Range. Quality and Reliability Engineering International, **27**, 15–25 (2011)

[369] Pedro, T.: One size does not fit all: a case for heterogeneous multiprocessor systems. IADIS International Conference Applied Computing, 15–22 (2005)

[370] Pellerey, F.: On univariate and bivariate aging for dependent lifetimes with Archimedean survival copulas. Kybernetica, **44**, 795–806 (2008)

[371] Pellerey, F. and Petakos, K.: On closure property of the NBUC class under formation of parallel systems. IEEE Transaction on Reliability, **51**, 452–454 (2002)

[372] Pellerey, F., Shaked, M. and Zinn, J.: Nonhomogeneous Poisson processes and logconcavity. Probability in the Engineering and Informational Sciences, **14**, 353–373 (2000)

[373] Pfeifer, D.: 'Record values' in einem stochastischen Modell mit nicht-identischen Verteilungen. Dissertation, Aachen University of Technology (1979)

[374] Pickands, J.: Multivariate negative exponential and extreme value distributions. Unpublished manuscript, University of Pennsylvania (1980)

[375] Pickands, J.: Multivariate extreme value distributions. 43rd Session International Statistical Institute (1981)

[376] Pignatiello, J. J., Acosta-Mejia, C. A. and Rao, B. V.: The Performance of Control charts for Monitoring Process Dispersion.: Proceedings of the 4th Industrial Engineering Research Conference, 320–328 (1995)

[377] Pledger, G. and Proschan, F.: Comparisons of order satistics from heterogeneous populations, with applications in reliability. Optimizing Methods in Statistics, 89–113. Academic Press, New York (1971)

[378] Press, W. H., Tuekolsky, S. A., Vetterling, W. T. and Flannery, B. P.: Numerical Recipes in FORTRAN 77: The Art of Scientific Computing. Cambridge University Press, New York (2001)

[379] Qiu, G. and Wu, J.: Some comparisons between generalized order statistics. Applied Mathematics: A Journal of Chinese Universities, Series B, **22**, 325–333 (2007)

[380] R Development Core Team. R Project, accessed December 5, 2012, http://www.r-project.org (2012)

[381] Ramalhoto, M. F. and Morais, M.: EWMA control charts for the scale parameter of a Weibull control variable with fixed and variable sampling intervals. Economic Quality Control, **13**, 23–46 (1998)

[382] Ramalhoto, M. F. and Morais, M.: Shewhart Control Charts for the Scale Parameter of a Weibull Control Variable with Fixed and Variable Sampling Intervals. Journal of Applied Statistics, **26**, 129–160 (1999)

[383] Ramos, A. and Ledford, A.: A new class of models for bivariate joint tails. Journal of the Royal Statistical Society: Series B (Statistical Methodology), **71**, 219–241 (2009)

[384] Rao, M.: More on a new concept of entropy and information. Journal of Theoretical Probability, **18**, 967–981 (2005)

[385] Rao, M., Chen, Y., Vemuri, B. C. and Wang, F.: Cumulative residual entropy: a new measure of information. IEEE Transactions on Information Theory, **50**, 1220–1228 (2004)

[386] Resnick, S. I.: Extreme Values, Regular Variation, and Point Processes. Springer-Verlag, New York (1987)

[387] Resnick, S. I.: Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer, New York (2007)

[388] Rinne, H.: The Weibull Distribution: A Handbook, Chapman and Hall/CRC, London (2008)

[389] Rojo, J.: A pure tail ordering based on the ratio of the quantile functions. Annals of Statistics, **20**(1), 570–579 (1992)

[390] Rojo, J.: On the preservation of some pure-tail orderings by reliability operations. Statistics and Probability Letters, **17**(5), 189–198 (1993)

[391] Rojo, J.: Relationships between pure tail orderings of lifetime distributions and some concepts of residual life. Annals of the Institute of Statistical Mathematics, **48**(2), 247–255 (1996)

[392] Rolski, T., Schmidli, H., Schmidt, V. and Teugels, J.: Stochastic Processes for Insurance and Finance. Wiley, New York (1999)

[393] Ross, S.: Applied probability models with optimization applications. Holden Day, San Francisco (1970)

[394] Rothschild, M. and Stiglitz, J. E.: Increasing risk. I. A definition. Journal of Economic Theory, **2**, 225–243 (1970)

[395] Rothschild, M. and Stiglitz, J. E.: Increasing risk. II. Its economic consequences. Journal of Economic Theory, **3**, 66–84 (1971)

[396] Rothschild, M. and Stiglitz, J. E.: Addendum to: "Increasing risk. I. A definition". Journal of Economic Theory, **5**, 306 (1972)

[397] Rüschendorf, L. and de Valk, V.: On regression representations of stochastic processes. Stochastic Processes and their Applications, **46**, 183–198 (1993)

[398] Rüschendorf, L., Schweizer, B. and Taylor, M. D.: Distribution with Fixed Marginals and Related Topics. Lecture Notes-Manograph Series, IMS, Hayward, CA (1996)

[399] Rykov, V. V.: Markov decision processes with finite state and decision spaces. Theory of Probability and its Applications, **11**, no. 2, 302–311 (1975)

[400] Rykov, V. V.: Controllable queueing systems. Itogi nauki i techniki. Teoria verojatn. Matem. Statist. Teoretich. Kibern., **12**, 45–152 (1975) (In Russian. There is English translation in Journal of Soviet Mathematics)

[401] Rykov, V. V.: Multi-server controllable queueing system with heterogeneous servers (In Russian). Vestnik of the Russian Peoples' Friendship University. Series Applied Mathematics and Information, **1**, 83–92 (1994)

[402] Rykov, V. V.: On monotonicity conditions of optimal control policies for controllable Queueing Systems. Automation and Remote Control, **9**, 92–106 (1999)

[403] Rykov, V. V.: Monotone Control of Queueing System with heterogeneous servers. Queueing Systems, **37**, 391–403 (2001)

[404] Rykov, V. V. and Efrosinin, D. V.: Numerical study of optimal control of queueing system with heterogeneous servers. Automation and Remote Control, **64**(2), 302–309 (2003)

[405] Rykov, V. V. and Efrosinin, D. V.: On slow Server Problem. Automation and Remote Control, **70**(12), 2013–2023 (2009)

[406] Sadegh, M. K.: Mean past and mean residual life functions of a parallel system with nonidentical components. Communications in Statistics – Theory and Methods, **37**, 1134–1145 (2008)

[407] Samaniego, F. J.: On closure of the IFR class under formation of coherent systems. IEEE Transactions on Reliability, **34**, 69–72 (1985)

[408] Samaniego, F. J.: System Signatures and Their Applications in Engineering Reliability. New York: Springer (2007)

[409] Samaniego, F. J., Balakrishnan, N. and Navarro, J.: Dynamic signatures and their use in comparing the reliability of new and used systems. Naval Research Logistics, **56**, 577–591 (2009)

[410] Sampson, A. R. and Whitaker, L. R.: Estimation of multivariate distributions under stochastic ordering. Journal of the American Statistical Association, **84**, 541–548 (1989)

[411] Scarsini, M.: An ordering of dependence. Topics in Statistical Dependence (edited by Block, H. W., Sampson, A. R. and Savits, T. H.), IMS Lecture Notes-Monograph Series 16, Hayward, CA, 403–414 (1990)

[412] Schmeidler, D.: Subjective probability and expected utility without additivity. Econometrica, **57**, 571–587 (1989)

[413] Schmidt, R.: Tail dependence for elliptically contoured distributions. Mathematical Methods of Operations Research, **55**, 301–327 (2002)

[414] Schroeder, M. J.: An alternative to entropy in the measurement of information. Entropy, **6**, 388–412 (2004)

[415] Schweizer, B. and Sklar, A.: Probabilistic Metric Spaces. North-Holland Series in Probability and Applied Mathematics. North-Holland Publishing Co., New York (1983)

[416] Sennott, L. I.: Stochastic Dynamic Programming and the Control of Queueing Systems. Wiley, New York (1999)

[417] Serfling, R. J.: Approximation Theorems of Mathematical Statistics. Wiley, New York (1980)

[418] Shaked, M. and Shanthikumar, J. G.: Multivariate imperfect repair. Operations Research, **34**, 437–448 (1986)

[419] Shaked, M. and Shanthikumar, J. G.: IFRA properties of some Markov jump processes with general state space. Mathematics of Operations Research, **12**, 562–568 (1987)

[420] Shaked, M. and Shanthikumar, J. G.: Dynamic multivariate aging notions in reliability theory. Stochastic Processes and their Applications, **38**, 85–97 (1991)

[421] Shaked, M. and Shanthikumar, J. G.: Optimal allocation of resources to nodes of series and parallel systems. Advances in Applied Probability, **24**, 894–914. (1992)

[422] Shaked, M. and Shanthikumar, J. G.: Stochastic Orders and Their Applications. Academic Press, San Diego (1994)

[423] Shaked, M. and Shanthikumar, J. G.: Hazard rate ordering of k-out-of-n systems. *Statistics and Probability Letters*, **23**, 1–8 (1995)

[424] Shaked, M. and Shanthikumar, J. G.: Supermodular stochastic orders and positive dependence of random vectors. Journal of Multivariate Analysis, **61**, 86–101 (1997)

[425] Shaked, M. and Shanthikumar, J. G.: Two variability orders. Probability in the Engineering and Informational Sciences, **12**, 1–23 (1998)

[426] Shaked, M. and Shanthikumar, J. G.: Stochastic Orders. Springer, New York (2007)

[427] Shaked, M. and Shanthikumar, J. G.: Stochastic orders and their applications (Second Edition). Springer Series in Statistics. Springer-Verlag, New York (2007)

[428] Shaked, M., Sordo, M. A. and Suárez-Llorens, A.: Global dependence stochastic orders. Methodology and Computing in Applied Probability, **14**, 617–648 (2012)

[429] Shaked, M. and Spizzichino, F.: Positive dependence properties of conditionally independent random lifetimes. Mathematics of Operations Research, **23**, 944–959 (1998)

[430] Shaked, M. and Szekli, R.: Comparison of replacement policies via point processes. Advances in Applied Probability, **27**, 1079–1103 (1995)

[431] Shaked, M. and Tong, Y. L.: Stochastic Inequalities. Lecture Notes-Manograph Series, IMS, Hayward, CA (1992)

[432] Shannon, C. E.: A mathematical theory of communication. Bell System Technical Journal, **27**, 279–423 (1948)

[433] Silvey, S. D.: On a measure of association. Annals of Mathematical Statistics, **35**, 1157–1166. (1964)

[434] Singh, H. and Misra, N.: On redundancy allocation in systems. Journal of Applied Probability **31**, 1004–1014 (1994).

[435] Singh, H. and Singh, R. S.: On allocation of spares at component level vs system level. Journal of Applied Probability, **34**, 283–287 (1997)

[436] Singh, H. and Singh, R. S.: Optimal allocation of resourses to nodes of series systems with respect to failure-rate ordering. Naval Research Logistics, **44**, 147–152 (1997)

[437] Sklar, A.: Fonctions de répartition à n dimensions et leurs marges. Publications de l'Institut de Statistique de l'Université de Paris, **8**, 229–231 (1959)

[438] Song, D. and Gupta, A. K.: Lp-norm uniform distribution. Proceedings of the American Mathematical Society, **125**, 595–602 (1997)

[439] Sordo, M. A.: Comparing tail variabilities of risks by means of the excess wealth order. Insurance: Mathematics and Economics, **45**, 466–469 (2009)

[440] Spizzichino, F.: Subjective Probability Models for Lifetimes. Monographs on Statistics and Applied Probability, vol. 91, Chapman & Hall/CRC, Boca Raton, FL (2001)

[441] Stidham, S. Jr.: Optimal control of admission to a queueing system. IEEE Transactions on Automatic Control, **30**, 705–713 (1985)

[442] Stidham, S. Jr. and Weber, R.: A survey of Markov decision models for control of networks of queues. Queueing Systems, **13**, 291–314 (1993)

[443] Stoyan, D.: Comparison methods for queues and other stochastic models. John Wiley & Sons, Berlin (1983)

[444] Strassen, V.: The existance of probability measures with given marginals. Annals of Mathematical Statistics, **36**, 423–439 (1965)

[445] Sukhatme, P. V.: Tests of significance for samples of the χ^2 population with two degrees of freedom. Annals of Eugenics, **8**, 52–56 (1937)

[446] Szekli, R.: Stochastic ordering and dependence in applied probability. Lecture Notes in Statistics (Springer). 97. Berlin: Springer-Verlag. viii (1995)

[447] Tawn, J. A.: Modelling multivariate extreme value distributions. Biometrika, **77**, 245–253 (1990)

[448] Tchen, A. H.: Inequalities for distributions with given marginals. Annals of Probability, **8**, 814–827 (1980)

[449] Tong, Y. L.: Probability Inequalities in Multivariate Distributions. Academic Press, New York (1980)

[450] Topkis, D.: Minimizing a submodular function on a lattice. Operations Research, **26** (2), 305–321 (1978)

[451] Torrado, N. and Lillo, R. E.: Comparisons among spacings from two populations. UC3M Working papers. Statistics and Econometrics 12-07(04), 1–18 (2012)

[452] Torrado, N. and Lillo, R. E.: Likelihood ratio order of spacings from two heterogeneous samples. Journal of Multivariate Analysis, **114**, 338–348 (2013)

[453] Torrado, N., Lillo, R. E. and Wiper, M. P.: On the conjecture of Kochar and Kowar. Journal of Multivariate Analysis, **101**, 1274–1283 (2010)

[454] Trivedi, P. K. and Zimmer, D. M.: Copula modeling: An introduction for practitioners. Foundations and Trends in Econometrics, **1**, 1–111 (2005)

[455] Trufin J., Albrecher, H. and Denuit., M.: Properties of a risk measure derived from ruin theory. The Geneva Risk and Insurance Review. Advance online publication, August 24, 2010; doi:10.1057/grir.2010.10.

[456] Unnikrishnan Nair, N. and Sankaran, P. G.: Properties of a mean residual life function arising from renewal theory. Naval Research Logistics, **57**, 373–379 (2010)

[457] Valdès, J. E., Arango, G., Zequeira, R. I. and Brito, G.: Some stochastic comparisons in series system with active redundancies. *Statistics and Probability Letters*, **80**, 945–949 (2010)

[458] Valdès, J. E. and Zequeira, R. I.: On the optimal allocation of an active redundancy in a two-component series system. Statistics and Probability Letters, **63**, 235–332 (2003)

[459] Valdès, J. E. and Zequeira, R. I.: On the optimal allocation of two active redundancies in a two-component series system. Operations Research Letters, **34**, 49–52 (2006)

[460] Vardeman, S. B. and Jobe, J. M.: Statistical Quality Assurance Methods for Engineers. John Wiley & Sons, Inc., New York (1999)

[461] Vaughan, R. J. and Venables, W. N.: Permanent expressions for order statistics densities. Journal of the Royal Statistical Society: Series B, **34**, 308–310 (1972)

[462] Vericourt, F. and Zhou, Y. P.: On the incomplete results for the heterogeneous server problem. Queueing Systems, **52**, 189–191 (2006)

[463] Vishnevsky, V. M. and Semenova, O. V.: Polling systems: theory and applications in bandwidth wireless networks (in Russian). M.: Tekhnosppere (2007)

[464] Vogt, V. H.: Zur Parameter- und Prozentpunktschätzung von Lebensdauerverteilungen bei kleinem Stichprobenumfang. Metrika, **14**, 117–131 (1969)

[465] Wang, S.: Insurance Pricing and Increased Limits Ratemaking by Proportional Hazards Transforms. Insurance: Mathematics and Economics, **17**, 43–54 (1995)

[466] Wang, S.: Ordering of Risks Under PH-Transforms. Insurance: Mathematics and Economics, **18**, 109–114 (1996)

[467] Wang, F. and Vemuri, B. C.: Non-rigid multi-modal image registration using cross-cumulative residual entropy. International Journal of Computer Vision, **74**, 201–215 (2007)

[468] Wang, F., Vemuri, B. C., Rao, M. and Chen, Y.: A new & robust information theoretic measure and its application to image alignment. Information Processing in Medical Imaging, Lecture Notes in Computer Science, **2732**, 388–400. Springer-Verlag, Heidelberg (2003)

[469] Wang, F., Vemuri, B. C., Rao, M. and Chen, Y.: Cumulative residual entropy, a new measure of information & its application to image alignment. Proceedings on the Ninth IEEE International Conference on Computer Vision (ICCV'03), **1**, 548–553. IEEE Computer Society (2003)

[470] Watson, G. S. and Wells, W. T.: On the possibility of improving the mean useful life of items by eliminating those with short lives. Technometrics, **3**, 281–298 (1961)

[471] Weber, R.: On a conjecture about assigning jobs to processors of different speeds. IEEE Transactions on Automatic Control, **38**, no. 1, 166–170 (1993)

[472] Weibull, W.: A Statistical Theory of the Strength of Material. Report No. 151, Ingeniors Vetenskaps Akademiens Handligar, Stockholm (1939)

[473] Weibull, W.: The Phenomenon of Rupture in Solids. Report No. 153, Ingeniors Vetenskaps Akademiens Handligar, Stockholm (1939)

[474] Wen, S., Lu, Q. and Hu, T.: Likelihood ratio orderings of spacings heterogeneous exponential random variables. Journal of Multivariate Analysis, **98**, 743–756 (2007)

[475] Williamson, R. E.: Multiply monotone functions and their Laplace transforms. Duke Mathematical Journal, **23**, 189–207 (1956)

[476] Wu, W. B. and Mielniczuk, J.: A new look at measuring dependence. Dependence in Probability and Statistics (Edited by Doukhan, P., Lang, G., Surgailis, D., and Teyssiere, G.). Springer, 123–142 (2010)

[477] Wu, X. and Yuen, K. C.: A discrete time risk model with interaction between classes of business. Insurance: Mathematics and Economics **33**, 117–133 (2003)

[478] Xie, H. and Hu, T.: Conditional ordering of generalized order statistics. Probability in the Engineering and Informational Sciences, **22**, 333–346 (2008)

[479] Xie, H. and Hu, T.: Ordering p-spacings of generalized order statistics revisited. Probability in the Engineering and Informational Sciences, **23**, 1–16 (2009)

[480] Xie, H. and Hu, T.: Some new results on multivariate dispersive ordering of generalized order statistics. Journal of Multivariate Analysis, **101**, 964–970 (2010)

[481] Xu, M., Li, X., Zhao, P. and Li, Z.: Likelihood ratio order of m-spacings in multiple-outlier models. Communications in Statistics: Theory and Methods, **36**, 1507–1525 (2007)

[482] Yanagimoto, T. and Okamoto, M.: Partial orderings of permutations and monotonicity of a rank correlation statistic. Annals of the Institute of Statistical Mathematics, **21**, 489–506 (1969)

[483] Yang, Z. and Xie, M.: Process monitoring of exponentially distributed characteristics through an optimal normalizing transformation. Journal of Applied Statistics, **27**, 1051–1063 (2000)

[484] Yu, Y.: Stochastic ordering of exponential family distributions and their mixtures. Journal of Applied Probability, **46**, 244–254 (2009)

[485] Zhang, L. and Chen, G.: EWMA Charts for Monitoring the Mean of Censored Weibull Lifetimes.: Journal of Quality Technology, **36**, 321–328 (2004)

[486] Zhang, Z.: Ordering conditional general coherent systems with exchangeable components. Journal of Statistical Planning and Inference, **140**, 454–460 (2010)

[487] Zhao, P.: Some new results on convolutions of heterogeneous gamma random variables. Journal of Multivariate Analysis, **102**, 958–976 (2011)

[488] Zhao, P. and Balakrishnan, N.: Mean residual life order of convolutions of heterogenous exponential random variables. Journal of Multivariate Analysis, **100**, 1792–1801 (2009)

[489] Zhao, P. and Balakrishnan, N.: Stochastic comparisons of conditional generalized order statistics. Journal of Statistical Planning and Inference, **139**, 2920–2932 (2009)

[490] Zhao, P. and Li, X.: Stochastic order of sample range from heterogeneous exponential random variables. Probability in the Engineering and Informational Sciences, **23**, 17–29 (2009)

[491] Zhao, P., Li, X. and Balakrishnan, N.: Conditional ordering of k-out-of-n systems with independent but nonidentical components. Journal of Applied Probability, **45**, 1113–1125 (2008)

[492] Zhao, P., Li, X. and Da, G.: Right spread order of the second order statistics from heterogeneous exponential distributions. Communications in Statistics - Theory and Methods, **47**, 3070–3081 (2011)

[493] Zhuang, W. and Hu, T.: Multivariate stochastic comparisons of sequential order statistics. Probability in the Engineering and Informational Sciences, **21**, 47–66 (2007)

[494] Zhuang, W. and Hu, T.: Multivariate dispersive ordering of spacings of generalized order statistics. Applied Mathematics Letters, **22**, 968–974 (2009)

[495] Zhuang, W., Yao, J. and Hu, T.: Conditional ordering of order statistics. Journal of Multivariate Analysis, **101**, 640–644 (2010)

Author Index

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9,

Subject Index

H. Li and X. Li (eds.), *Stochastic Orders in Reliability and Risk*, Lecture Notes in Statistics 208, DOI 10.1007/978-1-4614-6892-9,