Advances in Statistics, Combinatorics and Related Areas

This page is intentionally left blank

Selected Papers from the SCRA2001-FIM VIII, Wollogong Conference

Advances in Statistics, Combinatorics and Related Areas

University of Wollongong, Australia, 19–21 December 2001

Editors

Chandra Gulati
University of Wollongong, Australia

Yan-Xia Lin
University of Wollongong, Australia

Satya Mishra
University of South Alabama, USA

John Rayner
University of Wollongong, Australia

World Scientific
New Jersey • London • Singapore • Hong Kong

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: Suite 202, 1060 Main Street, River Edge, NJ 07661

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

ADVANCES IN STATISTICS, COMBINATORICS & RELATED AREAS

ISBN 981-238-201-1

Printed in Singapore by Mainland Press

In memory of my father Badri Nath and to my mother Savitri Devi, and Lilian, Danny and Jason: (CMG).

In memory of my father Zhensheng Lin and to my mother Wenying Yu, my husband Yi Mu and son John C. Mu: (YXL).

In memory of my parents:
Pandit Raghubeer Mishra and Srimati Kishori Mishra: (SNM).

To the Rayners: Fred, Carol, Glen, and Eric: (JCWR).

PREFACE

The School of Mathematics and Applied Statistics at the University of Wollongong played host to the International Conference on Statistics, Combinatorics and Related Areas and the Eighth International Conference of the Forum for the Interdisciplinary Mathematics (SCRA2001-FIMVIII). This conference took place from December 19-21, 2001 and it was cosponsored by the International Association of Statistical Computing and the University of South Alabama.

The Forum for Interdisciplinary Mathematics, FIM is an international society of scholars working in mathematical, statistical, information and system sciences. By the means of organizing international conferences, FIM is attempting to effectively integrate all quantitative areas of modern times.

The Forum for Interdisciplinary Mathematics (FIM) was formed in 1975 by a group of academicians at the University of Delhi under the leadership of Bhu Dev Sharma. Despite his leadership role as the current president of the Hindu University of America, Bhu Dev continues to provide direction in the growth of FIM. FIM leaders over the years include J. K. Ghosh, D. D. Joshi, C. R. Rao, M. M. Rao, D. K. Ray-Choudhuri, S.S. Shrikhande, N. M. Singhi and J. N. Srivastava. Satya Mishra is the current president of FIM and has been influential in organizing FIM's international conferences since 1997.

The following scholars served on the International Organization Committee (IOC) for this conference: Michel Deza, E.J. Dudewicz, Prem Goel, Michael Greenacre, C.C Heyde, Sadanori Konishi, J.C. Lee, Bryan Manly, K.V. Mardia, John Rayner, C. R. Rao, P.K. Sen, B.D. Sharma, Kunio Shimizu, N M. Singhi and M. S. Srivastava. We are very grateful to IOC and the following scholars for organizing the symposia in their respective research areas of expertise: Olcay Akman, S.B. Bagchi, Anoop Chaturvedi, Peter Cerone, Susmita Datta, David Edelman, Sat Gupta, Malcolm Hudson, Bing Li, Bryan Manly, Toshio Sakata, Ashis SenGupta, Takashi Seo, Bikas Sinha, Ross Sparks, David Steel, Walter Wallis and Alan Wan. Although events of September 11, 2001 stirred many cancellations, over 200 delegates from 25 countries took part in the conference activities and they were the true reason for its success.

During the conference festivities, three outstanding statisticians were recognised for their life-long and profound contributions to the field of Statistics: Professor P.K. Sen of USA, Professor Joe Gani of Australia and Professor Minoru Siotani of Japan. Each one has contributed profusely to the profession in their own countries as well as on the world stage.

Prior to the conference, Professor C. R. Rao was awarded an Honorary Doctor of Science degree by the University of Wollongong. It was his twenty seventh such award. He is the recipient of many other awards and honours. On June 12, 2002, he was given a National Medal of Science by President George W. Bush at the White House. We are delighted that he could contribute to this volume.

The programme comprised 6 Plenary Sessions, 32 Invited Sessions, 6 Contributed Sessions, 2 Students Competition Sessions, and one Poster Session. A total of 117 papers were presented at the invited sessions, with 32 papers at the contributed sessions, 9 papers at the FIM student competition sessions, and 3 papers at the poster session.

At the conclusion of the Conference, student competition session awards were presented to the nine participating students from Australia, India, Japan and USA. Papers were judged by a panel of judges with Professor Joe Gani as chair and Ashis SenGupta as the major force behind the organization of the competition. To them we say thank you! Tomohiro Ando (Kyushu University), Riccardo Biondini (University of Wollongong), and Yee Hwa Yang (University of California, Berkeley) shared the first prize pool of $600. Papers by Tomohiro Ando and Riccardo Biondini are included in this volume.

Some papers from this conference are included in a special issue of the Journal of Applied Mathematics and Decision Sciences, Vol. 6, Number 4. We wish to record our special thanks to John Best for his involvement in editing and seeing through this whole project.

We wish to express our sincerest thanks to the following referees for their help in the reviewing process: Raid Amin, Dheeraj Bansal, Adrian Baddeley, S. Bagui, John Best, Jennifer Brown, Martin Bunder, Pam Davy, Somnath Datta, Susmita Datta, Carol Feltz, Joe Gani, Prem Goel, Alister Gray, David Griffiths, Sat Gupta, Markus Hegland, Jim Hill, Bing Li, Liwan Liyanage, Xiao Ping Lu, Geoff MacLachlan, Bryan Manly, Denny Meyer, A. Nanthakumar, Mantong Ncube, Rod Nillsen, Shelton Peiris, Ken Russell, Olivier Thas, Daniel T. Voss, Alan Wan, Neville Weber and Susan Wilson. Their effort has resulted in a considerable strengthening of many of the papers. We also take this opportunity to thank Riccardo Biondini for his considerable hard work in the conversion of files to the format required by the publishers.

Members of the local organizing committee (LOC) deserve special thanks for looking after the delegates at the conference. Included in the LOC, besides members of School of Mathematics and Applied Statistics, were Kevin Donegan, Janette Green, Kevin Price and Yong-Goo Lee. Eric Beh looked after delegates accommodation and travel with meticulous planning. Janette Green acted as treasurer of this conference. In this task, she interacted with

most of the delegates and sponsors with utmost diligence and promptness. Yong-Goo Lee joined LOC activities in Wollongong upon arrival in early December. Kevin Price was the helping hand for many of the delegates and this trio took over many responsibilities of helping the delegates during the conference, which in turn drew special praise from the conference delegates for their dedication and untiring prompt help during the conference. Without their efforts, this conference would have been a bit less enjoyable. Yong-Goo Lee (having worked on LOC of ISI, Seoul 2001) was puzzled at the many menial tasks performed by LOC and remarked that in Korea such tasks will be carried out by students rather than academic staff. However, he, along with other members of LOC carried out such tasks with smile. We wish to record our appreciation for Carol Rayner (assisted by Glen and Vanessa) for her untiring work in arranging the logistics of morning and afternoon teas.

Some photos taken by Anne Porter during the conference can be viewed at http://www.uow.edu.au/informatics/maths/statconference/files/photos/photos.html. Financial sponsors for the conference included the University of South Alabama (College of Arts and Sciences, Department of Mathematics and Statistics and Vice President for Research), IMMACS, and School of Mathematics and Applied Statistics at the University of Wollongong. We thank all our sponsors for their impressive support.

Editors

Chandra Gulati	Yan-Xia Lin	Satya Mishra	John Rayner
UOW, Australia	UOW, Australia	Mobile, USA	UOW, Australia

July 31, 2002

CONTENTS

EFFICIENT ESTIMATORS OF DIFFERENT TYPES FOR DOMAINS

M.C. AGRAWAL
Department of Statistics
University of Delhi
Delhi 110007, India
E-mail: mc_agrawal@yahoo.com

CHAND K. MIDHA
Department of Statistics
University of Akron
Ohio, USA

In this paper, we have proposed efficient unbiased estimators of three types - direct, synthetic and composite - for the domain mean (or total), and compared them with certain customary ones to establish their viability, superior performance and practicability.

Keywords: domain (small area) estimation, direct estimators, synthetic estimators and composite estimators.

1 Introduction

Estimation at the levels of small domains and areas has been engaging the attention of survey practitioners for quite some time. It would be apt to stress that the estimation for small domains has to be geared from the point of view of precision, validity and applicability. In the context of the estimation of the domain mean, we may underscore the point that the structuring of the estimators with the intent of exploiting the available data in totality should be deemed to be a desirable feature (see Sarndal *et al.*, 1992, p.387). In essence, special methods are required to provide reliable estimators for small domains.

2 Two efficient estimators of the domain mean

Consider a population P of size N which contains a small domain, say, d of interest having size N_d. A simple random sample of size n is drawn from this population. Let s_d denote the sample of n_d units, which belong to domain d, from amongst the n units, i.e., n_d units are common to s and d. If the study

variable y assumes the value y_i on the ith unit (i = 1, 2, ..., N), then the domain mean $\overline{Y}_d$ and the domain variance S_d^2 are given by

$$\overline{Y}_d = \frac{1}{N_d}\sum_{i\in d} y_i, S_d^2 = \frac{1}{N_d - 1}\sum_{i\in d}(y_i - \overline{Y}_d)^2.$$

To estimate $\overline{Y}_d$, the estimator

$$\overline{y}_d = \frac{1}{n_d}\sum_{i\in s_d} y_i \qquad (n_d \geq 1) \tag{1}$$

is the usual one that has been considered in the literature (see Hedayat and Sinha (1991, pp.342-345) and Sarndal et al. (1992, pp.392-93)). This estimator is unbiased for $n_d \geq 1$ and its large-sample variance is expressible as

$$V(\overline{y}_d) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{S_d^2}{Q_d} + \frac{V(\alpha_d)}{nf^2Q_d}S_d^2 \doteq \left(\frac{1}{n} - \frac{1}{N}\right)\frac{S_d^2}{Q_d}\left(1 + \frac{\overline{Q}_d}{nQ_d}\right) \tag{2}$$

where

$$Q_d = \frac{N_d}{N}, \quad \alpha_d = \frac{n_d}{N_d}, \quad f = \frac{n}{N}$$

and

$$V(\alpha_d) \doteq \left(\frac{1}{n} - \frac{1}{N}\right) f\frac{\overline{Q}_d}{nQ_d}, \quad (\overline{Q}_d = 1 - Q_d).$$

Another estimator of $\overline{Y}_d$ (when N_d is known) is obtained as

$$\overline{y}_d^* = \frac{\overline{y}^*}{Q_d} \tag{3}$$

where

$$\overline{y}^* = \frac{1}{n}\sum_{i=1}^{n} y_i^*, \qquad y_i^* = \begin{cases} y_i, & \text{if } i \in d \\ 0, & \text{otherwise .} \end{cases}$$

It is unbiased and its large-sample variance is

$$V(\overline{y}_d^*) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{Q_d}(S_d^2 + \overline{Q}_d\overline{Y}_d^2) \tag{4}$$

For related details, the readers are referred to the literature as indicated above in the context of $\overline{y}_d$. It may be verified that, in practice, the estimator

$\overline{y}_d$ which does not exploit the knowledge of N_d scores over $\overline{y}_d^*$ in virtually all sampling situations. Interestingly, when N_d is not known, the estimator

$$\overline{y}_d^{**} = \frac{\overline{y}^*}{\hat{Q}_d} = \frac{\overline{y}^*}{\overline{x}},$$

where

$$\hat{Q}_d = \frac{n_d}{n} \quad \text{and} \quad x_i = \begin{cases} 1, & \text{if } i \in d \\ 0, & \text{otherwise} \end{cases},$$

has been treated by Hedayat and Sinha (1991, pp. 344-45) as a ratio-type estimator with distinct entity, while, as a matter of fact, $\overline{y}_d^{**}$ and $\overline{y}_d$ are exactly the same. Here, it would be apt to add that, when N_d is not known, the existing estimator of the domain total (say, Y_d) is obtainable from (3) as

$$\hat{Y}_d^* = N\overline{y}^* \tag{5}$$

whose large-sample variance is expressible as

$$V(\hat{Y}_d^*) = N^2 \left(\frac{1}{n} - \frac{1}{N}\right) Q_d(S_d^2 + \overline{Q}_d \overline{Y}_d^2)$$

A generalized direct estimator and its efficiency

We propose

$$\overline{y}_d^{(1)} = \frac{\beta_1 + \beta_2 \alpha_d}{\beta_1 + \beta_2 f} \overline{y}_d, \qquad (n_d \geq 1) \tag{6}$$

as an unbiased estimator of the domain mean $\overline{Y}_d$ where β_1 and β_2 are suitable weights, the other notations having already been defined in Section 2. Note that, when N_d is not known, we may consider the following twin special cases arising from (6):

(i) for estimation of the domain mean $\overline{Y}_d$, choose $\beta_2 = 0$ $(\beta_1 \neq 0)$, thus yielding $\overline{y}_d$ given by (1);

(ii) for estimation of the domain total Y_d, choose $\beta_1 = 0$ $(\beta_2 \neq 0)$, thus culminating in $\hat{Y}_d^* = N\overline{y}^*$ given by (5).

Besides, $\overline{y}_d^{(1)}$ reduces to $\overline{y}_d^*$ given by (3), when N_d is known, for $\beta_1 = 0$ $(\beta_2 \neq 0)$. The estimator $\overline{y}_d^{(1)}$ reflects the desirable statistical properties when

we consider (a) $n = N$ and (b) $N_d = N$. The variance of this estimator can be expressed as

$$V(\overline{y}_d^{(1)}) = \frac{1}{(\beta+f)^2}\left[V(\alpha_d)\left\{\overline{Y}_d^2 - \frac{S_d^2}{N_d}\right\} + \beta^2 V(\overline{y}_d)\right.$$

$$\left. + \frac{f(1-f)}{N_d}S_d^2 + 2\beta\frac{1-f}{N_d}S_d^2\right] \tag{7}$$

where $\beta = \beta_1/\beta_2$. We can choose a suitable β which renders $V(\overline{y}_d^{(1)}) \leq V(\overline{y}_d)$. However, the optimum value of β that minimizes the variance of $\overline{y}_d^{(1)}$ is given by

$$\beta_{opt} = f^2\left(\frac{N_d}{C_d^2} - 1\right)$$

where C_d^2 is the square of the coefficient of variation for domain d defined by

$$C_d^2 = \frac{S_d^2}{\overline{Y}_d^2}.$$

The coefficient of variation being a stable quantity over the time-space facilitates the computation of β_{opt} because an idea about C_d^2 from the past data would not generally be a problem. Further, since C_d^2, virtually in all practical sampling situations, would not exceed N_d, we can expect (and hence assume) $\beta_{opt} > 0$. Now, expressing the variance of $\overline{y}_d^{(1)}$ in terms of β_{opt} as

$$V_{min}(\overline{y}_d^{(1)}) = \frac{\beta_{opt}V(\overline{y}_d) + \frac{1-f}{N_d}S_d^2}{\beta_{opt}+f} = V(\overline{y}_d) - \left(\frac{1}{n} - \frac{1}{N}\right)\frac{\overline{Q}_d}{(\beta_{opt}+f)\,N_d Q_d}S_d^2,$$

we conclude the $\overline{y}_d^{(1)}$ is superior to $\overline{y}_d$. However, we present below the performance sensitivity of $\overline{y}_d^{(1)}$ in the face of perturbations in the exact value of β_{opt}.

Performance - sensitivity of $\overline{y}_d^{(1)}$

Let P_I be the proportional inflation in the variance of $\overline{y}_d^{(1)}$ if we employ some $\beta(>0)$ other than β_{opt}, i.e.,

$$P_I = \frac{V(\overline{y}_d^{(1)}) - V_{min}(\overline{y}_d^{(1)})}{V_{min}(\overline{y}_d^{(1)})}$$

wherein we consider

$$\nu = \frac{\beta - \beta_{opt}}{\beta_{opt}} \Rightarrow \beta = (1+\nu)\beta_{opt}$$

as the proportional deviation in β_{opt}, thus leading to

$$P_I = G\left(1 - \frac{\beta_{opt}^{*2}}{\left(\beta_{opt}^{*} + \nu\beta_{opt}\right)^2}\right) - \frac{2\nu\beta_{opt}\left(\frac{1}{n} - \frac{1}{N}\right)\frac{\overline{Q}_d}{N\overline{Q}_d^2}S_d^2}{\left(\beta_{opt}^{*} + \nu\beta_{opt}\right)^2 V_{min}(\overline{y}_d^{(1)})} \tag{8}$$

where G is the gain in the efficiency of $\overline{y}_d^{(1)}$ over $\overline{y}_d$, i.e.,

$$G = \frac{V(\overline{y}_d) - V_{min}(\overline{y}_d^{(1)})}{V_{min}(\overline{y}_d^{(1)})}$$

and $\beta_{opt}^{*} = \beta_{opt} + f$. It is evident that, for $\nu \geq 0$, the proposed estimator $\overline{y}_d^{(1)}$ will continue to perform better than $\overline{y}_d$ despite deviations in β_{opt} because the proportional inflation P_I would never exceed G for $\nu > 0$ as reflected by (8). However, for $\nu < 0$, we note that $\beta \geq 0 \Rightarrow \nu \geq -1$, and, then, the condition that follows from (8) to ensure $P_I \leq G$ is

$$\nu \geq -\frac{\beta_{opt} + f}{2\beta_{opt}} \tag{9}$$

which, in combination with $\nu \geq -1$, yields

$$f \leq \beta_{opt} \tag{10}$$

- a condition that is usually found to hold in practice. Hence, whenever $\beta_{opt} \geq f$, $\overline{y}_d^{(1)}$ can be looked upon as being robust to deviations in β_{opt} (as envisaged and spelt out above) in the sense of maintaining its superiority vis-a-vis $\overline{y}_d$ whatever be ν that ensures some $\beta \geq 0$.

Unbiased variance-estimation and conditional variance

An unbiased variance-estimator of $\overline{y}_d^{(1)}$ is

$$\hat{V}(\overline{y}_d^{(1)}) = \frac{1}{(\beta+f)^2}\left[\left(\overline{y}_d^2 - \frac{s_d^2}{n_d}\right)V(\alpha_d) + \beta^2\left(\frac{1}{n_d} - \frac{1}{N_d}\right)s_d^2 + \frac{f(1-f)}{N_d}s_d^2 + 2\beta\frac{1-f}{N_d}s_d^2\right]$$

where

$$s_d^2 = \frac{1}{n_d - 1}\sum_{i \in s_d}(y_i - \overline{y}_d)^2.$$

In the context of estimation for domains, it would be apt to consider, for a fixed n_d, the conditional variance of $\overline{y}_d^{(1)}$ given by

$$V_c(\overline{y}_d^{(1)}|n_d) = \left(\frac{\beta_1 + \beta_2\alpha_d}{\beta_1 + \beta_2 f}\right)^2 \left(\frac{1}{n_d} - \frac{1}{N_d}\right) S_d^2.$$

Interestingly, the optimum value of β $(= \beta_1/\beta_2)$ with a view to minimizing the conditional variance is attained at $\alpha_d = f$, and this yields

$$V_{c,min}(\overline{y}_d^{(1)}|n_d) = \left(\frac{1}{n_d} - \frac{1}{N_d}\right) S_d^2 = V_c(\overline{y}_d|n_d)$$

where $V_c(\overline{y}_d|n_d)$ is the conditional variance of $\overline{y}_d$ for a fixed n_d.

A predictive synthetic estimator and its efficiency

To arrive at a synthetic estimator, we start by splitting the population total as

$$\sum_{i\in P} y_i = \sum_{i\in d} y_i + \sum_{i\in \overline{d}} y_i = Y_d + \sum_{i\in \overline{d}} y_i \Rightarrow \sum_{i\in d} y_i = \sum_{i\in P} y_i - \sum_{i\in \overline{d}} y_i,$$

and then, to estimate Y_d, we proceed as

$$\hat{Y} = \sum_{i\in P} \hat{y}_i - \sum_{i\in \overline{d}} \hat{y}_i \tag{11}$$

where $\overline{d}$ is the complement of d and $\hat{y}_i (i \in P)$ and $\hat{y}_i (i \in \overline{d})$ are the respective predictors of $y_i (i \in P)$ and $y_i (i \in \overline{d})$. Against this backdrop, it is sensible to use

$$\hat{y}_i = \overline{y}(i \in P), \quad \hat{y}_i = \overline{y}_{\overline{d}}(i \in \overline{d})$$

and, thus, (11) leads to

$$\hat{Y}_d = N\overline{y} - (N - N_d)\overline{y}_{\overline{d}} \Rightarrow \hat{\overline{Y}}_d = \frac{\overline{y}}{Q_d} - \frac{1 - Q_d}{Q_d}\overline{y}_{\overline{d}} = \overline{y}_d^{(2)}, \quad \text{say} \tag{12}$$

where $\overline{y}_{\overline{d}}$ is the sample-based mean of the non-domain $(\overline{d})$ given by

$$\overline{y}_{\overline{d}} = \frac{1}{n_{\overline{d}}} \sum_{i\in s\cap \overline{s}_d} y_i \quad , \qquad n_{\overline{d}} = n - n_d$$

which unbiasedly estimates the population mean of $\overline{d}$, i.e., $\overline{Y}_{\overline{d}}$ defined as

$$\overline{Y}_{\overline{d}} = \frac{1}{N_{\overline{d}}} \sum_{i\in \overline{d}} y_i \quad , \qquad N_{\overline{d}} = N - N_d.$$

The estimator $\overline{y}_d^{(2)}$ is unbiased, and its approximate variance is

$$V(\overline{y}_d^{(2)}) = \frac{1}{Q_d^2}\left[\left(\frac{1}{n}-\frac{1}{N}\right)\{S^2-(1-Q_d)S_{\bar{d}}^2\}\right]$$

$$= V(\overline{y}_d) - \left(\frac{1}{n}-\frac{1}{N}\right)\frac{\overline{Q}_d}{nQ_d^2}S_d^2 + \left(\frac{1}{n}-\frac{1}{N}\right)\frac{1}{Q_d\overline{Q}_d}(\overline{Y}-\overline{Y}_d)^2 \qquad (13)$$

where $S_{\bar{d}}^2$ and S^2 are, respectively, the variances of the non-domain $(\bar{d})$ and the population defined by

$$S_{\bar{d}}^2 = \frac{1}{N_{\bar{d}}-1}\sum_{i\in\bar{d}}(y_i-\overline{Y}_{\bar{d}})^2 \text{and} S^2 = \frac{1}{N-1}\sum_{i\in P}(y_i-\overline{Y})^2.$$

In keeping with the spirit of synthetic estimation which is involved when the characteristics of the domain and those of the population are not different, we can expect $\overline{Y}_d$ to be close to $\overline{Y}$. Thus, in the case when $\overline{Y}_d = \overline{Y}$, it can be concluded from (13) that $\overline{y}_d^{(2)}$ will fare better than $\overline{y}_d$. Otherwise, the condition

$$C_d^2 \geq \frac{\overline{Q}_d^2(1-\overline{Y}/\overline{Y}_d)^2}{nQ_d},$$

which will hold if n is large, is required to be met for the superiority of $\overline{y}_d^{(2)}$ over $\overline{y}_d$. The variance-estimator of $\overline{y}_d^{(2)}$ is

$$\hat{V}(\overline{y}_d^{(2)}) = \frac{1}{Q_d^2}\left[\left(\frac{1}{n}-\frac{1}{N}\right)\{s^2-(1-Q_d)s_{\bar{d}}^2\}\right]$$

where $s_{\bar{d}}^2$ and s^2 are the estimators of $S_{\bar{d}}^2$ and S^2, respectively, defined by

$$s_{\bar{d}}^2 = \frac{1}{n_{\bar{d}}-1}\sum_{i\in s\cap\bar{s}_d}(y_i-\overline{y}_{\bar{d}})^2, \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\overline{y})^2$$

An efficient composite estimator

In order to arrive at a composite estimator, we combine the synthetic estimator $\overline{y}_d^{(2)}$ with the direct estimator $\overline{y}_d$ (defined in Section 2), thus obtaining

$$\overline{y}_d^{(3)} = W\left(\frac{\overline{y}}{Q_d} - \frac{(1-Q_d)\overline{y}_{\bar{d}}}{Q_d}\right) + (1-W)\overline{y}_d \qquad (14)$$

where W is a suitably chosen weight. It can be verified that $\overline{y}_d^{(3)}$ is unbiased. As a matter of fact, if we decide to exploit the available data relating to the

sample s, the domain d and the non-domain $\bar{d}$ in an optimal fashion with a view to achieving minimum variance and unbiasedness, we are led to $\overline{y}_d^{(3)}$. To demonstrate this, let us consider

$$\overline{y}_d^{(4)} = A\overline{y} + B\overline{y}_{\bar{d}} + C\overline{y}_d$$

where A, B and C are the constants to be suitably determined. The condition of unbiasedness for $\overline{y}_d^{(4)}$ yields

$$C = -AQ_d, \quad B = -A(1 - Q_d)$$

This renders

$$\overline{y}_d^{(4)} = A\{\overline{y} - (1 - Q_d)\overline{y}_{\bar{d}}\} + (1 - AQ_d)\overline{y}_d$$

which is the same as $\overline{y}_d^{(3)}$. Now, we find the variance of $\overline{y}_d^{(3)}$ as

$$V(\overline{y}_d^{(3)}) = \frac{W^2}{Q_d^2}\left[\left(\frac{1}{n} - \frac{1}{N}\right)\{S^2 - (1 - Q_d)S_d^{*2}\} + (1 - W)^2 V(\overline{y}_d)\right.$$

$$\left. +2\frac{W}{Q_d}(1 - W)\left(\frac{1}{n} - \frac{1}{N}\right)S_d^2\right] \tag{15}$$

To obtain the optimum value of W which minimizes the variance of $\overline{y}_d^{(3)}$, we can either proceed with the variance expression given by (15) by differentiating it and equating to zero, or, alternatively, invoke a zero function defined by

$$z = \overline{y} - (1 - Q_d)\overline{y}_d^* - Q_d\overline{y}_d$$

and apply the necessary and sufficient condition

$$Cov(\overline{y}_d^{(3)}, z) = 0$$

to arrive at the requisite optimum value of W. The solution for optimum W can be worked out as

$$W_{opt} = \frac{\overline{Q}_d^2 C_d^2}{nQ_d(1 - \overline{Y}/\overline{Y}_d)^2 + \overline{Q}_d^2 C_d^2},$$

leading to the minimum variance

$$V_{min}(\overline{y}_d^{(3)}) = Cov(\overline{y}_d^{(3)}, \overline{y}) = V(\overline{y}_d) - W_{opt}\left(V(\overline{y}_d) - \left(\frac{1}{n} - \frac{1}{N}\right)\frac{S_d^2}{Q_d}\right)$$

$$= V(\overline{y}_d) - \frac{\left(\frac{1}{n} - \frac{1}{N}\right)\overline{Q}_d^3 C_d^4 \overline{Y}_d^2}{n^2 Q_d^2(1 - \overline{Y}/\overline{Y}_d)^2 + n\overline{Q}_d^2 C_d^2 \overline{Y}_d^2},$$

implying thereby that $\overline{y}_d^{(3)}$ is more efficient than $\overline{y}_d$ if W is optimally chosen. Besides, $\overline{y}_d^{(3)}$ fares better than $\overline{y}_d^{(1)}$ if

$$C_d^2 \leq N_d \left\{ 1 - \frac{(1 - \overline{Y}/\overline{Y}_d)^2}{\overline{Q}_d^2} \right\}.$$

However, if $\overline{Y}/\overline{Y}_d \leq Q_d$, then $\overline{y}_d^{(1)}$ always performs better than $\overline{y}_d^{(3)}$.

Performance-sensitivity of $\overline{y}_d^{(3)}$

In a manner analogous to the one followed with regard to $\overline{y}_d^{(1)}$, we examine the performance-sensitivity of $\overline{y}_d^{(3)}$ by appraising the proportional inflation for $\overline{y}_d^{(3)}$ defined by

$$P_I' = \frac{V(\overline{y}_d^{(3)}) - V_{min}(\overline{y}_d^{(3)})}{V_{min}(\overline{y}_d^{(3)})}$$

which, after some algebra, simplifies to

$$P_I' = \nu'^2 G'$$

where

$$G' = \frac{V\overline{y}_{\overline{d}} - V_{min}(\overline{y}_d^{(3)})}{V_{min}(\overline{y}_d^{(3)})}$$

and

$$\nu' = \frac{W - W_{opt}}{W_{opt}} \Rightarrow W = (1 + \nu')W_{opt}$$

are, respectively, the gain in efficiency of $\overline{y}_d^{(3)}$ over $\overline{y}_d$ and the proportional deviation in W_{opt}. Thus, the estimator $\overline{y}_d^{(3)}$, despite the use of some $W(> 0)$ other than W_{opt}, will continue to fare better than $\overline{y}_d$ provided $P_I' \leq G' \Rightarrow \nu'^2 \leq 1$, i.e., some W different from W_{opt} that results in proportional deviation of upto 100% will still ensure superiority of $\overline{y}_d^{(3)}$ vis-a-vis $\overline{y}_d$. This, indeed, puts a premium on the use of the proposed estimator $\overline{y}_d^{(3)}$ in preference to the conventional estimator $\overline{y}_d$.

3 Numerical illustration

To illustrate the findings of the preceding section, we consider the following example from Sarndal et al. (1992, pp. 414-415).

Example: The following data relate to the study variable y (number of conservative seats in Municipal Council) of a Swedish population that comprises three regions (domains):

Region	N_d	$\sum_{i \in d} y_i$	$\sum_{i \in d} y_i^2$
1	47	349	3375
2	50	437	4159
3	45	194	1000

For the purpose of application of the results of Section 2, we consider the region 1 as the domain d of our interest. The following quantities relate to the above population or the region 1:

$N = 142$, $\overline{Y} = 6.9014$, $\overline{Y}_d = 7.4255$, $S^2 = 12.5576$, $S_d^2 = 17.0324$ and $Q_d = 0.3310$.

Taking a sample of size 15, we compute the variances of estimators - $\overline{y}_d$, $\overline{y}_d^{(1)}$, $\overline{y}_d^{(2)}$ and $\overline{y}_d^{(3)}$ as

$V(\overline{y}_d) = 3.4816, V(\overline{y}_d^{(1)}) = 3.4592$ (based on $\beta_{opt} = 1.6866$), $V(\overline{y}_d^{(2)}) = 3.1422$ and $V(\overline{y}_d^{(3)}) = 3.1309$ (based on $W_{opt} = 0.8482$),

thus concluding that the estimator $\overline{y}_d^{(3)}$ is the best followed, in decreasing order of efficiency, by $\overline{y}_d^{(2)}, \overline{y}_d^{(1)}$ and $\overline{y}_d$. It would be apt to add that choice of some W (different from W_{opt}) for continued superiority of $\overline{y}_d^{(3)}$ can be regulated by the findings of Section 2.

REFERENCES

Hedayat, A.S. and Sinha, B.K. (1991). *Design and Inference in Finite Population Sampling.* New York: John Wiley and Sons.

Sarndal, C.E., Swensson, B. and Wretman, J. (1992). *Model Assisted Survey Sampling.* New York: Springer-Verlag.

NEURAL NETWORK NONLINEAR REGRESSION MODELING AND INFORMATION CRITERIA

TOMOHIRO ANDO AND SADANORI KONISHI
Graduate School of Mathematics
Kyushu University, 6-10-1 Hakozaki
Higashi-ku, Fukuoka 812-8581, Japan.
E-mail: ando@math.kyushu-u.ac.jp

We consider the problem of constructing nonlinear regression models, using multilayer perceptrons and radial basis function network with the help of the technique of regularization. Crucial issues in the model building process are the choices of the number of basis functions, the number of hidden units and a regularization parameter. We consider the properties of nonlinear regression modeling based on neural networks, and investigate the performance of model selection criteria from an information-theoretic point of view.

Keywords: information criteria, model misspecification, multilayer perceptrons, radial basis function networks, regularization.

1 Introduction

Recently, intensive investigations have been made concerning the problem of constructing various types of nonlinear models such as neural networks, kernel, splines, etc. In the field of artificial neural networks multilayer perceptrons and radial basis function networks have emerged as multilayer networks, and the advantages and disadvantages have been pointed out both in theoretical and practical aspects (Bishop (1995), Ripley (1996), Webb (1999), Yee and Haykin (2001) and the references given therein). The purpose of the paper is constructing nonlinear regression models based on these two networks with the help of the technique of regularization.

The problems still remain in constructing neural network regression models from a finite and noisy data; determining the number of hidden units in the multilayer perceptron network, the number of basis functions in the radial basis function network and choosing an appropriate value of a regularization parameter. Cross validation (Stone (1974)) is often referred in the neural network literature. An advantage of cross validation lies in its independence of probabilistic assumptions. The computational time is however enormous for a large sample size due to the use of nonlinear optimization schemes. Neural networks often treat a huge amount of data, so we take an analytical ap-

proach to determining the adjusted parameters from an information-theoretic viewpoint.

Murata et al. (1994) introduced a network information criterion (NIC) to evaluate artificial neural network models under a general loss function including a regularization term. Ando et al. (2001) proposed nonlinear regression models based on radial basis function networks with hyperparameter, and presented an information-theoretic criterion for evaluating network models estimated by regularization.

We consider the properties of nonlinear regression modeling based on the multilayer perceptron and the radial basis function network, and investigate the performance of model selection criteria from an information-theoretic point of view. This article is organized as follows. In Section 2, we briefly review the two types of network modeling strategies through a nonlinear regression with Gaussian noise. Section 3 describes model evaluation and selection criteria for network statistical models. In Section 4, we conduct a Monte Carlo experiment to investigate the performance of model evaluation criteria through the radial basis function network Gaussian and non-Gaussian regression models.

2 Neural Network Nonlinear Regression Modeling

Models

Suppose we have n independent observations $\{(y_\alpha, \boldsymbol{x}_\alpha); \alpha = 1, 2, ..., n\}$, where y_α are random response variables and $\boldsymbol{x}_\alpha$ are vectors of d-dimensional explanatory variables. In order to draw information from the data, we use the Gaussian nonlinear regression model

$$y_\alpha = u(\boldsymbol{x}_\alpha) + \varepsilon_\alpha, \qquad \alpha = 1, ..., n, \tag{1}$$

where $u(\cdot)$ is an unknown smooth function and errors ε_α are independently and normally distributed with mean zero and variance σ^2. The problem is how to estimate the function $u(\cdot)$ from the observed data, for which we use multilayer perceptrons and radial basis function networks.

EMultilayer perceptrons

The multilayer perceptron with a single hidden layer is of the form

$$u_{\mathrm{MLP}}(\boldsymbol{x}; \boldsymbol{w}) = \sum_{k=1}^{m} \beta_k \psi_k \left(\alpha_{k0} + \sum_{i=1}^{d} \alpha_{ki} x_i \right) + \beta_0, \tag{2}$$

where $\boldsymbol{x} = (x_1, ..., x_d)^T$ is a vector of d-dimensional input variables, $\alpha_{k\cdot}$ are

weight parameters for k-th hidden unit in hidden layer, β_k are weight parameters for output unit, and m represents the number of hidden units. The nonlinear function $\psi_k(\cdot)$ is usually taken to be the logistic form $\psi_k(x) = 1/\{1 + \exp(-x)\}$. The Gaussian nonlinear regression model based on the multilayer perceptron is then given by

$$y_\alpha = \sum_{k=1}^{m} \beta_k \psi_k \left(\alpha_{k0} + \boldsymbol{\alpha}_k^T \boldsymbol{x}_\alpha\right) + \beta_0 + \varepsilon_\alpha, \qquad \alpha = 1, ..., n, \tag{3}$$

where $\boldsymbol{\alpha}_k = (\alpha_{k1}, ..., \alpha_{kd})^T$. The network parameters to be estimated in the model is $\boldsymbol{w} = (\alpha_{10}, \boldsymbol{\alpha}_1^T, ...\alpha_{m0}, \boldsymbol{\alpha}_m^T, \beta_0, ..., \beta_m)^T$, the $\{(d+1) \times m + m + 1\}$-dimensional vector. We also have to determine the number of hidden units.

ERadial basis function networks

The radial basis function network is expressed as a linear combination of radially symmetric nonlinear basis functions and takes the following form;

$$u_{\mathrm{RBF}}(\boldsymbol{x}; \boldsymbol{w}) = \sum_{k=1}^{m} w_k \phi_k(\boldsymbol{x}) + w_0, \tag{4}$$

where $\boldsymbol{w} = (w_0, ..., w_m)^T$ and $\phi_k(\boldsymbol{x})$ are a set of basis functions. Ando et al. (2001) introduced the Gaussian basis functions with hyperparameter ν given in the following;

$$\phi_k(\boldsymbol{x}; \boldsymbol{\mu}_k, \sigma_k, \nu) = \exp\left\{-||\boldsymbol{x} - \boldsymbol{\mu}_k||^2/(2\nu\sigma_k^2)\right\}, \qquad k = 1, ..., m, \tag{5}$$

where $\boldsymbol{\mu}_k$ is the d-dimensional vector determining the center of the radial basis function for unit k, σ_k is the width parameter and $||\cdot||$ is the Euclidean norm. The hyperparameter ν adjusts the amount of overlapping among basis functions so that the estimated regression function captures the structure in the data over the region of the input space.

Among possible strategies for determining the centers and widths of the basis functions we use a k-means clustering algorithm. This algorithm divides the input data set $\{\boldsymbol{x}_\alpha; \alpha = 1, ..., n\}$ into m clusters $A_1, ..., A_m$ that correspond to the number of the basis functions. The centers and width parameters are then determined by $\boldsymbol{c}_k = \sum_{\alpha \in A_k} \boldsymbol{x}_\alpha / n_k$, $s_k^2 = \sum_{\alpha \in A_k} ||\boldsymbol{x}_\alpha - \boldsymbol{c}_k||^2 / n_k$, respectively, where n_k is the number of the observations which belong to the k-th cluster A_k. By replacing the center $\boldsymbol{\mu}_k$ and the width parameter σ_k in (5) by $\boldsymbol{c}_k$ and s_k respectively, we have the radial basis function network regression model with hyperparameter ν in the form

$$y_\alpha = \sum_{k=1}^{m} w_k \exp\left\{-||\boldsymbol{x}_\alpha - \boldsymbol{c}_k||^2/(2\nu s_k^2)\right\} + \varepsilon_\alpha, \qquad \alpha = 1, ..., n. \tag{6}$$

In order to implement this modeling strategy we must estimate the weight parameters $\boldsymbol{w} = (w_0, ..., w_m)^T$ and choose the hyperparameter ν and the number of basis functions. For radial basis function network we refer to Broomhead and Lowe (1988), Moody and Darken (1989), Poggio and Girosi (1990), Bishop (1995), Webb (1999), Yee and Haykin (2001).

Model parameter estimation

The log-likelihood function of the Gaussian nonlinear regression model may be in general expressed as

$$l(\boldsymbol{w}, \sigma^2) = -n \log(2\pi\sigma^2)/2 - \sum_{\alpha=1}^{n} \{y_\alpha - u_{(.)}(\boldsymbol{x}_\alpha; \boldsymbol{w})\}^2/(2\sigma^2).$$

In fitting data with complex structure and high-dimensional data, the maximum likelihood method does not yield satisfactory results, since it often involves overfitting and yields unstable parameter estimates. We therefore estimate the unknown weights and the error variance by regularization. Instead of maximizing the log-likelihood function, the network parameter vector $\boldsymbol{w}$ and the error variance σ^2 are estimated such that they maximize the penalized log-likelihood function

$$l_\lambda(\boldsymbol{w}, \sigma^2) = l(\boldsymbol{w}, \sigma^2) - \frac{n\lambda}{2} \boldsymbol{w}^T Q \boldsymbol{w}, \tag{7}$$

where λ is a regularization parameter and Q is a known positive semidefinite matrix. The regularization parameter λ and the matrix Q have an effect on reducing the variances of the network parameter estimates.

In the multilayer perceptron network (3) the unknown parameters are estimated by employing a nonlinear optimization scheme, so the computation time is excessive and a careful initialization of parameter values is required. In contrast to the multilayer perceptron, the maximum penalized likelihood estimates of the weight parameters in the radial basis function network model (6) are explicitly given by

$$\hat{\boldsymbol{w}} = \left(\Phi_\nu^T \Phi_\nu + n\beta Q\right)^{-1} \Phi_\nu^T \boldsymbol{y},$$

where $\Phi_\nu = (\boldsymbol{\phi}(\boldsymbol{x}_1), ..., \boldsymbol{\phi}(\boldsymbol{x}_n))^T$, $\boldsymbol{\phi}(\boldsymbol{x}_\alpha) = (1, \phi_1(\boldsymbol{x}_\alpha; \nu), ..., \phi_m(\boldsymbol{x}_\alpha; \nu))^T$, $\boldsymbol{y} = (y_1, ..., y_n)^T$ and $\beta = \lambda\sigma^2$ is a regularization parameter. Hence the radial basis function network model is free from the nonconvergence and identification problems.

After estimating the network parameters, the error variance σ^2 is estimated by $\hat{\sigma}^2 = \sum_{\alpha=1}^{n} \{y_\alpha - \hat{y}_\alpha\}^2 / n$ where $\hat{y}_\alpha = u_{(.)}(\boldsymbol{x}_\alpha; \hat{\boldsymbol{w}})$. The statistical

model constructed by the regularization method is then

$$f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}}) = (2\pi\hat{\sigma}^2)^{-1/2}\exp\left[-\left\{y_\alpha - u_{(\cdot)}(\boldsymbol{x}_\alpha;\hat{\boldsymbol{w}})\right\}^2/(2\hat{\sigma}^2)\right], \tag{8}$$

where $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{w}}^T, \hat{\sigma}^2)^T$.

The regression models (3) and (6) with small number of parameters might not have enough flexibility to capture the nonlinearity in the data. The model with too many parameters adapts too closely to the observed data. In these cases the estimated regression models do not produce acceptable prediction values for future observations. Concerning the regularization parameter, smaller values will produce a curve with too much fitting and lager values lead the network regression model to be underfitted. In general, as the model complexity is increased, the variance of the estimated model tends to increase and the bias tends to decrease, and conversely as the model complexity is decreased.

The problem is how to choose the regularization parameter λ, the number of basis functions m and the hyperparameter ν in the radial basis function network, and the number of hidden units in the multilayer perceptron network by a suitable model evaluation criterion.

3 Information criteria

Akaike's (1973, 1974) information criterion AIC provides a useful tool for constructing statistical models. AIC is, however, a criterion for evaluating models estimated by the maximum likelihood method, and it can be derived under the assumption that the true distribution belongs to the specified parametric family of probability distributions. Hence the criterion AIC is not theoretically justified for the evaluation of a model estimated by regularization, even if the neural network model encompasses the true structure.

Information-theoretic criteria may be constructed as estimators of the Kullback-Leibler measure of discriminatory information (Kullback and Leibler (1951)) between the true model, $g(y|\boldsymbol{x})$, and the fitted model $f(y|\boldsymbol{x};\hat{\boldsymbol{\theta}})$ or, equivalently, the expected log-likelihood

$$\eta\left(g : f(y|\boldsymbol{x};\hat{\boldsymbol{\theta}})\right) = \int g(y|\boldsymbol{x})\log f(y|\boldsymbol{x};\hat{\boldsymbol{\theta}})dy \tag{9}$$

(see, e.g., Konishi and Kitagawa (1996)). Ando et al. (2001) obtained an information criterion for evaluating the radial basis function network regression model estimated by regularization under model misspecification. Let

$$\psi(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta}) = \log f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{\theta}) - \lambda\boldsymbol{w}^T Q\boldsymbol{w}/2, \tag{10}$$

where $\boldsymbol{\theta}$ denotes the parameter included in the model. Then the result is given by

$$\mathrm{GIC}(m,\lambda,\nu) = -2\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}}) + 2\mathrm{tr}\left(I_G J_G^{-1}\right), \tag{11}$$

where I_G and J_G are $(m+2)\times(m+2)$ matrices given by

$$I_G = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial\psi(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}}\frac{\partial\log f(y_\alpha|\boldsymbol{x}_\alpha,\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}^T}, \quad J_G = -\frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial^2\psi(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}.$$

This criterion was obtained by correcting the asymptotic bias of the log-likelihood $\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})/n$ in estimating the expected log-likelihood (9).

Murata et al. (1994) proposed the network information criterion NIC. For evaluating the radial basis function network regression model estimated by regularization, NIC is given in the following;

$$\begin{aligned}\mathrm{NIC}(m,\lambda,\nu) &= -2\sum_{\alpha=1}^{n}\psi(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}}) + 2\mathrm{tr}\left(I_N J_N^{-1}\right)\\ &= -2\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}}) + n\lambda\hat{\boldsymbol{w}}Q\hat{\boldsymbol{w}} + 2\mathrm{tr}\left(I_N J_N^{-1}\right),\end{aligned} \tag{12}$$

where I_N and J_N are $(m+2)\times(m+2)$ matrices given by

$$I_N = \frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial\psi(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}}\frac{\partial\psi(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}^T}, \quad J_N = -\frac{1}{n}\sum_{\alpha=1}^{n}\frac{\partial^2\psi(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}.$$

The adjusted parameters λ, ν and m are determined as the minimizer of GIC or NIC. A comprehensive survey of model selection was given by Rao and Wu (2001).

4 Numerical results

Monte Carlo experiments were conducted to examine the performance of the information criteria GIC in (11) and NIC in (12) through the radial basis function network Gaussian and non-Gaussian regression models estimated by the maximum penalized likelihood method.

Gaussian nonlinear regression models

In the Monte Carlo simulations, data sets $\{(x_\alpha,y_\alpha);\alpha=1,...,n\}$ are repeatedly generated from the true regression model $y_\alpha = u(x_\alpha)+\varepsilon_\alpha$ for

$x_\alpha = (2\alpha - 1)/(2n)$. For the true function, $u(x)$, we arbitrarily assumed the following functions: (a) $u(x) = \sin(2\pi x)$, (b) $u(x) = -4x(1-x) + \sin(4\pi x)/4$, (c) $u(x) = \exp(-2x)\sin(4\pi x)/4$. The errors ε_α are assumed to be independently distributed according to a mixture of two normal distributions

$$g(\varepsilon_\alpha) = \gamma \frac{1}{\sigma_1}\phi\left(\frac{\varepsilon_\alpha}{\sigma_1}\right) + (1-\gamma)\frac{1}{\sigma_2}\phi\left(\frac{\varepsilon_\alpha}{\sigma_2}\right), \tag{13}$$

where $\phi(x)$ denotes the density function of a standard normal distribution and the standard deviations are taken as $\sigma_1 = 0.4R_w$ and $\sigma_2 = 1.2R_w$ with R_w being the range of $u(x)$ over $[0,1]$.

The Gaussian nonlinear regression model in (8) is estimated based on a data set generated from the true model with the mixture density (13). In the modeling process, the weight parameters are estimated by maximizing the penalized log-likelihood function (7) with penalty term $\sum_{j=2}^{m}(\Delta^2 w_j)^2$ where Δ is a difference operator such as $\Delta w_j = w_j - w_{j-1}$. In this case positive semidefinite matrix Q in (7) can be represented by $Q = D_2^T D_2$ where D_2 is $(m-1)\times(m+1)$ matrix representation for the difference operator Δ^2.

The bias of the log-likelihood in estimating the expected log-likelihood is given by

$$\begin{aligned} nb(\hat{G}) &= E_{g(\boldsymbol{y}|X)}\left[\sum_{\alpha=1}^{n}\log f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}}) - n\eta\left(g : f(y_\alpha|\boldsymbol{x}_\alpha;\hat{\boldsymbol{\theta}})\right)\right] \\ &= E_{g(\boldsymbol{y}|X)}\left[-\frac{n}{2} + \frac{\gamma n\sigma_1^2 + (1-\gamma)n\sigma_2^2}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^2}\sum_{\alpha=1}^{n}\{u(x_\alpha) - \hat{\boldsymbol{w}}^T\boldsymbol{\phi}(x_\alpha;\nu)\}^2\right]. \end{aligned}$$

Figure 1 plots the true bias and the asymptotic bias estimates of GIC and NIC respectively given by equations (11) and (12) for values of a mixing proportion γ. The exact values of the bias were calculated numerically by using a Monte Carlo simulation with 100,000 repetitions.

We observe from Figure 1 that the log-likelihood of a fitted model has a significant bias as an estimate of the expected log-likelihood, and that the bias increases with the deviation from normality that is reflected by the value of a mixing proportion γ. The difference becomes larger as the deviation from normality increases with the value of γ. In the case of $n = 100$, GIC and NIC yield almost the same bias estimates. The bias estimates of GIC converge to the true bias, while the bias estimates of NIC tend to overestimate the true bias, increasing with the number of the sample size n.

Table 1 compares the average squared error ASE $= \sum_{\alpha=1}^{n}\{u(x_\alpha) - \hat{y}_\alpha\}^2/n$ between the true and estimated functions. The simulation results were obtained by averaging over 100 repeated Monte Carlo trials. It may be seen

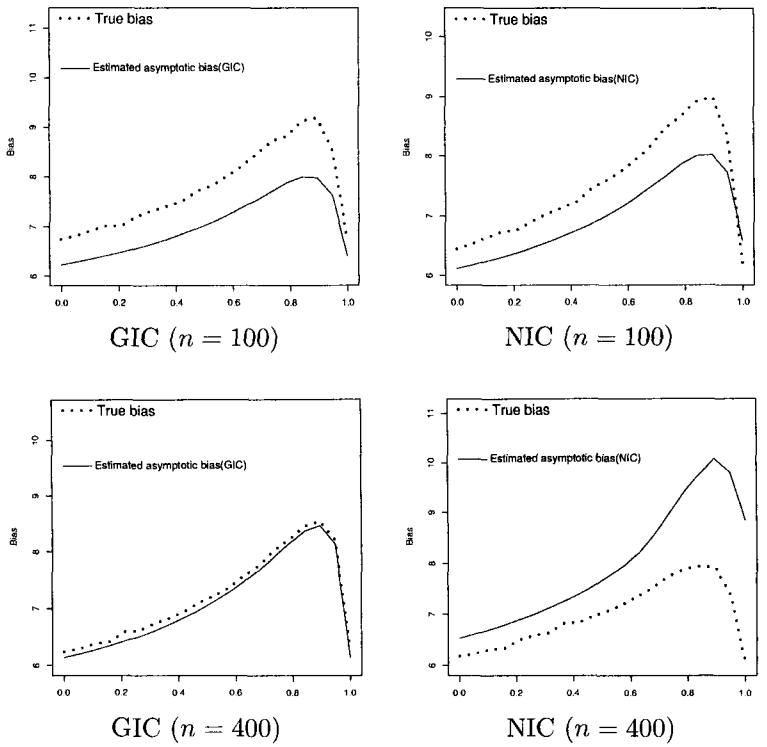


Figure 1. The true bias (- - -) and the asymptotic bias estimates of GIC and NIC (—) for values of a mixing proportion γ.

from Table 1 that the radial basis function network Gaussian regression models evaluated by GIC is superior to those by NIC in all of the situations; it gives the smallest mean value for ASE. In detail, the regularization parameter chosen by NIC was relatively smaller than that of GIC and leads to overfitting of the model. Similar comparisons were made for other combinations of sample sizes and true regression functions. We found the results described above to be essentially unchanged.

Table 1. Comparison of the average squared errors.

		True function (a)		True function (b)		True function (c)	
γ	n	GIC	NIC	GIC	NIC	GIC	NIC
0.5	100	0.2206	0.2421	0.0728	0.0785	0.0931	0.1131
0.6	100	0.1488	0.1658	0.0609	0.0681	0.0765	0.0853
0.7	100	0.1242	0.1317	0.0564	0.0616	0.0602	0.0710
0.8	100	0.0963	0.1083	0.0453	0.0536	0.0477	0.0559
0.9	100	0.0671	0.0747	0.0326	0.0400	0.0373	0.0422
1.0	100	0.0410	0.0439	0.0194	0.0212	0.0219	0.0248
0.5	200	0.1089	0.1112	0.0375	0.0405	0.0533	0.0591
0.6	200	0.0806	0.0878	0.0343	0.0380	0.0425	0.0469
0.7	200	0.0736	0.0821	0.0294	0.0311	0.0288	0.0322
0.8	200	0.0553	0.0613	0.0213	0.0285	0.0234	0.0278
0.9	200	0.0356	0.0394	0.0156	0.0187	0.0158	0.0207
1.0	200	0.0207	0.0232	0.0084	0.0100	0.0088	0.0111

Non-Gaussian regression model

A simulation study was conducted to assess the performance of the radial basis function network logistic regression model given by

$$f(y_\alpha|\boldsymbol{x}_\alpha;\boldsymbol{w}) = \pi(\boldsymbol{x}_\alpha)^{y_\alpha}(1-\pi(\boldsymbol{x}_\alpha))^{1-y_\alpha}, \qquad \alpha = 1,...,n, \tag{14}$$

where $\pi(\boldsymbol{x}_\alpha) = 1\Big/\left[1+\exp\{-\boldsymbol{w}^T\boldsymbol{\phi}(\boldsymbol{x}_\alpha;\nu)\}\right]$ is the conditional probability. We generate the binary observations $\{y_\alpha,(x_{1\alpha},x_{2\alpha});\alpha = 1,...,n\}$ according to models: (a) $\mathrm{pr}(Y=1|\boldsymbol{x}) = 1/[1+\exp\{-\sin(2\pi x_1)-\cos(2\pi x_2)\}]$, (b) $\mathrm{pr}(Y=1|\boldsymbol{x}) = 1/[1+\exp\{-\exp(x_1)\sin(\pi x_2)\}]$ and (c) $\mathrm{pr}(Y=1|\boldsymbol{x}) = 1/[1+\exp\{-x_1\sin(\pi x_2)\}]$, where the design points are uniformly distributed in $[0,2]\times[0,2]$.

We first fit the radial basis function network logistic regression model to a set of the 0-1 response data shown in Figure 2, which were generated from true structure (a). The adjusted parameters were chosen by using the criterion given by (11). Figure 3 compares the true surface with fitted conditional probability. We see that our modeling strategy captures the true structure well.

In Table 2 we compare the average squared errors ASE $= \sum_{\alpha=1}^{n}\{\mathrm{pr}(y_\alpha = 1|\boldsymbol{x}_\alpha) - \hat{\pi}(\boldsymbol{x}_\alpha)\}^2/n$ between the true and estimated conditional probabilities, in which the fitted functions are obtained using GIC and NIC. The simulation results were obtained by averaging over 100 repeated Monte Carlo trials. Table 2 shows that the radial basis function network logistic regression models evaluated by GIC is superior to those by NIC in all of the situations; it gives the smallest mean value for ASE.

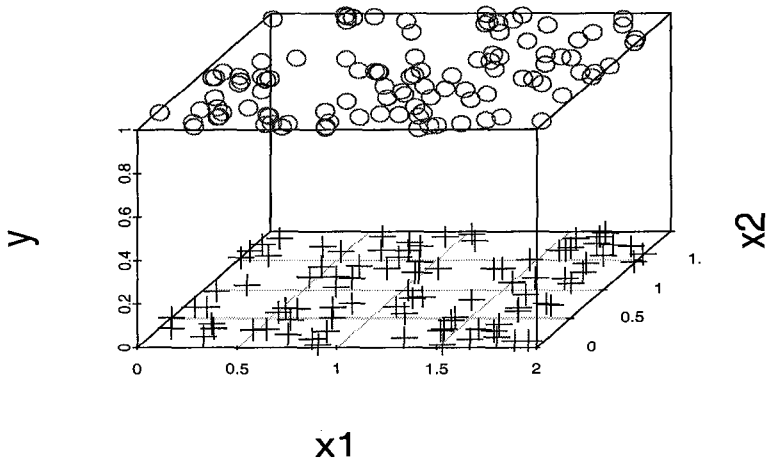
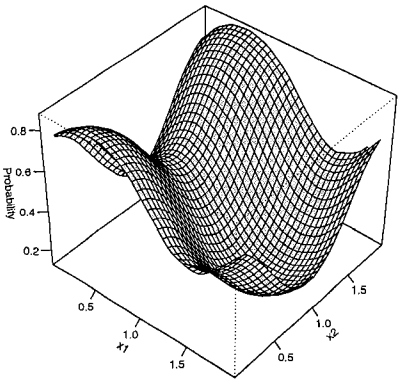


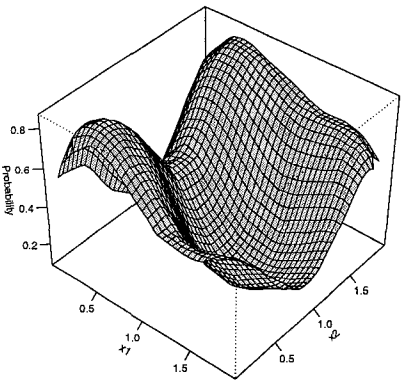
Figure 2. A set of the 0 - 1 response data generated from the true model (a) $\mathrm{pr}(Y = 1|\boldsymbol{x}) = 1/\left\{1 + \exp\left\{-\sin(2\pi x_1) - \cos(2\pi x_2)\right\}\right\}$.

Table 2. Comparison of the average squared errors ASE $= \sum_{\alpha=1}^{n}\{\mathrm{pr}(y_\alpha = 1|\boldsymbol{x}_\alpha) - \hat{\pi}(\boldsymbol{x}_\alpha)\}^2/n$.

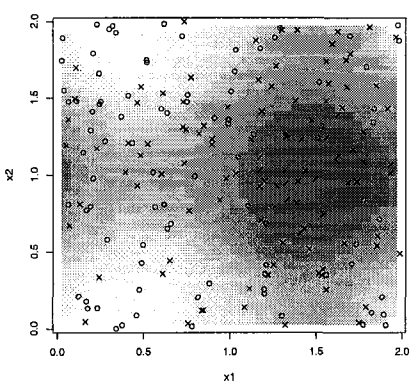
	True model (a)		True model (b)		True model (c)	
n	GIC	NIC	GIC	NIC	GIC	NIC
100	0.0412	0.0465	0.0247	0.0263	0.0216	0.0227
200	0.0289	0.0318	0.0098	0.0110	0.0133	0.0148



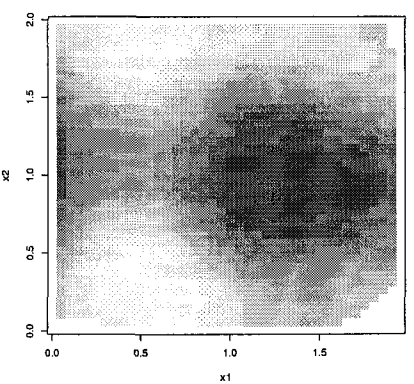
True probability: (a) $\mathrm{pr}(Y = 1|\boldsymbol{x})$.



Estimated probability: $f(y_\alpha|\boldsymbol{x}_\alpha; \hat{\boldsymbol{w}})$.



Generated data and contour image.



Estimated contour image.

Figure 3. Fitting radial basis function network logistic regression model $f(y_\alpha|\boldsymbol{x}_\alpha; \hat{\boldsymbol{w}})$ in (14) for the true model (a) $\mathrm{pr}(Y = 1|\boldsymbol{x}) = 1/\{1 + \exp\{-\sin(2\pi x_1) - \cos(2\pi x_2)\}\}$.

Acknowledgments

The authors are grateful to a referee and the Editor for their careful reading of the manuscript and helpful comments.

REFERENCES

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *2nd Inter. Symp. on Information Theory*, Petrov, B. N. and Csaki, F. (editors), Akademiai Kiado, Budapest. 267 - 281.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transaction Auto. Control.* AC - 19, 716 - 723.

Ando, T., Imoto, S. and Konishi, S. (2001). Estimating nonlinear regression models based on radial basis function networks (*in Japanese*). *Japanese J. Appl. Statist.* 30, 19 - 35.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition.* Oxford University Press.

Broomhead, D. S. and Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *Complex Systems.* 2, 321 - 335.

Konishi, S. and Kitagawa, G. (1996). Generalised information criteria in model selection. *Biometrika.* 83, 875 - 890.

Kullback, S. and Leibler, R. A. (1951). On Information and Sufficiency. *Annals of Math. Statist..* 22, 79 - 86.

Moody, J. and Darken, C. J. (1989). Fast learning in networks of locally-tuned processing units. *Neural Computation.* 1, 281 - 294.

Murata, N., Yoshizawa, S. and Amari, S. (1994). Network Information Criterion Determining the Number of Hidden Units for an Artificial Neural Network Model. *IEEE Transactions on Neural Networks.* 5, 865 - 872.

Poggio, T and Girosi, F. (1990). Networks for approximation and leaning. *Proceedings IEEE.* 78, 1481-1497.

Rao, C. R. and Wu, Y. (2001). On Model Selection (with Discussion), in *Model Selection.* Lahiri, P. (ed.). IMS Lecture Notes - Monograph Series. 38, 1 - 64.

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks.* Cambridge: Cambridge University Press.

Stone, C. J. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *J. Roy. Stat. Soc. Ser. B.* 36, 111 - 147.

Webb, A (1999). *Statistical Pattern Recognition.* London: Arnold.

Yee, P.V and Haykin, S. (2001). *Regularized Radial Basis Function Networks.* New York: Wiley.

CHISQUARED COMPONENTS AS TESTS OF FIT FOR DISCRETE DISTRIBUTIONS

D.J. BEST AND J.C.W. RAYNER
School of Mathematics and Applied Statistics
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: jbest@uow.edu.au, john_rayner@uow.edu.au

The Pearson chisquared test of fit for discrete distributions may be improved by partitioning the test statistic into useful components. Examples will be given to demonstrate the improvement possible. A brief simulation study will look at the small sample distribution of components for testing for the Poisson distribution. This paper updates material in Rayner and Best (1989).

Keywords: grouping data, orthogonal polynomials, partititon of chisquared, Poissonness.

1 Introduction

An excellent summary of chisquared goodness of fit tests is given by Moore (1986). Here we briefly discuss how to improve these tests of fit by partitioning the chisquared statistic when the categories are ordered.

Pearson (1900) used examples of ordered categories multinomial data in his famous paper introducing the chisquared goodness of fit test. Since then there have been many applications of such goodness of fit tests presented in scientific publications. A large number of these applications have concerned, as did the original examples in Pearson (1900), testing for normality when only pre-categorized or pregrouped data are available. The availability of the chisquared approximation to the distribution of Pearson's test statistic is an advantage for it when compared to competitor goodness of fit tests.

The main theme in this expository paper will be to suggest that the Pearson test can be improved by partitioning the test statistic into useful components. Colleagues have commented that this means we have a problem with significance levels, because we are not using one but many tests. As with analysis of variance, we suggest the utility of the components outweighs any necessity to stick to rigid rules of inference. The components have convenient approximate chisquared distributions, and their ready interpretation gives deep insight into the hypotheses of interest.

We emphasise that for ungrouped continuous data it may be better to

use a goodness of fit test that does not take the risk of losing information by grouping the data. In this paper our interest centres on discrete data.

2 Definitions

Suppose we wish to test the hypothesis that n observations come from an ordered categories multinomial distribution F. If there are $k+1$ categories and counts $N_j, j = 0, 1, ..., k$, then the familiar Pearson goodness of fit test is based on the statistic

$$X^2 = \sum_{j=0}^{k} (N_j - np_j)^2/(np_j),$$

where p_j is the probability under F of an observation lying in the jth category, $j = 0, 1, ..., k$, and where $n = N_0 + ... + N_k$. If the $p_0, p_1, ..., p_k$ are unknown then they are functions of the unknown parameters of the underlying hypothesised distribution. The maximum likelihood estimation of these parameters often requires iterative techniques. When this is so the easier to calculate method of moments estimators could be used. Yet another method of estimating the parameters requires the definition of the components of X^2, and this we now give.

Suppose we associate a score x_j with the jth category or class. Put

$$\mu = \sum_{j=0}^{k} x_j p_j \text{ and } \mu_r = \sum_{j=0}^{k} (x_j - \mu)^r p_j \text{ for } r = 2, 3, ...$$

If $a = \left(\mu_4 - \frac{\mu_3^2}{\mu_2} - \mu_2^2\right)^{-0.5}$ define the first three orthonormal polynomials by

$$g_0(x_j) = 1, g_1(x_j) = \frac{(x_j - \mu)}{\sqrt{\mu_2}} \text{ and } g_2(x_j) = a\left\{(x_j - \mu)^2 - \frac{(x_j - \mu)\mu_3}{\mu_2} - \mu_2\right\}.$$

These polynomials are orthonormal on the probabilities $p_0, p_1, ..., p_k$. Higher order orthonormal polynomials $g_3(x_j), ..., g_k(x_j)$ can be derived by using the recurrence relation in Emerson (1968) or the determinant formula of Lancaster (1969, p.49). If the categories or classes are equiprobable then the formulae simplify and we can use tables of orthogonal polynomials such as those in Fisher and Yates (1963). Quenouille, in the discussion of Watson (1958), gives another method of partitioning X^2 based on weighted regression which seems to be equivalent to what we suggest. See Appendix 1. Cramer

(1966, p.441) also suggests a similar idea based on an Edgeworth series for the special case of the categorized normal distribution.

Continuing with our orthogonal polynomial approach, take

$$V_m = \sum_{j=0}^{k} N_j g_m(x_j)/\sqrt{n}, \quad m = 1, ..., k.$$

It may be shown, as for example in Lancaster (1953), that X^2 may be partitioned using the V_m, so that

$$X^2 = \sum_{m=1}^{k} V_m^2.$$

We usually call the V_m^2 components of X^2.

If r parameters need estimation then estimates can be obtained by setting the first r V_m to zero and solving r (often nonlinear) equations. We conjecture that the remaining V_m are asymptotically independent standard normal. If moment or maximum likelihood estimators are calculated from the ungrouped data, then the first r V_m are not necessarily zero, and to get a statistic with an approximate chisquared distribution with $k - r$ degrees of freedom we should use $(X^2 - \sum_{m=1}^{r} V_m^2)$, as suggested by Quenouille in the discussion of Watson (1958). Often $\sum_{m=1}^{r} V_m^2$ is negligible, but this needs checking for each data set. We also conjecture that the $k - r$ remaining Vs are again asymptotically independent standard normal. The paper by Best and Rayner (2002) gives a simulation verification of this for the Geometric distribution. A similar verification for the ordered categories discrete uniform is given by Best and Rayner (1997a). Appendix 2 of this paper gives a brief simulation verification for the Poisson distribution. Best and Rayner (1997b) discuss the applicability of the asymptotic distributions for the binomial but further work is needed to verify this applicability for other distributions.

In practice data often seem to be consistent with alternatives that fall into three or four dimensions in the parameter space. Such alternatives are said to be of "low order"; see Rayner and Best (1989, section 4.3). For these alternatives the important components are often the first few. We are also interested in residuals to see if further components need to be calculated. For example, for the Poisson we focus on the residual $R = X^2 - V_2^2 - V_3^2 - V_4^2$.

3 Three examples

We now give three examples that demonstrate the utility of our approach. We obtain p-values using the approximate chisquared distribution. The justification for this was outlined in the penultimate paragraph of the previous section.

Cordial drink preference example

Best and Rayner (1997a) give an imaginary market research example where 40 consumers were asked for their colour preference of five orange-mango cordial drinks. The five drinks had different orange colours ranging from "pale" orange to "deep" orange. We may wish to test for no preferences between colours: that is, to test the fit of an ordered categories discrete uniform distribution where we associate a score of j with the jth class. The data were:

0	1	2	3	4
5	10	11	10	4

where the first row is the score and the second row gives the frequencies. We find $X^2 = 5.25$ (p-value 0.26) and $V_2^2 = 5.14$ (p-value 0.02). If we had used only a X^2 test we would have concluded that there was equal preference for the orange colours. Similarly, if we apply the Anderson-Darling statistic A^2, as discussed in Choulakian et al. (1994), we find $A^2 = 0.89$, again an insignificant value.

The significance of V_2^2 suggests departures in variance or dispersion from what is expected for a discrete uniform distribution, although some caution is needed in placing too much weight on this suggestion as it is possible to get large V_2 due to the effects of higher order moments; see, for example, Rayner et al. (1995). In the present case inspection of the data would also seem to indicate there is a significant lack of preference for both pale and deep orange, implying a variance different to that of the discrete uniform. There are no other effects as the residual $X^2 - V_2^2 = 0.11$, which is clearly nonsignificant. Power comparisons in Best and Rayner (1997a) indicate the chisquared components are often good tests of fit for the discrete uniform distribution.

The tables of orthogonal polynomials given, for example, by Fisher and Yates (1963) could be used to calculate V_2 in this particular case of equal null probabilities. However, in general the formulae above are needed and we now

give some detail of their calculation for the present example.

First we find $\mu = 3$, $\mu_2 = 2$, $\mu_3 = 0$, $\mu_4 = 6.8$ and $a = 0.5976$. Thus

$$g_1(j) = 0.7071(j - 3) \text{ and } g_2(j) = 0.5976(j^2 - 6j + 7).$$

We are given $(N_1, N_2, N_3, N_4, N_5) = (5, 10, 11, 10, 4)$ and $n = 40$. We then find $V_1 = -0.2236$ and $V_2 = -2.2678$.

Radioactive counts example

Rayner and Best (1989, p.94) quote the following data from a Rutherford and Geiger experiment relating to radioactive decay counts of Polonium.

Count:	0	1	2	3	4	5	6	7
Frequency:	57	203	383	525	532	408	273	139
Count:	8	9	10	11	12	13	14	
Frequency:	45	27	10	4	0	1	1	

Douglas (1994) and Best and Rayner (2002) both suggest grouping the data so that class expectations are greater than unity. There are two purposes: first, to ensure that the chisquared approximations for X^2 and its components are adequate, and second, that there is no great loss of information. As data-dependent classes have been constructed, this procedure will affect the distributions of the components and of X^2. However, simulations we have looked at to date suggest this effect is minimal. For example, see Appendix 2 below and simulations in Best and Rayner (2002).

In testing for a Poisson distribution with the Poisson parameter estimated by the mean of the ungrouped data, 3.87, we find $X^2 = 12.98$ (p-value 0.30, 11 degrees of freedom) when classes are pooled so that all expectations are at least unity. Further, $V_1^2 = 0.001, V_2^2 = 3.26$ (p-value 0.07), $V_3^2 = 0.16$ (p-value 0.69), $V_4^2 = 5.07$ (p-value 0.02), $R = 3.49$ (p-value 0.90). Spinelli and Stephens (1997) find the Anderson-Darling statistic $A^2 = 1.25$ with associated p-value 0.03 for these data. The large V_2 and V_4 values indicate the spread of the observed distribution does not match that of the Poisson. The value of R indicates there is likely to be at most one other large V value. As expected from the power studies in Best and Rayner (1997a), A^2 is sensitive for this data set as V_2 is large, and A^2 weights the early components more than the latter. The omnibus X^2 weights all of its components equally, and so is not as sensitive here as A^2. It appears that the effect of a few large V_m is diluted by the many nonsignificant V_m. Cramer (1966, p.436) also looked at these data and found a similar X^2 and p-value. However, he did not look at components and concluded that the Poisson was a good model for the data.

Best and Rayner (1989, pp.94-96) examined this data set using smooth tests of fit and also concluded there were dispersion and kurtosis deviations from the Poisson assumption. However, a problem with the smooth tests for the Poisson is that very large sample sizes are needed for the asymptotic distributions to be useful.

Dice data example

Pearson (1900) and Rayner and Best (1989, p.12), among others, consider Weldon's dice data where the number of 5s or 6s is counted for 26,306 throws of 12 dice:

Count:	0	1	2	3	4	5	
Frequency:	185	1149	3265	5475	6114	5194	
Count:	6	7	8	9	10	11	12
Frequency:	3067	1331	403	105	14	4	0

Rayner and Best (1989, p.136) test for a binomial with parameters $N = 12$ and $p = 1/3$, using Krawtchouk polynomials. However these are a special case of the polynomials, $g_m(x_j)$ defined above. Using the above formulae, $V_1 = 5.20$ (suggesting a mean change) and $V_2 = 0.75$. Clearly the data support a binomial parameter p greater than expected for a binomial distribution. Power comparisons in Best and Rayner (1997b) indicate that the chisquared components are often good tests of fit for the binomial distribution.

Kemp and Kemp (1991) observe that Rayner and Best (1989, p.136) give an incorrect formula for $h_2(x;p)$ and so incorrectly give $V_2 = 1.54$. The correct formula replaces the term Npq by Np^2. Guttorp (1992) observes that a likely reason for $p > 1/3$ is that "the markings of the faces of dice is commonly done using little indentations ... (which) change the center of gravity toward faces with fewer dots (and) implies an increased probability of the event (5 or 6)".

If p is estimated from the ungrouped data using maximum likelihood estimation, then $\hat{p} = 0.3377$. If counts are pooled so that all cell expectations are at least one, then there are 12 cells and $X^2 = 12.67$ with 10 degrees of freedom. The p-value is 0.24. Further $V_1 = -0.001, V_2 = 0.638, V_3 = 0.060$ and $V_4 = 0.148$, none of which are significant.

4 Conclusion

We suggest X^2 and its components are useful statistics for testing goodness of fit for discrete distributions. The components give indications of possible deviations from the hypothesised discrete distribution. We have given applica-

tions involving testing fit of the ordered categories discrete uniform, binomial and Poisson distributions. Convenient chisquared approximate distributions are available for X^2 and its components for these distributions. Goodness of fit for any univariate discrete distribution could be tested for fit via our methods for the ordered categories multinomial distribution. However further work needs to be done to verify the applicability of the chisquared approximations for other distributions. Although we do not give details here, we conjecture that the power of X^2 and its component tests can be well approximated using noncentral chisquared distributions. Best and Rayner (1997a) discuss this conjecture for the discrete uniform distribution. Rayner (2002) considers some categorized normal examples using a different method of partitioning. However, using the approach of this paper gives components almost identical to his. Further work needs to be done to establish the reason for this.

Appendix 1

Quenouille, in the discussion of Watson (1958), suggested partitioning X^2 using weighted polynomial regression. The approach given was to take $y_j = (N_j - np_j)/(np_j), x_j = j$, and weights $w_j = np_j$.

Suppose then that the regression coefficient, b say, for a straight line regression is estimated by weighted least squares so that

$$b = \sum_j w_j(x_j - \bar{x})(y_j - \bar{y}) / \sum_j w_j(x_j - \bar{x})^2$$

in which $\bar{x} = \sum_j w_j x_j / \sum_j w_j = \sum_j j p_j = \mu$ and $\bar{y} = \sum_j w_j y_j / \sum_j w_j = 0$.

Further

$$\begin{aligned} b &= \sum_j (j - \mu)(N_j - np_j) / \sum_j np_j(j - \mu)^2 \\ &= \sum_j N_j(j - \mu) / \sum_j np_j(j - \mu)^2 \\ &= \sum_j N_j(j - \mu)/(n\mu_2) = V_1/\sqrt{n\mu_2}. \end{aligned}$$

The usual regression sum of squares is $b^2 \sum_j w_j(j - \mu)^2$, which as we have just seen, is $n\mu_2 b^2 = V_1^2$. Thus the Quenouille approach is seen to be identical to an orthogonal polynomial approach for finding the first component.

Appendix 2

As a check on the validity of the chisquared approximation to the distributions of X^2, V_2^2, V_3^2, and $R = X^2 - V_2^2 - V_3^2$ in the case of testing for a Poisson distribution, Table 1 gives test sizes, based on 1,000 simulations, for α nominally 0.05, for a sample size n of 50, and for various values of the Poisson parameter lambda. As in Douglas (1994), the number of classes used in calculating X^2 varies so that each class has an expectation of at least one. An approximate standard error for the sizes in Table 1 is 0.007.

Table 1. Sizes for various statistics, λ as shown, $n = 50$ and nominal $\alpha = 0.05$.

λ	X^2	V_2^2	V_3^2	R
0.5	0.04	0.04	0.05	-
0.75	0.05	0.05	0.05	0.04
1.0	0.04	0.05	0.05	0.05
1.5	0.05	0.05	0.05	0.05
2.0	0.04	0.05	0.04	0.05

REFERENCES

Best, D.J. and Rayner, J.C.W. (1997a). Goodness of fit for the ordered categories discrete uniform distribution. *Communications in Statistics: Theory and Methods* 26, 899-909.

Best, D.J. and Rayner, J.C.W. (1997b). Goodness of fit for the binomial distribution. *Australian Journal of Statistics* 39, 355-364.

Best, D.J. and Rayner, J.C.W. (2002). Tests of fit for the geometric distribution. Submitted.

Choulakian, V., Lockhart, R.A. and Stephens, M.A. (1994). Cramer-Von Mises statistics for discrete distributions. *Canadian Journal of Statistics* 22, 125-137.

Cramer, H. (1966). *Mathematical Methods of Statistics.* Princeton: Princeton University Press.

Douglas, J.B. (1994). Empirical fitting of discrete distributions. *Biometrics* 50, 576-590.

Emerson, P.L. (1968). Numerical construction of orthogonal polynomials from a general recurrence formula. *Biometrics* 24, 695-701.

Fisher, R.A. and Yates, F. (1963). *Statistical Tables for Biological, Agricultural and Medical Research.* Edinburgh: Oliver and Boyd.

Guttorp, P. (1992). Letter to the Editor. *The American Statistician* 46, 239.

Kemp, A.W. and Kemp, C.D. (1991). Weldon's dice data revisited. *The American Statistician* 45, 216-222.

Lancaster, H.O. (1953). A reconciliation of χ^2 considered from metrical and enumerative aspects. *Sankhya* 3, 1-9.

Lancaster, H.O. (1969). *The Chi-Squared Distribution.* New York: Wiley.

Moore, D.S. (1986). Tests of chi-squared type. *Goodness of Fit Techniques*, D'Agostino, R.B. and Stephens, M.A. (editors), 63-95. New York: Dekker.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen from random sampling. *Philosophical Magazine* 5th series 50, 157-175.

Rayner, G.D. (2002). Components of the Pearson-Fisher chi-squared statistic. Submitted.

Rayner, J.C.W. and Best, D.J. (1989). *Smooth Tests of Goodness of Fit.* New York: Oxford University Press.

Rayner, J.C.W., Best, D.J. and Mathews, K.L. (1995). Interpreting the skewness coefficient. *Commun. Statist. A - Theory and Methods 24*, 593-600.

Spinelli, J.J. and Stephens, M.A. (1997). Cramer-von Mises tests of fit for the Poisson distribution. *The Canadian Journal of Statistics* 25, 257-268.

Watson, G. (1958). On chi-square goodness of fit tests for continuous distributions. *JRSS(B)* 20, 44-72 (includes discussion).

ON THE DETERMINATION OF AN APPROPRIATE HEDGING STRATEGY

RICCARDO BIONDINI
Department of Accounting and Finance
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: riccardo@uow.edu.au

YAN-XIA LIN
School of Mathematics and Applied Statistics
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: yanxia@uow.edu.au

MICHAEL MCCRAE
Department of Accounting and Finance
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: mccrae@uow.edu.au

In traditional finance theory, there is the underlying assumption that risk in both cash and futures markets is constant over time, thus disregarding the possible dynamic (time-varying) nature of the distribution of the asset returns. The implication of such an assumption is that the resultant hedge ratio will be constant through the hedging period. This article examines the criteria under which different hedging strategies are optimal from the perspective of a bona-fide hedger. The ability of dynamic approaches to minimise risk is compared to the static procedures of naive and conventional hedging. It is shown that if the hedge ratios are unstable, allowance for such stochastic movements will significantly increase hedging effectiveness by reducing the volatility of the hedged portfolio. Noting that the forecasted hedge ratio provides a constant alternative to the conventional hedge, a hedging rule is determined which enables a comparison of two constant hedge ratios, bypassing the need for transaction cost considerations. It is shown that the conventional hedge does not necessarily provide the most effective constant hedge.

Keywords: dynamic hedge, hedging effectiveness, futures markets.

1 Introduction

The most common method for determining the hedge ratio is known as the conventional hedging strategy, whereby the ratio is determined by dividing the

(unconditional) covariance between cash and futures returns by the (unconditional) variance of the futures returns. The hedge ratio may be equivalently obtained by regressing the returns to holding the spot asset ΔS on the returns to holding the futures asset ΔF. The resultant slope parameter provides the hedge ratio. In such a strategy both cash and futures positions are assumed to be held without adjustment during the cash-holding period. The hedger's subsequent position will remain unchanged (i.e. one would have a perfect hedge) provided the cash and futures returns both change by the same proportions. The assumption that the hedge ratio is constant over time translates to the covariance matrix being constant through the post-sample period (equal to the in-sample period's unconditional covariance matrix).

Dynamic hedges may be more effective than static techniques since the time path of the conditional correlation matrix may reveal information that cannot be obtained by simply computing the sample correlation statistic. In such situations static approaches fail to provide adequate hedging performance by failing to incorporate the dynamic interaction between cash and futures returns. Static hedging approaches therefore cannot generally produce optimal hedge ratios, giving rise to important concerns regarding the risk-reduction properties of such models and forcing market participants to consider the application of dynamic strategies in managing price risk. Dynamic hedge ratios have been shown to provide more optimal hedges by Bollerslev (1987) and Park and Switzer (1995) for stock index futures, Cecchetti, Cumby and Figlewski (1988) for treasury bonds, Baillie and Myers (1991) and Myers (1991) for commodities, Kroner and Sultan (1993) for foreign currency and Gagnon and Lypny (1995) for banker acceptances. However, Watt (1997) concluded that a static hedging strategy may indeed be the most effective hedging technique even though dynamic models may be found to have superior explanatory power.

The motivation behind this paper is to determine the conditions under which a particular hedging strategy provides a more optimal (in terms of risk-reduction) hedge. Specifically, in the case of constant hedging the conventional technique is shown to not necessarily be the most optimal constant hedge. A hedge based on the forecasting curve of future hedge ratios may provide an alternative static hedge to the conventional hedge in certain instances. This will depend upon appropriately modelling the in-sample data and obtaining an appropriate forecasting curve. This is referred to in this paper as the forecasting hedge.

The conventional hedging rule will not be optimal in cases where the basis is not constant. The reason is that the slope parameter from the conventional approach is merely a ratio of the unconditional covariance between cash and

futures returns to the unconditional variance of the futures returns. However, in situations where the basis is time-varying, the covariance and variance in the optimal hedging rule are clearly conditional moments that depend on information available at the time the hedging decision is made. A preferred alternative to static hedging is a generalised approach that takes proper account of relevant conditioning information. Employing ex-post data will not correctly estimate the ex-ante risk-minimising hedge ratios when the basis is not constant through the post-sample period. The calculation of the hedge ratio via the traditional method precludes dynamic updating of the hedge as circumstances change and new information arrives over time. Structural shifts between cash and futures returns are accountable for a general failure of static hedging approaches. The time-varying residual variances and covariances should be allowed to respond to price shocks and changes in volatility by systematically allowing the covariance matrix to be updated over time as new information arrives at the marketplace. Improvements to hedging performance may be made through the implementation of strategies that involve more than simply buying and holding a fixed futures position over the entire cash-holding period.

This paper is organised as follows. Section two outlines the notation used in the analysis, section three provides a mathematical comparison of different hedging strategies. In section four a constant alternative to the conventional hedge is proposed based on the forecasted hedge ratios. Section five provides an example of constant hedge comparisons via practical data and section six concludes.

2 Notation

The GARCH model is basically a time-dependent conditional variance model which allows the second moment of the distribution of returns to change through time (Bollerslev, 1986). In such models, the variance of the series is stationary over the long-term (i.e. it does not increase or decrease ad infinitum, or trend). However the variance does deviate from this stationary process in the short term. The time-varying residual variances and covariances are allowed to respond to price shocks and changes in volatility by systematically allowing the covariance matrix to be updated over time as new information arrives at the marketplace. GARCH models have been employed successfully in modelling dynamic volatilities and correlations by not only incorporating heteroscedasticity observed in economic and financial data but also having the advantage of capturing the tendency for such data to exhibit leptokurtosis, skewness and volatility clustering.

GARCH models extended to the multivariate setting are known as MGARCH models. Such models are often invoked to estimate dynamic correlations of financial assets. The MGARCH model (for the bivariate case) will now be introduced in detail. Define y_t as an observable 2×1 vector process of the returns on the cash and futures positions. This implies that extension of the univariate model must take into account time-variability not only of a vector of conditional means, but also of the 2×2 covariance matrix $\mathbf{H}_t$. A simple model of the interaction between the returns is given by

$$y_t = \mu + \epsilon_t, \qquad \epsilon_t \sim N(0, \mathbf{H}_t)$$

where μ is a 2×1 vector of constants, $\epsilon_t = [\epsilon_{1,t} \ \epsilon_{2,t}]'$ ($'$ denoting transpose) is a 2×1 vector of errors and $\mathbf{H}_t$ defines the symmetric (2×2) conditional covariance matrix between the returns where

$$\mathbf{H}_t = \begin{pmatrix} H_{11,t} & H_{12,t} \\ H_{21,t} & H_{22,t} \end{pmatrix}.$$

The next section will outline the situations under which various hedging strategies may be accepted as more optimal than others.

3 Comparison of hedging methods

Comparison of hedging methods centres on the calculation of the portfolio return at each post-sample period t. The portfolio return at time t is defined by

$$\Delta S_t - h_t \Delta F_t,$$

where h_t is the hedge ratio determined given information available at time t. The (unconditional) variance of these returns is then calculated and the hedging method which leads to the smallest variance of portfolio returns is regarded as the optimal hedging method. Furthermore, within this period it may be possible to ascertain certain time periods where various hedging strategies are likely to provide more optimal hedges. Hedging effectiveness is traditionally calculated by taking unconditional variances. The novelty of this analysis is that conditional variances will also be analysed, the logic behind this is that a certain hedging method may be optimal in certain periods but not in others.

In the case of naive hedging, $h_t = 1$ for all t. Therefore the conditional variance of such a hedged portfolio is

$$Var_{t-1}(\Delta S_t - \Delta F_t) = H_{11,t} - 2H_{12,t} + H_{22,t},$$

where $H_{11,t}$ denotes the conditional variance of cash returns at time t, $H_{12,t}$ denotes the conditional covariance between cash and futures returns at time t and $H_{22,t}$ denotes the conditional variance of futures returns at time t.

The conditional variance of a portfolio hedged by the conventional minimum-variance procedure is

$$Var_{t-1}(\Delta S_t - h_{ls}\Delta F_t) = H_{11,t} - 2h_{ls}H_{12,t} + h_{ls}^2 H_{22,t},$$

where h_{ls} is the hedge ratio obtained via the conventional (least-squares) technique.

Where the hedge is constructed via time-varying techniques, $h_t = \frac{H_{12,t}}{H_{22,t}}$. The conditional variance of the resultant hedged portfolio is

$$Var_{t-1}(\Delta S_t - h_t\Delta F_t) = H_{11,t} - \frac{H_{12,t}^2}{H_{22,t}}. \tag{1}$$

The quantity on the right-hand side of (1) will be smaller as $H_{12,t}^2$ approaches $H_{11,t}H_{22,t}$. Therefore the portfolio variance decreases as the conditional (dynamic) correlation between cash and futures returns increases. This correlation, according to Hegde (1982), is expected to rise during periods of increased volatility.

The conventional technique provides a more optimal hedge than the naive strategy if

$$(h_{ls} + 1)H_{22,t} - 2H_{12,t} > 0,$$

subject to $h_{ls} < 1$. If $H_{12,t} \approx H_{22,t}$ then the conventional strategy is approximately equivalent to the naive technique.

A dynamic hedging strategy is assumed to produce a more effective hedge than the naive technique if

$$-(H_{12,t} - H_{22,t})^2 < 0. \tag{2}$$

Since the left-hand side of (2) is always negative, a dynamic hedge is never less effective than the naive hedge. In the case where $H_{12,t} \approx H_{22,t}$ for all t (i.e. when the time-varying hedge ratio is approximately equal to 1 for all t) both the dynamic and naive techniques are equally effective, otherwise the time-varying hedge is closer to being optimal than the naive hedge.

A dynamic technique provides a better hedge than the minimum-variance technique if

$$H_{11,t} - \frac{H_{12,t}^2}{H_{22,t}} - H_{11,t} + 2h_{ls}H_{12,t} - h_{ls}^2 H_{22,t} < 0$$

which may be simplified to

$$-(\frac{H_{12,t}}{H_{22,t}} - h_{ls})^2 < 0 \tag{3}$$

since the quantity $H_{22,t}$ is strictly positive for all t. Since the left-hand side of (3) is always negative, the time-varying hedge is never less effective than the minimum-variance hedge. A dynamic hedge is only equally effective as the minimum-variance hedge when the quantity $\frac{H_{12,t}}{H_{22,t}}$ is approximately constant for all t.

Similarly a dynamic technique provides a more optimal hedge than any constant hedge ratio h if

$$-(\frac{H_{12,t}}{H_{22,t}} - h)^2 < 0$$

since the quantity $H_{22,t}$ is strictly positive for all t.

From the above comparisons it may be concluded that if the conditional ratio $\frac{H_{12,t}}{H_{22,t}}$ is approximately equal to 1, in some defined interval from time t_0 to time t_1, the naive hedge is effective. If, however, the ratio $\frac{H_{12,t}}{H_{22,t}}$ is approximately equal to a constant (not necessarily 1) a static hedge ratio h (not necessarily the conventional hedge) will be effective during the period from t_0 to t_1. If, however, the hedge ratios are not stable over time then dynamic techniques would be the preferred method for hedging a desired cash position. In this case, a plot of the quantity $\frac{H_{12,t}}{H_{22,t}}$ is likely to reveal nonstationarity and thus the optimal hedge will not remain constant during the life of the hedge. How effective a dynamic hedge is with respect to each of the constant hedging strategies will depend upon the extent that the optimal hedge ratio $\frac{H_{12,t}}{H_{22,t}}$ is unstable over time.

Criteria will now be introduced to compare constant hedges. This criteria will be utilised in Section 5 where a comparison between different hedging strategies is conducted. Specifically, two constant hedges will be compared. The criteria allows the hedger to determine which of the two constant hedge ratios should be applied. Assume there is the choice of two available hedge ratios, h_1 and h_2, both of which are constant. Furthermore assume that $h_1 > h_2$. Under such conditions h_1 is determined to be superior to h_2 if

$$(h_1 - h_2)((h_1 + h_2)H_{22,t} - 2H_{12,t}) < 0.$$

Therefore whether h_1 is accepted to be superior to h_2 depends upon whether the quantity $(h_1+h_2)H_{22,t}-2H_{12,t}$ is less than zero. Since $h_t = \frac{H_{12,t}}{H_{22,t}}$, this rule may be simplified to determining whether $\frac{h_1+h_2}{2} < h_t$. If so, h_1 is

accepted to be closer to optimal than h_2. This makes intuitive sense as it is known the optimal hedge ratio at any time period is h_t. In practical situations this hedge ratio cannot be altered frequently due to transaction costs. Therefore between two constant hedge ratios, h_1 and h_2, one should choose the hedge ratio that is closer to h_t for the majority of time periods.

In the next section a constant hedge ratio is proposed, calculated to be the limit of the forecasting curve of the hedge ratio (available at the present time, t). This constant hedge ratio will then be compared to the constant hedge ratio obtained via the conventional technique.

4 Forecasting hedge ratios

This section will derive a method for determining an alternative constant hedge ratio. The motive behind this is that the conventional hedge ratio may well not be the most optimal constant hedge ratio (and there may exist an alternative constant hedge ratio which, if invoked, will reduce the risk of the hedged portfolio).

A constant alternative to the conventional approach may be taken to be the limit of the forecast of conditional covariance of cash and futures returns ($H_{12,t}$) divided by the forecast of conditional variance of the futures returns ($H_{22,t}$). Define the present time by t and the resulting f-step ahead forecast as

$$\hat{\mathbf{H}}_{t+f} = E_t(\epsilon_{t+f}\epsilon'_{t+f}),$$

where E_t denotes the conditional expectation given information available at time t and $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are matrices in the MGARCH specification, then

$$\hat{\mathbf{H}}_{t+1} = E_t(\epsilon_{t+1}\epsilon'_{t+1}) = \mathbf{C}'\mathbf{C} + \mathbf{A}'\epsilon_t\epsilon'_t\mathbf{A} + \mathbf{B}'\mathbf{H}_t\mathbf{B},$$

The f-th step ahead forecast of the conditional variance can be written in the general form

$$\hat{\mathbf{H}}_{t+f} = E_t(\epsilon_{t+f}\epsilon'_{t+f}) = \sum_{k=0}^{f-1}(\mathbf{A}' + \mathbf{B}')^k * \mathbf{C}'\mathbf{C} * (\mathbf{A} + \mathbf{B})^k \tag{4}$$

$$+ (\mathbf{A}' + \mathbf{B}')^{f-1}(\mathbf{A}'\epsilon'_t\epsilon_t\mathbf{A})(\mathbf{A} + \mathbf{B})^{f-1} + (\mathbf{A}' + \mathbf{B}')^{f-1}(\mathbf{B}'\mathbf{H}_t\mathbf{B})(\mathbf{A} + \mathbf{B})^{f-1},$$

where * denotes element-by-element multiplication, e.g.

$$\begin{aligned}(\mathbf{A}' + \mathbf{B}') * \mathbf{C}'\mathbf{C} * (\mathbf{A} + \mathbf{B}) &= \mathbf{A}'\mathbf{C}'\mathbf{C}\mathbf{A} + \mathbf{B}'\mathbf{C}'\mathbf{C}\mathbf{B} \\ (\mathbf{A}' + \mathbf{B}')^2 * \mathbf{C}'\mathbf{C} * (\mathbf{A} + \mathbf{B})^2 &= \mathbf{A}'^2\mathbf{C}'\mathbf{C}\mathbf{A}^2 + \mathbf{A}'\mathbf{B}'\mathbf{C}'\mathbf{C}\mathbf{B}\mathbf{A} \\ &\quad + \mathbf{B}'\mathbf{A}'\mathbf{C}'\mathbf{C}\mathbf{A}\mathbf{B} + \mathbf{B}'^2\mathbf{C}'\mathbf{C}\mathbf{B}^2,\end{aligned}$$

since the ordering of the multiplication matters.

It can intuitively be seen that, assuming the determinant $|\mathbf{A}+\mathbf{B}|$ is less than one, the limit of the second and third terms of (4) tend to zero as f approaches infinity. Therefore as f increases, $\hat{H}_{t+f}$ will increasingly be impacted by the expression $\sum_{k=0}^{f-1}(\mathbf{A}'+\mathbf{B}')^k * \mathbf{C}'\mathbf{C} * (\mathbf{A}+\mathbf{B})^k$, although the impact of the individual terms of $(\mathbf{A}'+\mathbf{B}')^k * \mathbf{C}'\mathbf{C} * (\mathbf{A}+\mathbf{B})^k$ diminishes over the longer lag period (i.e. the limit of $(\mathbf{A}'+\mathbf{B}')^k * \mathbf{C}'\mathbf{C} * (\mathbf{A}+\mathbf{B})^k$ as k approaches infinity is equal to 0).

The next section will apply different hedging strategies to real-life data and, based on the application of the criteria developed in the previous section, will determine an appropriate hedging strategy.

5 Example

This example will demonstrate that, in the case of constant hedging, the conventional hedge is not necessarily the most optimal hedge. There may exist other constant hedge ratios which provide greater variance reduction of the hedged portfolio. Furthermore, via application of an alternate strategy (which involves the alternating of constant hedging ratios), it is seen that hedging effectiveness will significantly increase and will even be greater than that obtained using dynamic techniques.

To gain an understanding of the potential strength of the constant forecasting hedge, hedging effectiveness is compared between four types of hedges; the naive hedge, the conventional hedge, the forecasting (constant) hedge and the time-varying hedge (where this dynamic hedge is rebalanced every day). By rebalancing daily the dynamic hedge provides the optimal hedge ratio, h_t, based upon information available at t. To compare the performances of various hedging techniques, daily returns are constructed as implied by the computed hedge ratios and the variance of the returns of the constructed portfolios are calculated over the entire sample period. For the sake of completeness, the portfolio variance for the unhedged investor is also reported.

The cash and three month forward rates of the New Taiwanese Dollar are examined (vis-á-vis the United States Dollar). All data are obtained from *DATASTREAM*. The data period examined is from December 31, 1996 to December 31, 1999 inclusively. The period January 1, 1999 to December 31, 1999, is then withheld from the sample (and subsequently used for post-sample estimation) and all relevant hedge ratios are estimated over the December 31, 1996 - December 31, 1998 period. The returns (i.e. the difference of the natural logarithms) are analysed.

The time series plots of the return series are both time-varying (although

they appear to have a constant mean), with significantly more volatility in the second half of both series. Analysing the mean, median, skewness and kurtosis it may be concluded that both series are not normally distributed, the mean and median appear slightly different, the skewness is not zero, the kurtosis is not equal to 3, the Jarque-Bera and Shapiro-Wilk statistics for both series are both significant (the p-value is always 0). Strictly speaking, the Jarque-Bera and not the Shapiro-Wilk test should be invoked in this instance since the sample size is less than 2000. Therefore the null hypothesis that the distributions are normal may be rejected at the 1% level of significance. The signs for skewness are all positive, indicating the distribution for the rate of return is skewed right. Similarly, all the signs for kurtosis are positive, which indicates that the distributions of both cash and futures returns are heavy-tailed.

The attention now focusses on possible heteroscedasticity in the variances of the returns. The Box-Ljung (BL) test is used to test for such heteroscedasticity in the variances of the cash and futures returns respectively. For $p = 24$, the $BL(p)$ statistic for squared returns is significant, indicating the variance for cash returns is indeed heteroscedastic. Similarly, the Lagrange-Multiplier test is applied to both cash and futures returns. The test statistic is significant, thus rejecting the null hypothesis that the variances are homoscedastic.

A MGARCH model with an ARIMA(3,3) specification for the conditional mean and an ARIMA(1,1) specification for the conditional variance is fitted to the cash and futures returns, as this yields the lowest value of the AIC statistic (the resultant optimisation algorithm converges). All possible combinations up to an including order 4 of p and q were fitted to the conditional mean, the lowest AIC statistic was obtained when $p = q = 3$. The resultant estimates of each of the parameters in the MGARCH model are as follows;

$$\mathbf{A} = \begin{pmatrix} 0.2788 & -0.8505 \\ 0.0500 & 1.1870 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0.8989 & 0.2142 \\ -0.0033 & 0.6801 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 0.0012 & 0.0000 \\ 0.0012 & 0.0000 \end{pmatrix}.$$

The variances of both cash and futures returns are affected by shocks to the system. The shocks are permanent in nature since, extracting the coefficients from the matrices A and B, $|A_{11} + B_{11}| > 1$ and $|A_{22} + B_{22}| > 1$. The effect of these shocks on the covariance between cash and futures returns is transitory as $|A_{12} + B_{12}| < 1$ and $|A_{21} + B_{21}| < 1$. A hedge ratio of 0.8502 is obtained via the conventional hedging strategy. The forecasted hedge ratio is equal to 0.5112.

The variance of the unhedged portfolio is equal to 0.00000284. Via implementation of the naive hedge the portfolio variance is reduced by 15.8%. The hedging effectiveness via the conventional technique is equal to 37.9% and the

Table 1. Hedging Effectiveness of various hedging strategies for the New Taiwanese Dollar.

Hedging Strategy	Variance of Hedged Portfolio	Hedging Effectiveness
No Hedge	0.00000284	-
Naive	0.00000239	0.157901
Minimum Variance	0.00000176	0.379058
Constant Forecasting	0.00000125	0.561104
Time-varying	0.00000121	0.574264

effectiveness of the forecasted hedged portfolio is 56.1%. The dynamic hedge results in a reduction of 57.4% of the variance. The dynamic hedge is based upon the fitting of the $ARIMA(3,3)$ model at each post-sample period. In reality it is not known whether this model yields the lowest AIC statistic at each period (and it is impractical to determine this). If the most appropriate $ARIMA$ model is fitted at each period the hedging effectiveness of the dynamic technique will, in all likelihood, increase, e.g. when the $ARIMA(2,0)$ model is fitted at each period the hedging effectiveness value increases to 62.2%. The results for various hedging strategies are included in Table 1.

The relationship between any constant hedge ratio and the effectiveness of a hedge is shown in Figure 1. The forecasting hedge yields a hedging effectiveness value of 56.1%. The maximum variance reduction a hedger can obtain via the implementation of a constant hedge is 56.3%. This is attainable when the hedge ratio is equal to 0.54. The forecasting hedge ratio is closer to the optimal (constant) hedge ratio than the conventional hedge.

The difference in conditional variance between the conventional technique and the constant forecasting strategy for the post-sample period is plotted in Figure 2. From this plot it may be seen that even though the conditional variance of the conventional technique is less than the conditional variance of the constant forecasting strategy for the majority of time periods, the larger deviations from zero occur in the positive direction (i.e. where the conditional variance of the conventional approach is greater than that via the forecasting hedging technique). This is also seen through the mean value of the difference of conditional variances which is equal to 0.00000030. It may therefore make intuitive sense to initially adopt the conventional approach until such time the difference between conventional and forecasting strategies becomes significant so as to warrant a change in regime. This is an interesting issue to discuss but is beyond the scope of this paper.

The hedger may want to switch from the conventional technique to an im-

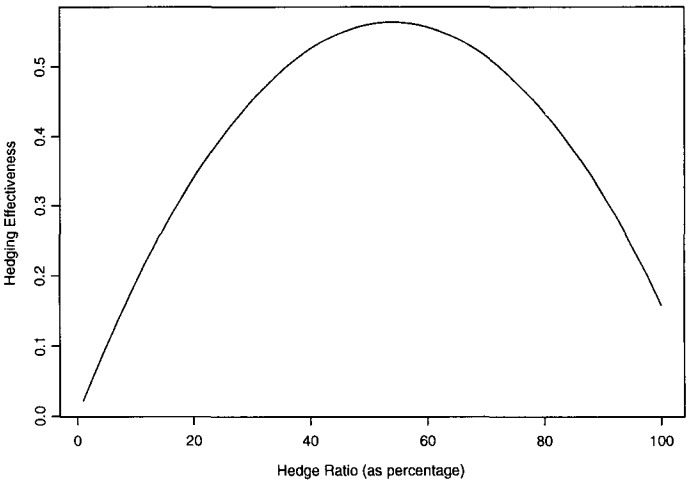


Figure 1. Relationship between any constant hedge ratio and hedging effectiveness for the New Taiwanese Dollar.

plementation of the forecasting strategy if the difference between conditional variances reaches 0.000005 (and similarly switch back to the conventional strategy once the difference becomes smaller than -0.000005). If this strategy is adopted the effectiveness of the hedge will increase to 63.7%, as the hedge will change from the conventional to the forecasting hedge strategy at approximately $t = 50$. Such effectiveness is not even obtained via the updating of the hedge at each time period. The hedger would not only minimise their risk (via the minimisation of the variance of the hedge portfolio) but also reduce the cost of implementation of the hedge since a smaller proportion of the cash portfolio is hedged in the futures market. The hedge ratio via this strategy would only have to be updated once during the post-sample period. This quasi-static hedging strategy proves to be the most optimal strategy in this instance.

6 Conclusion

Conditional hedging techniques make an over-simplified assumption of a constant covariance matrix between cash and futures returns, leading to sub-optimal hedging decisions. The implication of incorrectly assuming constant variability of returns are a loss of efficiency in terms of portfolio allocations and inappropriate decisions in terms of hedging the associated risk. There-

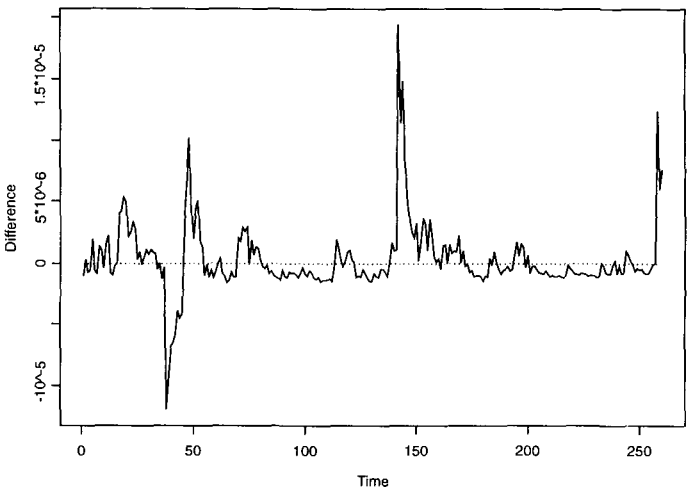


Figure 2. Conditional variance of the conventional hedge minus the conditional variance of the forecasting hedge for the New Taiwanese Dollar.

fore, in order to achieve optimal hedge ratios, allowance should be made for possible dynamic characteristics of the variances of cash and futures returns. This paper has shown that the conventional hedge does not necessarily provide the most effective constant hedge and that a constant hedge based on the limit of the forecasted hedge ratios may provide greater variance reduction. Furthermore this paper has also shown that it may be possible to obtain further variance reduction via the switching of (constant hedge) regimes from the conventional technique to the forecasting hedge strategy (or vice-versa). It may also be pointed out that as futures are derivative instruments with symmetric pay-offs, the hedging problem is far less pronounced than say, for hedging with options, as options have asymmetric pay-offs. Further research would involve the extension of the proposed methodology to cover the problem of dynamic hedging (i.e. delta hedging) with options.

REFERENCES

Baillie, R.T. and Myers, R.J. (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics* 6, 109-124.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.

Bollerslev, T. (1987). A conditional heteroscedastic time series model for

speculative prices and rates of return. *Review of Economics and Statistics* 31, 307-327.

Cecchetti, S.G., Cumby, R.E. and Figlewski, S. (1988). Estimating the optimal futures hedge. *Review of Economics and Statistics* 70, 623-630.

Gagnon, L. and Lypny, G. (1995). Hedging short-term interest risk under time-varying distributions. *The Journal of Futures Markets* 15, 767-783.

Hegde, S.P. (1982). The impact of interest rate level and volatility on the performance of interest rate hedges. *The Journal of Futures Markets* (December), 341-356.

Kroner, K.F. and Sultan, J. (1993). Time varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis* 25, 535-551.

Myers, R.J. (1991). Estimating time varying optimal hedge ratios on futures markets. *The Journal of Futures Markets* 11, 39-53.

Park, T. and Switzer, L. (1995). Bivariate GARCH estimation of the optimal hedge ratios for stock index futures: A note. *The Journal of Futures Markets* 15, 61-67.

Watt, D.G. (1997). Canadian short-term interest rates and the BAX futures market: An analysis of the impact of volatility on hedging activity and the correlation of returns between markets. *Bank of Canada Working Paper*, 97-18.

SIMULATING TRANSECTS THROUGH TWO-PHASE PARTICLES

BRIAN M. BRAY
School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW 2522, Australia
email: bmb02@uow.edu.au

The composition of a transect through a two-phase particle has a probability distribution which depends on the shape, composition and phase structure within the particle, as well as the type of randomness used to generate the transect. This paper will illustrate how the transect composition can be simulated in *Mathematica* for some simple geometric models. In particular, the liberation index suggested by Davy (1984) will be evaluated for two-dimensional models, thereby extending the results of Coleman (1991).

Keywords: geometric probability, random secant.

1 Introduction

For a population of two-phase particles with phases denoted by 0 and 1, Davy (1984a) introduced the index of liberation defined by

$$\Lambda_d = 1 - \frac{\mathcal{E}(X_0X_1)\mathcal{E}(X^2)}{\mathcal{E}(X_0X)\mathcal{E}(X_1X)}$$

where $\mathcal{E}(\cdot)$ denotes expectation, X denotes length L, area A or volume V for $d = 1$, 2 or 3 respectively, and X_i $(i = 0, 1)$ denotes the portion of X that belongs to phase i $(X_0 + X_1 = X)$. The index is symmetrical with respect to the two phases and ranges in value from 0 (when X_1/X is constant) to 1 (when the only possible values of X_1 are 0 and X). It is therefore a measure of variation of particle composition within the population. The index evaluated within lower dimensional cross-sections or transects tends to be higher than that of the original populatiion, as will be demonstrated in Section 5.

Consider a fixed two-dimensional convex body $\mathcal{K}$ with area A and perimeter B, divided into two phases by a random line. Four different types of random line will be discussed in Section 2. For any such line, the two possible allocations of phases 0 and 1 are assumed to be equally likely, and therefore A_0 and $A_1 = A - A_0$ have the same distribution.

The two-phase particle is then probed by an independent random line, resulting in a transect of length $L = L_0 + L_1$. Let α and β denote the

types of randomness of the interphase boundary line and the transect respectively. $\mathcal{E}_\alpha(\cdot)$, $\mathcal{E}_\beta(\cdot)$ and $\mathcal{E}_{\alpha,\beta}(\cdot)$ will denote expectation with respect to α-randomness of the interphase boundary line, β-randomness of the transect and joint (α,β)-randomness. Similarly $\Lambda_2(\alpha)$ and $\Lambda_1(\alpha,\beta)$ denote liberation indices with respect to the specified types of randomness.

By symmetry,

$$\mathcal{E}_\alpha(A_0A) = \mathcal{E}_\alpha(A_1A) = \frac{1}{2}A^2$$

and therefore

$$\Lambda_2(\alpha) = 1 - \frac{4\mathcal{E}_\alpha(A_0A_1)}{A^2} \tag{1}$$

In other words, $1-\Lambda_2(\alpha)$ is twice the probability that two independent uniform random points of $\mathcal{K}$ lie in different phases.

By a similar symmetry argument

$$\Lambda_1(\alpha,\beta) = 1 - \frac{4\mathcal{E}_{\alpha,\beta}(L_0L_1)}{\mathcal{E}_\beta(L^2)} \tag{2}$$

2 Random secants

The intersection of a random line with a non-empty, compact convex body $\mathcal{K}$ in an n dimensional Euclidean space is called a random secant of $\mathcal{K}$. Before any statistical properties involving random secants may be calculated, the type of randomness defining the secants must be specified. Several authors have discussed random secants, including Bate and Pillow (1947), Kendall and Moran (1963), Horowitz (1965), Kingman (1965, 1969) and Coleman (1969). Four particular types of randomness will be considered in this paper.

The diagrams following each definition show two simulated random secants in a unit square (circle), and their intersection (or not) with a random interphase boundary line.

2.1 ∂ - Randomness: surface radiator randomness

A secant of a convex body $\mathcal{K}$ is defined by a point on its surface and a direction. The point and direction are from independent uniform distributions.

2.2 λ - Randomness

A secant is defined as the straight line through two points chosen uniformly and independently in the interior of $\mathcal{K}$.

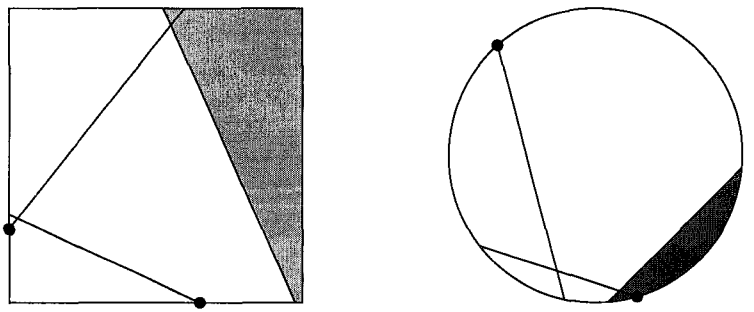


Figure 1. ∂ - random secants

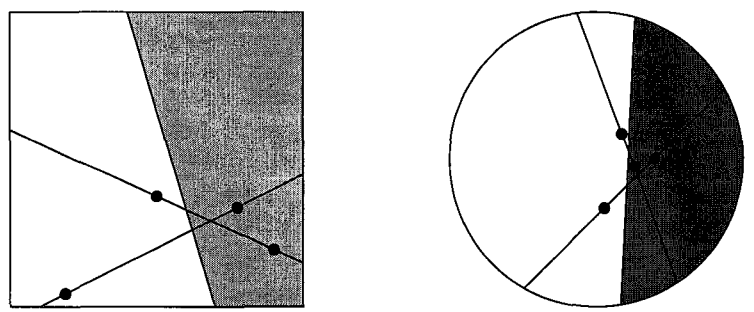


Figure 2. λ - random secants.

2.3 μ - *Randomness: mean free path randomness (Isotropic Uniform Randomness)*

A secant is defined by its *distance* from the origin and the *direction* of its normal. For a disc, the distance and direction are independently and uniformly distributed. For general $\mathcal{K}$, a secant is generated by intersection with a μ-random secant within a disc containing $\mathcal{K}$.

2.4 ν - *Randomness: interior radiator randomness (also I - randomness)*

A secant is defined by a *uniform random point in the interior* of $\mathcal{K}$ and an *isotropic direction.*

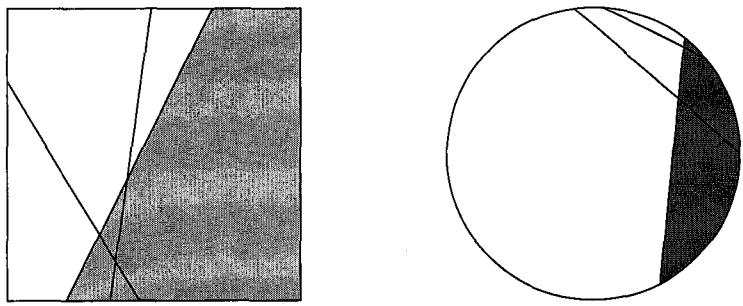


Figure 3. μ - random secants

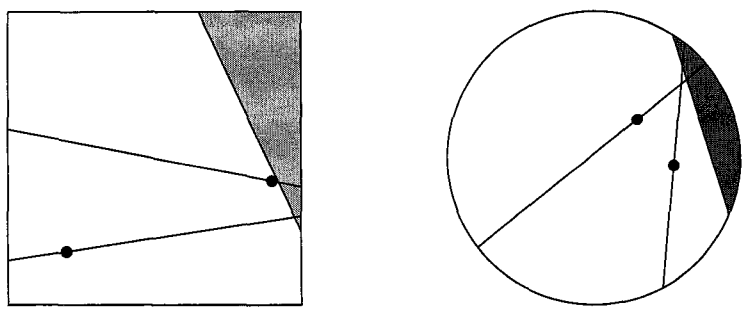


Figure 4. ν - random secants

3 Mean random lengths

Let

$$F_\alpha(l) = P\{L \leq l | \alpha - \text{randomness}\}, \qquad \alpha \in \{\nu, \mu, \lambda\}$$

and let

$$\bar{F}_\alpha = 1 - F_\alpha(l)$$

with corresponding probability density function $f_\alpha(l)$.

Kingman (1969) has shown that

$$f_\nu(l) \propto l f_\mu(l) \propto \frac{1}{l^n} f_\lambda(l) \tag{3}$$

where n is the dimensionality of $\mathcal{K}$ and the proportionality constant depends only on $\mathcal{K}$.

Coleman (1969) summarises random paths through a disc of unit radius by defining

$$f_\alpha(l) = \frac{C_\alpha l^{m_\alpha}}{\sqrt{1 - \frac{l^2}{4}}} \qquad (0 < l < 2). \tag{4}$$

The corresponding expected values, together with those for a square with unit side are shown in Table 1.

Table 1. Expected secant lengths for unit circle and unit square

α	C_α	m_α	$\mathcal{E}_\alpha(L)$ (circle)	$\mathcal{E}_\alpha(L)$ (square)
∂	$\frac{1}{\pi}$	0	$\frac{4}{\pi}$	$\frac{3\log(1+\sqrt{2}) - \sqrt{2} + 1}{\pi}$
μ	$\frac{1}{4}$	1	$\frac{\pi}{2}$	$\frac{\pi}{4}$
ν	$\frac{1}{2\pi}$	2	$\frac{16}{3\pi}$	$\frac{4\left[3\log(1+\sqrt{2}) - \sqrt{2} + 1\right]}{3\pi}$
λ	$\frac{1}{6\pi}$	4	$\frac{256}{45\pi}$	$\frac{2}{3}\left[\log(1+\sqrt{2}) + \frac{2+\sqrt{2}}{5}\right]$

4 Theoretical liberation indices

For some types of randomness, theoretical expressions for liberation indices can be derived. In the case of a μ-random interphase boundary line, the probability that two points of $\mathcal{K}$ belong to different phases is just the probability that the line segment between the two points is intersected by the interphase boundary line. For μ measure, this probability is just the ratio of the "perimeter" of the line segment to the perimeter of $\mathcal{K}$. By using the interpretation of Λ_2 immediately following (1),

$$\Lambda_2(\mu) = 1 - \frac{4\mathcal{E}(R)}{B}$$

where R is the distance between two independent uniform random points of $\mathcal{K}$. From Enns and Ehlers (1978), this can be re-expressed as

$$\Lambda_2(\mu) = 1 - \frac{2\mathcal{E}_\lambda(L)}{B} \qquad . \tag{5}$$

The corresponding numerical values are 0.424 for a circle and 0.479 for a square.

Similarly, the probability that a secant of length L intersects a μ-random interphase boundary is given by

$$P_\mu(\text{secant is composite } |L) = \frac{2L}{B} \quad .$$

Conditional upon the secant being composite, the interphase boundary point is uniformly distributed over the interval $(0, L)$. Therefore

$$\mathcal{E}_\mu(L_0L_1|\text{composite secant of length } L) = \frac{L^2}{6} \quad .$$

If we assume β-randomness for the random transect,

$$\mathcal{E}_{\mu,\beta}(L_0L_1) = \mathcal{E}_\beta\left\{\left(\frac{L^2}{6}\right)\left(\frac{2L}{B}\right)\right\} = \frac{\mathcal{E}_\beta(L^3)}{3B} \quad .$$

For a μ-random transect, Crofton's Theorem (See, e.g., Santalo, 1976) $\mathcal{E}_\mu(L^3) = 3A^2/B$ yields

$$\mathcal{E}_{\mu,\mu}(L_0L_1) = \frac{A^2}{B^2}$$

Substituting this value into (2) gives

$$\Lambda_1(\mu,\mu) = 1 - \frac{4A^2}{\mathcal{E}_\mu(L^2)B^2} = 1 - \frac{4A}{\pi\mathcal{E}_\nu(L)B} \quad .$$

Davy (1984b) proved the inequality

$$\mathcal{E}_\nu(L) \leq \frac{16}{3\pi}\sqrt{\frac{A}{\pi}}$$

so that

$$\Lambda_1(\mu,\mu) \leq 1 - \frac{3\sqrt{\pi A}}{4B} \quad . \tag{6}$$

In the case of the disc, the upper bound is achieved exactly with a numerical value of $\Lambda_1(\mu,\mu) = 0.625$. For a square, $\Lambda_1(\mu,\mu) = 0.664$ which is very close to the upper bound of 0.668. For both the circle and the square, the 1-dimensional liberation index exceeds the 2-dimensional index by large margin.

5 Simulations of the liberation index

A series of random trials using each of the above forms of randomness in a two-phase particle in the shape of a square and a circle were simulated. The average secant length $\overline{L}$ and corresponding theoretical result $\mathcal{E}_\beta(L)$ from Table 1 are tabulated in Table 2, together with the 1-dimensional Liberation Index. (Each trial was run 1000 times.)

The simulations confirmed the theoretical 1-dimensional (μ, μ) indices found in the previous section, and showed relatively minor differences between indices for the circle and square. Coleman (1991) simulated μ-random linear transects through capped spheres of fixed composition, and likewise found a large discrepancy between 1-dimensional and full dimensional indices.

Table 2. Simulated liberation indices

Randomness α, β	$\mathcal{K}$	$\overline{L}$	$\mathcal{E}_\beta(L)$	$\hat{\Lambda}_1$
∂, ∂	Square	0.71593	0.710	0.68387
∂, ∂	Circle	1.26557	1.273	0.69537
μ, μ	Square	0.78803	0.785	0.66593
μ, μ	Circle	1.57739	1.571	0.66720
ν, ν	Square	0.94580	0.946	0.64692
ν, ν	Circle	1.69300	1.698	0.64358
λ, λ	Square	1.0394	1.043	0.62361
λ, λ	Circle	1.80960	1.811	0.62207

6 Conclusion

For the simple phase structure considered in this paper, the liberation index obtained from a linear transect is much greater than the corresponding two-dimensional index, but does not vary greatly between circles and squares.

REFERENCES

Bate, A.E. and Pillow, M.E. (1947) Mean free path of sound in an auditorium. *Proc. Phys. Soc.* 59, 535-541.

Coleman, R. (1969) Random paths through convex bodies. *J. Appl. Prob.* 6, 430-444.

Coleman, R. (1991) The Davy index of mineral liberation *Acta Stereol.* 10, 103-112.

Davy, Pamela J. (1984a) Probability models for liberation. *J. Appl. Prob.* 21, 260-269.

Davy, Pamela J. (1984b) Inequalities for Moments of Secant Length. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 68, 243-246.

Enns, E.G. and Ehlers, P.F. (1978) Random paths through a convex region. *J. Appl. Prob.* 15, 144-152.

Horowitz, M. (1965) Probability of random paths across elementary geometrical shapes. *J. Appl. Prob.* 2, 169-77.

Kendall, M.G. and Moran, P.A.P. (1963) *Geometrical Probability.* London: Griffin.

Kingman, J.F.C. (1965) Mean free paths in a convex reflecting region. *J. Appl. Prob.* 2, 162-68.

Kingman, J.F.C. (1969) Random secants of a convex body. *J. Appl. Prob.* 6, 660-72.

Santalo, L.A. (1976) *Integral Geometry and Geometric Probability.* Massachusetts: Addison-Wesley.

NEW BOUNDS FOR THE THREE-POINT RULE INVOLVING THE RIEMANN-STIELTJES INTEGRAL

P. CERONE AND S. S. DRAGOMIR
School of Communications and Informatics
Victoria University of Technology
PO Box 14428, Melbourne City MC 8001, Victoria, Australia.
E-mail: pc@matilda.vu.edu.au
E-mail: sever@matilda.vu.edu.au

The solution to many problems in applied probability requires the evaluation of Riemann-Stieltjes integrals involving the convolution of cumulative distribution functions. Closed form expressions for the solution are very rare indeed. In this paper we examine the evaluation of the Stieltjes integral, which plays a fundamental role in the numerical solution of Volterra-Stieltjes integral equations that appear frequently in renewal theoretic problems. The evaluation of Stieltjes integrals on which this paper concentrates is thus, it is argued, germane to the solution of such problems. A generalised trapezoidal rule is utilised and *a priori* error bounds are determined in the current development.

Keywords: Riemann-Stieltjes integrals, Ostrowski, trapezoidal, three-point rules, bounds, Volterra-Stieltjes integral equations, renewal.

1 Introduction

The complete solution to many problems in applied probability requires evaluation of convolution integrals of the form

$$\int_{\alpha}^{\beta} F(t-x)\,dG(x) \qquad (\alpha,\beta \text{ finite or infinite}), \tag{1}$$

where F and G are cumulative distribution functions (c.d.f.) (see, Tortorella, 1990). Closed form expressions for the convolution (1) are available in a very few special cases. There is, therefore, a strong motivation for developing techniques for evaluating the convolution of c.d.f.'s numerically (see, Tortorella, 1990).

The Stieltjes integral

$$\int_{a}^{b} f(x)\,du(x), \tag{2}$$

where f and u are density functions, is also of considerable importance in problems of reliability theory (see, Boehme et al., 1991). Boehme et al. (1991)

proceed to apply the evaluation of the Stieltjes integral to the numerical solution of Volterra-Stieltjes integral equations that appear frequently in renewal theoretic problems.

The evaluation of the Volterra integral equation of renewal type

$$r(t) = h(t) + \int_0^t r(t-x)\, dF(x) \tag{3}$$

has been, and continues to be of great interest. For $h(t) = F(t)$, a life distribution, $r(t)$ is the renewal function which has a rich history with may fruitful applications in a variety of fields, including reliability, inventory management, queueing, insurance and many other industrial engineering and operations research problems. Baxter (1981) gives a review of the numerical solution of Volterra integral equations of convolution type as applied in the reliability arena. Xie (1989) evaluates integrals of the form (3) but does not provide an explicit expression for a bound on the error.

The evaluation of Stieltjes integrals on which this paper concentrates is thus, it is argued, fundamental to the solution of such problems.

For two functions $f, u : [a,b] \to \mathbf{R}$ and $x \in [a,b]$, consider the *generalised Ostrowski functional* (see, Dragomir, 2000):

$$\theta(f,u;a,x,b) := [u(b) - u(a)]\, f(x) - \int_a^b f(t)\, du(t), \tag{4}$$

where the Riemann-Stieltjes integral $\int_a^b f(t)\, du(t)$ is assumed to exist.

In Dragomir, (2001), the second author proved the following inequality

$$|\theta(f,u;a,x,b)| \leq H\left[\frac{1}{2}(b-a) + \left|x - \frac{a+b}{2}\right|\right]^r \bigvee_a^b (u) \tag{5}$$

for all $x \in [a,b]$, where $f : [a,b] \to \mathbf{R}$ is of r-H-Hölder type, u a function of bounded variation and $\bigvee_a^b (u)$ is its total variation on $[a,b]$. We recall this to mean,

$$|f(x) - f(y)| \leq H|x-y|^r \quad \text{for any} \quad x, y \in [a,b]; \tag{6}$$

$H > 0$, $r \in (0,1]$ are given. He has also shown that the constant $\frac{1}{2}$, the coefficient of $(b-a)$, is the best possible for all $r \in (0,1]$.

Dragomir (2000), by the use of a different technique, has proved the following complementary result

$$|\theta(f,u;a,x,b)| \leq H\left[(x-a)^r \bigvee_a^x (f) + (b-x)^r \bigvee_x^b (f)\right] \tag{7}$$

$$\leq \begin{cases} H\left[(x-a)^r+(b-x)^r\right]\left[\frac{1}{2}\bigvee_a^b(f)+\frac{1}{2}\left|\bigvee_a^x(f)-\bigvee_x^b(f)\right|\right], \\ H\left[(x-a)^{qr}+(b-x)^{qr}\right]^{\frac{1}{q}}\left[\left(\bigvee_a^x(f)\right)^p+\left(\bigvee_x^b(f)\right)^p\right]^{\frac{1}{p}} \\ \qquad\qquad \text{where } p>1,\ \frac{1}{p}+\frac{1}{q}=1; \\ H\left[\frac{1}{2}(b-a)+\left|x-\frac{a+b}{2}\right|\right]^r\bigvee_a^b(f), \end{cases}$$

provided f is a mapping of bounded variation and u is of Hölder type with the constant $r \in (0,1]$ and $H>0$.

Dragomir et al. (2000) have also considered another approach in approximating the Riemann-Stieltjes integral. Namely, they introduced the *generalised trapezoid functional*

$$GT(f,u;a,x,b) := [u(x)-u(a)]f(a)+[u(b)-u(x)]f(b)-\int_a^b f(t)\,du(t) \tag{8}$$

and using the identity

$$GT(f,u;a,x,b) = \int_a^b (u(t)-u(x))\,df(t) \tag{9}$$

they proved the result

$$|GT(f,u;a,x,b)| \leq H\left[\frac{1}{2}(b-a)+\left|x-\frac{a+b}{2}\right|\right]^r\bigvee_a^b(f), \tag{10}$$

provided u is of r-H–Hölder type ($H>0$, $r\in(0,1]$) and f is of bounded variation. Here the constant $\frac{1}{2}$, the coefficient of $(b-a)$, is also sharp. A partitioning of the interval of integration allowed the estimation of the error to be determined, enabling *a priori* knowledge for a desired accuracy.

In Sections 2 and 3, further bounds are obtained for the generalised trapezoid functional $GT(f;u,a,x,b)$. The work is extended to three point rules, the results of which are shown to recapture earlier Ostrowski results involving function evaluation at an interior point. The trapezoid results are also shown to be particular cases of the three point rules.

2 Results for trapezoidal type rules

We start with the following result:

Theorem 1 *Let $f : [a,b] \to \mathbf{R}$ be a function of r-H−Hölder type and $u : [a,b] \to \mathbf{R}$ a function of bounded variation on $[a,b]$. Then we have the inequality:*

$$|GT(f;u,a,x,b)| \le H\left[(x-a)^r \bigvee_a^x (u) + (b-x)^r \bigvee_x^b (u)\right] \tag{11}$$

$$\le \begin{cases} H \bigvee_a^b (u) \left[\frac{1}{2} + \left|\frac{x - \frac{a+b}{2}}{b-a}\right|\right]^r (b-a)^r, \\ H\left[\left(\bigvee_a^x (u)\right)^\beta + \left(\bigvee_x^b (u)\right)^\beta\right]^{\frac{1}{\beta}} \left[\left(\frac{x-a}{b-a}\right)^{\alpha r} + \left(\frac{b-x}{b-a}\right)^{\alpha r}\right]^{\frac{1}{\alpha}} (b-a)^r \\ \qquad\qquad \text{where } \alpha > 1,\ \frac{1}{\alpha} + \frac{1}{\beta} = 1; \\ H\left[\frac{1}{2}\bigvee_a^b (u) + \frac{1}{2}\left|\bigvee_a^x (u) - \bigvee_x^b (u)\right|\right]\left[\left(\frac{x-a}{b-a}\right)^r + \left(\frac{b-x}{b-a}\right)^r\right](b-a)^r, \end{cases}$$

for all $x \in [a,b]$.

Proof: It is easy to see that

$$\begin{aligned} -GT(f;u,a,x,b) &= \int_a^x f(t)\,du(t) + \int_x^b f(t)\,du(t) \\ &\quad - f(b)\int_x^b du(t) - f(a)\int_a^x du(t) \\ &= \int_a^x [f(t) - f(a)]\,du(t) + \int_x^b [f(t) - f(b)]\,du(t), \end{aligned} \tag{12}$$

for all $x \in [a,b]$.

Taking the modulus in (12), we obtain

$$|GT(f;u,a,x,b)| \le \left|\int_a^x [f(t) - f(a)]\,du(t)\right| + \left|\int_x^b [f(t) - f(b)]\,du(t)\right|. \tag{13}$$

We know that if $m : [a,b] \to \mathbf{R}$ is continuous and $n : [a,b] \to \mathbf{R}$ is of bounded variation, then the Riemann-Stieltjes integral $\int_a^b m(t)\,dn(t)$ exists and we have the inequality:

$$\left|\int_a^b m(t)\,dn(t)\right| \le \sup_{t\in[a,b]} |m(t)| \bigvee_a^b (n). \tag{14}$$

Now, using the inequality (14) and the fact that f is r-H−Hölder, we may state that

$$\left|\int_a^x [f(t) - f(a)]\,du(t)\right| \le \sup_{t\in[a,x]} |f(t) - f(a)| \bigvee_a^x (u) \le H(x-a)^r \bigvee_a^x (u),$$

and, in a similar way,

$$\left| \int_x^b [f(t) - f(b)]\, du(t) \right| \leq H (b-x)^r \bigvee_x^b (u).$$

Thus, utilising (13), we deduce the first inequality in (11).

Denote by $J(x)$ the expression,

$$J(x) := H \left[(x-a)^r \bigvee_a^x (u) + (b-x)^r \bigvee_x^b (u) \right], \quad x \in [a,b].$$

We observe that

$$J(x) \leq H \max \{(x-a)^r, (b-x)^r\} \left[\bigvee_a^x (u) + \bigvee_x^b (u) \right]$$

$$= H \left[\max \{x-a, b-x\} \right]^r \bigvee_a^b (u) = H \left[\frac{1}{2} + \left| \frac{x - \frac{a+b}{2}}{b-a} \right| \right]^r (b-a)^r \bigvee_a^b (u)$$

proving the first part of the second inequality in (11).

Using Hölder's discrete inequality for $\alpha > 1$, $\frac{1}{\alpha} + \frac{1}{\beta} = 1$, we may state that

$$J(x) \leq H \left[(x-a)^{\alpha r} + (b-x)^{\alpha r} \right]^{\frac{1}{\alpha}} \left[\left(\bigvee_a^x (u) \right)^\beta + \left(\bigvee_x^b (u) \right)^\beta \right]^{\frac{1}{\beta}}$$

$$= H \left[\left(\frac{x-a}{b-a} \right)^{\alpha r} + \left(\frac{b-x}{b-a} \right)^{\alpha r} \right]^{\frac{1}{\alpha}} \left[\left(\bigvee_a^x (u) \right)^\beta + \left(\bigvee_x^b (u) \right)^\beta \right]^{\frac{1}{\beta}} (b-a)^r,$$

proving the second part of the second inequality in (11).

Finally, we observe that

$$J(x) \leq H \max \left\{ \bigvee_a^x (u), \bigvee_x^b (u) \right\} \left[(x-a)^r + (b-x)^r \right]$$

$$= H \left[\frac{1}{2} \bigvee_a^b (u) + \frac{1}{2} \left| \bigvee_a^x (u) - \bigvee_x^b (u) \right| \right] \left[\left(\frac{x-a}{b-a} \right)^r + \left(\frac{b-x}{b-a} \right)^r \right] (b-a)^r$$

and the last part of (11) is proved.

Corollary 1 *If u is as in Theorem 1, but f is $L-$Lipschitzian, then,*

$$|GT(f,u;a,x,b)| \leq L \left[(x-a) \bigvee_a^x (u) + (b-x) \bigvee_x^b (u) \right]. \tag{15}$$

Remark 1 *If u is as in Theorem 1, but f is $L-$Lipschitzian, which is effectively 1-$L-$Hölder type satisfying (6), then simplification of the result (11) occurs.*

If in (11) we take $x = \frac{a+b}{2}$, then we get the tightest bounds,

$$\left|GT\left(f,u;a,\frac{a+b}{2},b\right)\right| \leq \frac{1}{2^r} H \bigvee_a^b (u)\,(b-a)^r \tag{16}$$

for f of r-$H-$Hölder and, if f is $L-$Lipschitzian,

$$\left|GT\left(f,u;a,\frac{a+b}{2},b\right)\right| \leq \frac{1}{2} L\,(b-a) \bigvee_a^b (u)\,. \tag{17}$$

Another particular case of interest may be obtained from Theorem 1. If $x_0 \in [a,b]$ is a point for which we have

$$\bigvee_a^{x_0} (u) = \bigvee_{x_0}^b (u)\,, \tag{18}$$

then we have from (11)

$$|GT\,(f,u;a,x_0,b)| \leq \frac{1}{2} H \bigvee_a^b (u) \left[\left(\frac{x_0-a}{b-a}\right)^r + \left(\frac{b-x_0}{b-a}\right)^r\right](b-a)^r \tag{19}$$

and, if f is $L-$Lipschitzian, then

$$|GT\,(f,u;a,x_0,b)| \leq \frac{1}{2} L \bigvee_a^b (u)\,(b-a)\,. \tag{20}$$

Now, if we assume that $u : [a,b] \to \mathbf{R}$, $u\,(t) = \int_a^t g\,(s)\,ds$ where $g : [a,b] \to \mathbf{R}$ is continuous, then

$$\begin{aligned} \int_a^b f\,(t)\,du\,(t) &= \int_a^b f\,(t)\,g\,(t)\,dt, \\ -GT\,(f,u;a,x,b) &= \int_a^b f\,(t)\,g\,(t)\,dt - \left[f\,(a)\int_a^x g\,(s)\,ds + f\,(b)\int_x^b g\,(s)\,ds\right] \\ &=: \zeta\,(f,g;a,x,b)\,. \end{aligned} \tag{21}$$

Also, we have

$$\bigvee_a^b (u) = \int_a^b |u'\,(s)|\,ds = \int_a^b |g\,(s)|\,ds.$$

We may now state the following corollary in approximating the integral of the product of two functions.

Corollary 2 *Assume that* $f:[a,b]\to\mathbf{R}$ *is as in Theorem 1 and* $g:[a,b]\to\mathbf{R}$ *is continuous on* $[a,b]$. *Then*

$$|\zeta(f,g;a,x,b)|\leq\left[(x-a)^r\int_a^x g(s)\,ds+(b-x)^r\int_x^b g(s)\,ds\right] \tag{22}$$

$$\leq\begin{cases} H\int_a^b|g(s)|\,ds\left[\frac{1}{2}+\left|\frac{x-\frac{a+b}{2}}{b-a}\right|\right]^r(b-a)^r, \\ H\left[\left(\int_a^x|g(s)|\,ds\right)^\beta+\left(\int_x^b|g(s)|\,ds\right)^\beta\right]^{\frac{1}{\beta}}\left[\left(\frac{x-a}{b-a}\right)^{\alpha r}+\left(\frac{b-x}{b-a}\right)^{\alpha r}\right]^{\frac{1}{\alpha}}(b-a)^r, \\ H\left[\frac{1}{2}\int_a^b|g(s)|\,ds+\frac{1}{2}\left|\int_a^x|g(s)|\,ds-\int_x^b|g(s)|\,ds\right|\right] \\ \qquad\times\left[\left(\frac{x-a}{b-a}\right)^r+\left(\frac{b-x}{b-a}\right)^r\right](b-a)^r, \end{cases}$$

for all $x\in[a,b]$.

Remark 2 *If* f *is* L*–Lipschitzian, then replacing* H *by* L *and* r *by* 1 *produces some simplification, especially for the last inequality in (22).*

If $x=\frac{a+b}{2}$, then, obviously from (21)

$$\left|\zeta\left(f,g;a,\frac{a+b}{2},b\right)\right|=\left|\int_a^b f(t)g(t)\,dt-f(a)\int_a^{\frac{a+b}{2}}g(s)\,ds-f(b)\int_{\frac{a+b}{2}}^b g(s)\,ds\right|$$

and so from (22) of Corollary 2 we may write

$$\left|\zeta\left(f,g;a,\frac{a+b}{2},b\right)\right|\leq\frac{1}{2^r}H(b-a)^r\int_a^b|g(s)|\,ds, \tag{23}$$

which, for an L–Lipschitzian function f, specialises to

$$\left|\zeta\left(f,g;a,\frac{a+b}{2},b\right)\right|\leq\frac{1}{2}L(b-a)\int_a^b|g(s)|\,ds. \tag{24}$$

Further, if $x_0\in[a,b]$ is a point for which $\int_a^{x_0}|g(s)|\,ds=\int_{x_0}^b|g(s)|\,ds$, then we have

$$|\zeta(f,g;a,x_0,b)|=\left|\int_a^b f(t)\,g(t)\,dt-f(a)\int_a^{x_0}g(s)\,ds-f(b)\int_{x_0}^b g(s)\,ds\right|$$

$$\leq\frac{1}{2}H\int_a^b|g(s)|\,ds\cdot\left[\left(\frac{x_0-a}{b-a}\right)^r+\left(\frac{b-x_0}{b-a}\right)^r\right](b-a)^r. \tag{25}$$

In particular, if f is $L-$Lipschitzian and x_0 is as above, then

$$|\zeta(f,g;a,x_0,b)| \leq \frac{1}{2}L(b-a)\int_a^b |g(s)|\,ds. \tag{26}$$

3 Results for three point rules

In this section it is demonstrated that the generalised trapezoid functional $GT(f,u;a,x,b)$ defined by (8) may be utilised to obtain three point rules which involve function evaluation at the ends of the interval and at an interior point. Let the *generalised Simpson functional* be defined by

$$GS(f,u;a,x,z,y,b) = [u(x)-u(a)]f(a)+[u(y)-u(x)]f(z) \tag{27}$$
$$+[u(b)-u(y)]f(b)-\int_a^b f(t)\,du(t),$$

where $a \leq x \leq z \leq y \leq b$. We note that if $u(t)=t$, $x=\frac{5a+b}{6}$, $y=\frac{a+5b}{6}$ and $z=\frac{a+b}{2}$, the traditional Simpson rule results.

It may be easily shown that

$$GS(f,u;a,x,z,y,b) = GT(f,u;a,x,z)+GT(f,u;z,y,b)$$

and hence since $GT(f,u;a,x,b)$ satisfies identities (9) and (12) we have that

$$GS(f,u;a,x,z,y,b) = \int_a^z (u(t)-u(x))\,df(t)+\int_z^b (u(t)-u(y))\,df(t) \tag{28}$$

and

$$-GS(f,u;a,x,z,y,b) = \int_a^x (f(t)-f(a))\,du(t)+\int_x^z (f(t)-f(z))\,du(t)$$
$$+\int_z^y (f(t)-f(z))\,du(t)+\int_y^b (f(t)-f(b))\,du(t). \tag{29}$$

The following two theorems are obtained by utilising identities (28) and (29) respectively.

Theorem 2 *Let $f:[a,b]\to\mathbf{R}$ be a function of bounded variation on $[a,b]$ and $u:[a,b]\to\mathbf{R}$ be of r-$H-$Hölder type. We then have*

$$|GS(f,u;a,x,z,y,b)| \leq H\left[M^r(x-a,z-x)\bigvee_a^z(f)+M^r(y-z,b-y)\bigvee_z^b(f)\right]$$

$$\leq H[\max\{x-a,z-x,y-z,b-y\}]^r\bigvee_a^b(f), \tag{30}$$

where $M(A,B)=\frac{A+B}{2}+\frac{|A-B|}{2}$.

Proof: We have from (28) using the triangle inequality

$$|GS(f,u;a,x,z,y,b)|\leq \sup_{t\in[a,z]}|u(t)-u(x)|\bigvee_a^z(f)+\sup_{t\in[z,b]}|u(t)-u(y)|\bigvee_z^b(f)$$

$$\leq H\sup_{t\in[a,z]}|t-x|^r\bigvee_a^z(f)+H\sup_{t\in[z,b]}|t-y|^r\bigvee_z^b(f)$$

since u is r-$H-$Hölder and so satisfies (6).

Now, using the fact that for $w\in[c,d]$, $\sup_{t\in[c,d]}|t-w|^r=[\max\{w-c,d-w\}]^r$ gives the first result in (30) as stated. The coarser bound is obvious.

Theorem 3 *Let* $f:[a,b]\to\mathbf{R}$ *be a* r-$H-$*Hölder type and* $u:[a,b]\to\mathbf{R}$ *be a function of bounded variation. Then we have*

$$|GS(f,u;a,x,z,y,b)| \tag{31}$$

$$\leq H\left[(x-a)^r\bigvee_a^x(u)+(z-x)^r\bigvee_x^z(u)+(y-z)^r\bigvee_z^y(u)+(b-y)^r\bigvee_y^b(u)\right]$$

$$\leq\begin{cases} H[\max\{x-a,z-x,y-z,b-y\}]^r\bigvee_a^b(u); \\ H[(x-a)^{\alpha r}+(z-x)^{\alpha r}+(y-z)^{\alpha r}+(b-y)^{\alpha r}]^{\frac{1}{\alpha}}\left[\left(\bigvee_a^x(u)\right)^\beta\right. \\ \left.+\left(\bigvee_x^z(u)\right)^\beta+\left(\bigvee_z^y(u)\right)^\beta+\left(\bigvee_y^b(u)\right)^\beta\right]^{\frac{1}{\beta}} \quad \text{where } \alpha>1,\ \frac{1}{\alpha}+\frac{1}{\beta}=1; \\ H[(x-a)^r+(z-x)^r+(y-z)^r+(b-y)^r]\max\left\{\bigvee_a^x(u),\bigvee_x^z(u),\bigvee_z^y(u),\bigvee_y^b(u)\right\}. \end{cases}$$

Proof: The proof follows that of Theorem 1 closely upon utilising (29) and the triangle inequality.

Remark 3 *We note that if* x,z *and* y *are taken at their respective midpoints, then from (30) and (31) the tightest bounds are obtained*

$$\left|GS\left(f,u;a,\frac{3a+b}{2},\frac{a+b}{2},\frac{a+3b}{2},b\right)\right|$$

$$\leq\begin{cases} H\cdot\left(\frac{b-a}{4}\right)^r\bigvee_a^b(f),\ f \text{ of bounded variation, } u \text{ of } r-H-\text{Hölder} \\ H\cdot\left(\frac{b-a}{4}\right)^r\bigvee_a^b(u),\ \text{vice versa.} \end{cases}$$

These bounds may be noticed to be sharper than the tightest bounds from the Ostrowski and trapezoidal results from (7), (8), (10) and (11)

$$\left|\theta\left(f,u;a,\frac{a+b}{2},b\right)\right| = \left|GT\left(f,u;a,\frac{a+b}{2},b\right)\right| \leq \begin{cases} H\cdot\left(\frac{b-a}{2}\right)^r \bigvee_a^b(f), \\ H\cdot\left(\frac{b-a}{2}\right)^r \bigvee_a^b(u). \end{cases}$$

Remark 4 *If we take* $x = y = z$, *then we reproduce the results for the trapezoidal type rules of Section 2. On the other hand, taking* $x = a$ *and* $y = b$ *recaptures the Ostrowski type results (5) and (7). That is,* $GS(f,u;a,z,z,z,b) = GT(f,u;a,z,b)$ *and* $GS(f,u;a,a,z,b,b) = \theta(f,u;a,z,b)$.

Finally, the above results may be utilised to obtain composite rules by partitioning the interval $[a,b]$ so that $a = \xi_0 < \xi_1 < \ldots < \xi_{n-1} < \xi_n = b$ with $x_i \in [\xi_i, \xi_{i+1}]$ $i = 0,\ldots,n-1$. The bounds for $GT(f,u;\xi_i,x_i,\xi_{i+1})$ are then used to give *a priori* error estimates over the entire interval.

REFERENCES

Baxter, L.A. (1981). Some remarks on numerical convolution. *Commun. Statist. - Simula. Computa.* B 10, 281-288.

Boehme, T.K., Preuss, W., Van der Wall, V. (1991). On a simple numerical method for computing Stieltjes integrals in reliability theory. *Probab. Engrg. Inform. Sci.* 5, 113-128.

Dragomir, S.S. (2001). On the Ostrowski inequality for the Riemann-Stieltjes integral $\int_a^b f(u)\,du(t)$, where f is of Hölder type and u is of bounded variation and applications. *J. KSIAM* 5, 35-45.

Dragomir, S.S. (2000). On the Ostrowski's inequality for Riemann-Stieltjes integral and applications. *Korean J. Comput. & Appl. Math.* 7, 477-485.

Dragomir, S.S., Buşe, C., Boldea, M.V., and Braescu, L. (2001). A generalisation of the trapezoidal rule for the Riemann-Stieltjes integral and applications. *Nonlinear Analysis Forum* 6, 337-351. *RGMIA Res. Rep. Coll.* 3, Article 2. http://rgmia.vu.edu.au/v3n4.html

Tortorella, M (1990). Closed Newton-Cotes quadrature rules for Stieltjes integrals and numerical convolution of life distributions. *SIAM J. Sci. Statist. Comput.* 11, 732-748.

Xie, M. (1989). On the solution of renewal type integral equations. *Commun. Statist. - Simula* 18, 281-293.

SOME COMPARISONS OF CLUSTERING AND CLASSIFICATION TECHNIQUES APPLIED TO TRANSCRIPTIONAL PROFILING DATA

SUSMITA DATTA
Department of Mathematics and Statistics
Georgia State University
Atlanta, GA, USA
E-mail: sdatta@mathstat.gsu.edu

JONATHAN ARNOLD
Department of Genetics
University of Georgia
Athens, GA 30602, USA

This paper gives some comparisons of statistical clustering and classification methods applied to transcriptional profiling data obtained from a microarray experiment. A number of such techniques are illustrated using a well known data set on the yeast genome.

Keywords: transcriptional profiling, yeast genome, sporulation, clustering, classification.

1 Introduction

In the new millennium, the challenge is using sequenced genomes to predict how living systems function. Genomics is now a data-driven science in which the complete genetic blueprint enables all RNA and protein levels in the cell (i.e. RNA and protein profiles) to be determined in microbial systems (DeRisi, Iyer and Brown, 1997). The promise of genomics is to integrate all of this information about DNA, RNA, and proteins to understand complex traits like development, sex, biological clocks, RNA, and pathogenicity in microbial systems initially. The challenge of genomics is to integrate all of the information about DNA, RNA, and proteins on a cell to summarize and predict these complex traits and to compute emergent properties of microbial systems from chemical reaction network models of DNA, RNA, and proteins (Bhalla and Iyengar, 1999). These models will be identified by fitting them to the observed RNA and protein profiles of microbial species with the result that genomics is shifted from data-driven discovery to hypothesis-driven science.

One of the keys to this transformation of genomics is developing new statistical models and tools for the analysis of transcriptional profiling data (Eisen *et al.*, 1998). In microarray studies, the expression levels of thousands of genes are studied simultaneously in terms of the abundance of the mRNA level during transcription at various time points (as a biological process progresses). The resulting longitudinal data for a gene will be referred to as its expression profile or transcriptional profile. Transcriptional profiling data will serve two roles. One, complex living systems have many components that are hierarchically arranged and many interactions between these components and environmental inputs. Transcriptional profiling data will be key to identifying topologies of reaction networks and regulatory hierarchies structuring an organism's adaptive response or reaction norm (Pilpel, Sudarsanam and Church, 2001). For example, information flow in the cell is hierarchically arranged along the pathway of the Central Dogma: DNA to RNA to protein. An environmental signal like a sugar may signal the cell to turn on certain genes, which produce RNA, which in turn produce certain proteins (like those on the glycolytic pathway and Kreb Cycle) to metabolize glucose, and some other proteins which may activate/repress other genes involved in glucose utilization or the utilization of alternate sugars. What genes come on over time as indicated by their RNA profiles gives clues to their function and with whom they and their products interact (DeRisi *et al.*, 1997). So, the RNA profiles become a useful tool to infer the topology of metabolic networks.

Secondly, in some reaction networks sufficient genetic and biochemical information has accumulated that an initial topology for the reaction network is available like with glucose utilization (DeRisi et al., 1997). Other examples include the lac operon (Jacob and Monod, 1961), trp operon (Yanofsky and Kolter, 1982), GAL cluster (Johnston, 1987), qa cluster (Geever et al., 1989), lytic-lysogenic cycles of phage λ (Ptashne, 1992), cell cycle (Sveiczer et al., 2000), biological clock (Lee et al., 2000) and others. In this setting the RNA profile becomes the data input for fitting and testing formal models describing a particular reaction network. RNA and protein levels can be measured after perturbing the system in varied ways including knocking out components of the network (genes or proteins) or by altering the environmental signal. In this setting the RNA transcriptional profiles are part of the measured response of the cell, which the biochemical network must correctly predict, in much the same way as electronic circuits must match specified outputs in their design given specified inputs.

The purpose of this paper is to present some comparisons of a number of well known statistical clustering and classification techniques in grouping

genes by functional roles utilizing RNA profiles. This clustering or classification of genes by their similarity of their transcriptional profiles in turn can be used to validate a reaction network in hypothesis-driven genomics.

2 Microarrays

Several thousand genes are arrayed robotically on glass slides or nylon membranes on a precise grid and intensities are scanned on and off the grid under one experimental condition in which the treatment mRNAs are compared to the control mRNAs. Minimally there are four intensities per gene (red or green on or off the grid). In the case of the yeast experiments 6118 genes are arrayed, and 4 intensities per gene are recorded (http://cmgm.stanford.edu/pbrown). The 2 intensities off the grid are used to "adjust" the 2 signal intensities on the grid. For example, the off-grid intensities might be subtracted from the on-grid intensities. Then the adjusted treatment intensity is "normalized" with respect to the adjusted control intensity and transformed. For example adjusted treatment intensity (red) might be divided by the adjusted control intensity (green) and then log of the ratio taken. This final transformed ratio is measured then on all of the arrayed genes. In the yeast experiments over 6000 transformed ratios of adjusted intensities are available under different experimental conditions during the diauxic shift or sporulation. Each transformed ratio of adjusted signal intensities is a measurement of the relative expression of a particular gene.

The whole experiment is thus repeated under different treatment conditions. For example, the final data matrix is then G x C in dimensions, where G is the number of arrayed genes and C, the number of treatment conditions. In the sporulation example, there were at least 7 time points (conditions).

The process of statistical analysis begins with visualizing and manipulating the raw data matrices or more typically with manipulations of the final single G x C data matrix with the transformed and normalized ratios of expression for each gene. Both the genes (rows) and treatment conditions (columns) can be permuted to highlight the similarities and differences in the patterns of gene expression (Ross-McDonald et al., 1999) or for displaying the effects of independent variables like time since the diauxic shift or position during the cell cycle. The G x C data matrix itself becomes a data-rich vehicle for the display of all subsequent analyses (Tufte, 1983).

An excellent account on data normalization that should be done to remove systematic bias prior to any statistical analysis is given in Yang et al. (2001).

3 Sporulation data

We consider a well known experiment on the yeast genome and the resulting microarray data. The experiment (Chu et al., 1998) was designed to study the changing expression patterns of nearly all genes during meiosis and sporulation in S. cerevisiae (budding yeast). The meiotic cell cycle is fundamental to most eukaryotes in which diploid cells give rise to haploid germ cells (spores in the case of yeast), and this example illustrates how transcriptional profiling illuminates an emergent property of a living system, reproduction.

About 6 thousand yeast genes (nearly 97% of all known or predicted genes) were arrayed on glass slides and used to assay the levels of mRNAs at various time points in the biological process of sporulation. Changes in the mRNA levels from each genes were measured at seven time points ($t = 0$, 0.5, 2, 5, 7, 9, 11.5 hours).

Over 1000 genes show significant changes in mRNA levels over time. About half of these gens were expressed and the rest were depressed during sporulation. There is a striking pattern to the expression profiles of these genes during the cell cycle. Chu et al. (1998) grouped the positively expressed genes into 7 clusters in an ad hoc way on the basis of similarity of their expression profiles to genes known to be involved in sporulation.

These 7 clusters give clues to genes function in sporulation and which genes work together. One cluster of 158 "Middle genes" during the middle induction of the cell cycle responded like target genes under the regulatory control of a master regulator gene, Ndt80. It was confirmed that 70% of these 158 target genes had at least one upstream DNA sequence that acts as the switch by which Ndt80 turns on a target gene. Additional experiments were done to confirm that when Ndt80 was experimentally turned on at the wrong time, then two-thirds of the target genes were turned on at the wrong time as well. In this way transcriptional profiling was used as a means to specify part of the circuit controlling sporulation. A formal circuit is under construction for the cell cycle in yeast (Sveiczer *et al.*, 2000).

Some clues to how genes function and their role in the circuit were elucidated in other stages of sporulation. As an example during mid-late induction, 61 genes came on. Over a third of these genes had the Ndt80 on/off switch, suggesting the formal possibility that a repressor gene may be present delaying the expression of some genes. A formal test of this hypothesis can be performed by constructing a circuit with such a repressor present and by additional perturbation experiments to the circuit to identify such a repressor.

Prior to the completion of these transcriptional profiling only 50 genes were implicated in sporulation. Now there are over 500 genes implicated in

controlling the complex trait of sporulation. In addition the global pattern of expression confirmed the presence of two major stages, meiotic prophase in which recombination operates and meiotic division and gamete morphogenesis in which genes involved in division and spore formation operate.

4 Comparison of clustering methods

Even in well studied model systems like the yeast S. cerevisiae or bacterium Escherichia coli (found in our waste and sewage) approximately half of the genes are of unknown function. If genes of unknown function can be associated with genes of known function, then clues as to the roles of the unknown genes are the result. It is, therefore, desirable to exploit available tools for clustering from numerical taxonomy and statistics (Sokal and Sneath, 1963, and Hartigan, 1975).

Initial efforts have utilized standard clustering tools like UPGMA (Unweighted pair group method using arithmetic averages) and K-Means to reorder the transcriptional profiling data matrix by rows so that genes with similar profiles appear next to each other (Eisen *et al.*, 1998, Ross-McDonald *et al.*, 1999). These two examples fall into two classes of clustering algorithms, "distance-based" clustering tools like UPGMA that use the data matrix to compute distances between cases and "character-based" clustering tools that utilize the data matrix more directly. In either approach the results can be displayed by overlaying the clusters on the transcriptional profiling data matrix or some lower dimensional representation of the data matrix like the principal components associated with the transcriptional profiling data matrix.

We reanalyze this data set (sporulation) using a number of existing statistical clustering techniques. Instead of clustering only the positively expressed genes (Chu *et al.*, 1998) we cluster all differentially expressed (positively or negatively) genes satisfying the same variation filter as in Chu *et al.* (1998). Overall, we select ten clusters which is larger than what Chu *et al.* (1998) had used. These extra clusters were used to accommodate some additional expression profiles resulting from the negatively expressed genes.

The following clustering techniques were considered. S-plus implementation of all these techniques are available in the library MASS and are described in Venables and Ripley (1999).

(i) Hierarchical Clustering with Correlation: This is perhaps the most commonly used clustering techniques with microarray data and uses a distance or dissimilarity measure. We choose one minus absolute correlation between two genes for this purpose.

(ii) Clustering by K-Means: Usually another clustering algorithm, such as

the one above is run to determine the cluster centers to be used in the K-means algorithm. The algorithm then assigns the observations into various clusters in order to minimize the within-class sum of squares from these centers.

(iii) Model Based Clustering: The idea behind model based clustering is to regard the data as coming from a mixture (say of multivariate normal) distribution. The unknown group levels are then selected by the methods like maximum likelihood.

(iv) Fanny: It uses fuzzy logic and produces a probability vector for each observation. A hard cluster is determined by assigning an observation to a group which has the highest probability. (See Kaufman and Rousseeuw (1990) for the details). Typically, relatively fewer clusters are produced by this method. Like distance based methods, one has a choice of using a general dissimilarity measure. We have used the L_1 distance which is more robust than the Euclidean distance.

(v) Diana: It is a divisive clustering method where initially all the observations are clustered together. Subsequently, the bigger groups are broken down into smaller groups so that genes with larger distance or dissimilarity are placed in different clusters.

The results are shown in Figure 1. These plots are useful for a visual comparison. As expected, there are some differences in the results of the various algorithms. Overall, K-means and Diana seem to be most effective in achieving distinct class boundaries.

5 Classification

In the language of machine learning, classification is referred to as "supervised learning" since one has a "training" or representative set with known group membership.

One of the major challenges of any genome project (Bennett and Arnold, 2001) is the annotation of sequence data. As a variety of sequence, RNA, and protein information accumulate on model systems, one is faced with the daunting task of annotating all of the 35,000 human genes in this very large periodic table of life (Venter *et al.*, 2001). How do we organize the table? As the wealth of sequence information has come online, a variety of classification systems have arisen to classify genes by function. One of these is the EGAD system (http://www.tigr.org/tdb/egad/egad.shtml), and the other is the MIPS system (http://mips.gsf.de/mips/sitemap/). The MIPS system is more detailed. One approach to assigning genes to such a classification is on the basis of sequence similarity (Altschul *et al.*, 1990, 1997), but transcriptional profiling information provide another avenue of classification.

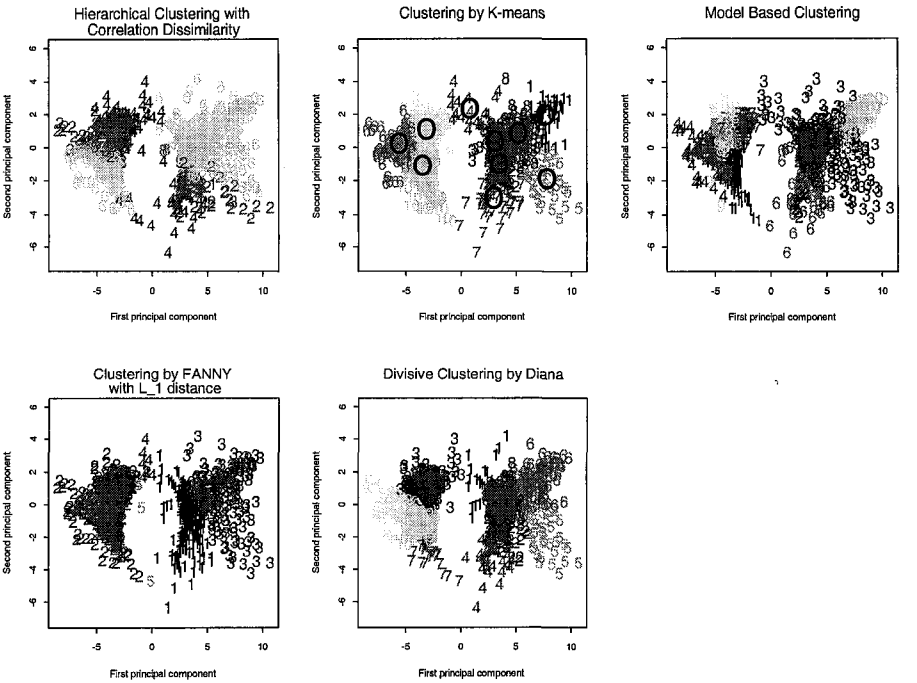


Figure 1. Clustering of all differentially expressed genes during sporulation.

A classification system requires the availability of prior information to train the "classification function". One approach is to relate sequences in one target organism to those in a model organism like E. coli, S. cerevisiae, C. elegans, or D. melanogaster. When sequence comes available in the target organism, similarity searches against these model systems can then yield an initial classification of genes within the target organism. Second, functional studies in the target organism may provide the prior information. Typically this approach leads to the classification of at most 50% of the genes in the target organism (Prade *et al.*, 2000). With a training set in hand, then one can use transcriptional profiling data as a classification tool to group the remaining genes of unknown function.

Alternatively, one may be interested in some particular process like fermentation and or sporulation and have a small training set based on

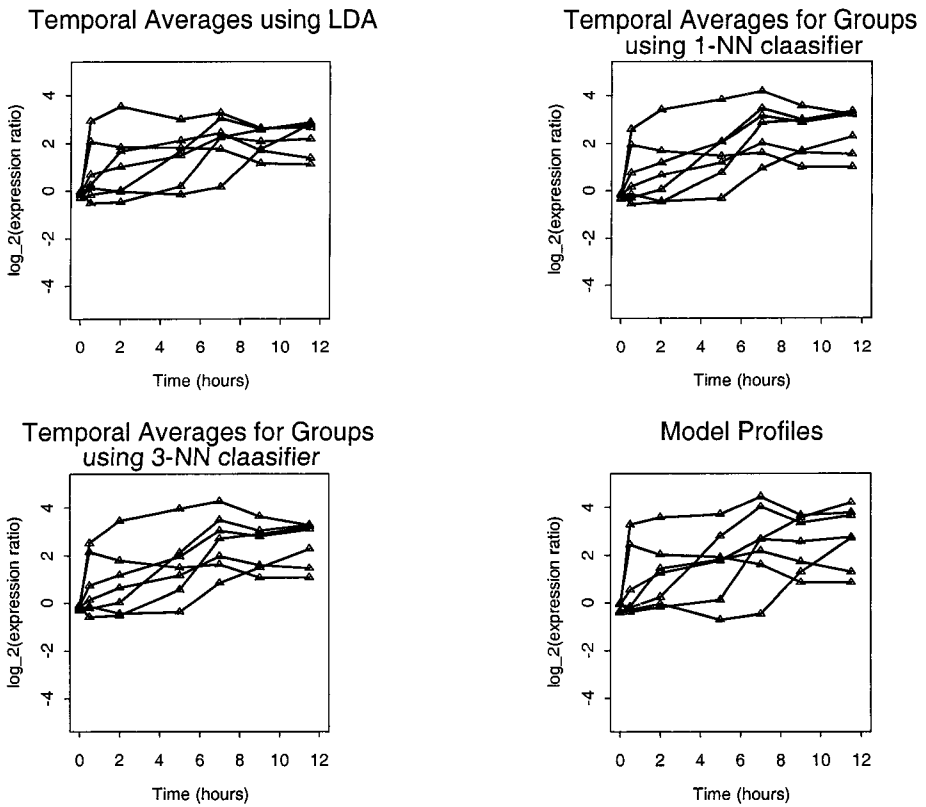


Figure 2. Classification of positively expresses yeast genes during sporulation and their average temporal pattern.

prior experiments or have created an initial clustering of genes by prior transcriptional profiling experiments or by means of experiments distinct from transcriptional profiling. The goal is then to classify genes by function not in the training set.

We investigate how standard statistical methods such as the linear discriminant analysis (LDA) and nearest neighbor (NN) work with the sporulation data. We consider the set of handpicked genes with known functions as given in Chu *et al.* (1998) to be the training set. Since all of these genes are positively expressed, we restrict our attention to positively expressed genes in the classification illustration. After classification, the average temporal pro-

files of each of the seven classes are compared with the model profile (Figure 2).

All three classification tools yield similar results (Fig. 2). The 3NN-nearest neighbourhood may give the most distinct cluster boundaries (results not shown). The average temporal profiles resemble the model profile to a reasonable degree given the training set is small.

6 Validation

The dataset considered here is one of the first reliable and publicly available microarray data. Perhaps due to the heavy cost involved, the experiment was not replicated. As we all know, replication is important and essential in accessing the variability and validation of a statistical procedure. The relatively recent microarray studies incorporate at least two replications. A cross validation type approach can be formulated for checking the consistency of a grouping procedure in the case of replicated dataset. In the case of the sporulation data set of Chu *et al.* (1998), the data for a single time may be deleted to formulate a cross validation approach. Some interesting finding emerge with this approach whose details will appear elsewhere.

7 Concluding remarks

We have provided an introduction to a microarray experiment and a well known publicly available dataset on S. cerevisiae. We illustrated how statistical clustering and classification techniques are used to group genes in functional classes on the similarity of their expression profiles. Five existing statistical clustering algorithms and two classification algorithms have been considered for a comparative study.

These techniques can also be used to group tissue samples instead of genes using the entire set of gene expression levels (considered as a multivariate data of a rather large dimension) for each tissue sample (McLachlan, *et al.* 2002; Weinstein, J. N., 1992). For example, a cancer cell can be clustered or classified based on the expression profile of all the genes for that sample.

Clustering and classification techniques provide helpful preliminary guidance regarding a global picture of the relative similarity of the expression profiles. However, for studying the interactions between genes and to find the regulatory networks, additional statistical techniques such as partial least squares are needed (Datta, 2001a). The number of papers dealing with statistical methods for microarray data has been steadily increasing. A partial overview can be obtained in Datta (2001b).

The number of papers dealing with good validation strategies for clustering algorithms is rather limited at the moment. We are pursuing this issue and the results will appear elsewhere.

Acknowledgments

We thank the referee for many useful suggestions which led to a better paper. Susmita Datta's research was supported in part by US National Science Foundation grant DBI-0074642.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403-410.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Millar, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acid Research* 25, 3389-3402.

Bennett, J.W. and Arnold, J. (2001). Genomics for Fungi. In *The Mycota VIII: Biology of the Fungal Cell.* Howard, R.J. and Gow, N.A.R. (editors). New York: Springer-Verlag.

Bhalla, V.S. and Iyengar, R. (1999). Emergent properties of networks of biological signaling pathways. *Science* 283, 381-387.

Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P.O. and Herskowitz, I. (1998). The transcriptional program of sporulation in budding yeast. *Science* 282, 699-705.

Datta, S. (2001a). Exploring relationships in gene expressions: A partial least squares approach. *Gene Expression* 9, 257-264.

Datta, S. (2001b). Statistical techniques for microarray data: A partial overview, submitted.

DeRisi, J.L., Iyer, V.R. and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680-686.

Eisen, M.B., Spellman, P.T., Brown, P.O. and Botstein, D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sc. USA* 95, 14863-14868.

Geever *et al.* (1989). DNA sequence, organization and regulation of the qa gene cluster of Neurospora crassa. *J. Mol. Biol.* 207, 15-34.

Hartigan, J. (1975). *Clustering Algorithms.* New York: Wiley.

Hartigan, J. (1978). A K-means clustering algorithm. *Applied Statistics* 28, 100-108.

Jacob, F. and Monod, J. (1961). Genetic regulatory mechanisms in the synthesis of proteins. *J. Mol. Biol.* 3, 318-356.

Johnston, M. (1987). A model fungal regulatory mechanism: the GAL genes of Saccharomyces cerevisiae. *Microbiological Reviews* 51, 458-476.

Kaufman, L. and Rousseeuw, P.J. (1990). *Fitting Groups in Data.* An Introduction to Cluster Analysis. New York: John Wiley.

Lee, K., Loros, J.J. and Dunlap, J.C. (2000). Interconnected feedback loops in the Neurospora circadian system. *Science* 289, 107-110.

McLachlan, G.J., Bean, R.W. and Peel, D. (2002). A mixture model-based approach to the clustering of microarray expression data. *Bioinformatics* 18, 1-10.

Pilpel, Y., Sudarsanam, P. and Church, G.M. (2001). Identifying regulatory networks by combinatorial analysis of promoter elements. *Nature Genetics* 29, 153-159.

Prade, R.A., Ayoubi, P., Krsihnan, S., Macwana, S. and Russell, H. (2001). Accumulation of stress and inducer-dependent plant-cell-wall-degrading enzymes during asexual development in Aspergillus nidulans. *Genetics* 157, 957-967.

Ptashne, M. (1992). *A Genetic Switch: Phage Lambda and Higher Organisms.* Oxford, UK: Cell Press and Blackwell Scientific Publications.

Ross-Macdonald, P., Coelho, P.S.R., Roemer, T., Agarwal, S., Kumar, A., Jansen, R., Cheung, K.-H, Sheehan, A., Symoniatis, D., Vmansky, L., Heldtman, M., Nelson, F.K., Iwasaki, K., Hager, M., Gerstein, P., Miller, P., Roeder, G.S and Snyder, M. (1999). Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* 482, 413-417.

Sokal, R.R. and Sneath, P.H.A. (1963). *Principles of Numerical Taxonomy.* New York: Freeman.

Sveiczer, A., Csikasz-Nagy, A., Gyorffy, B., Tyson, J.J. and B. Novak (2000). Modeling the fission yeast cell cycle: quantized cycle times in wee1-cdc25D mutant cells. *Proc. Natl. Acad. Sc. USA* 97, 7865-7870.

Tufte, E.R. (1983). *The Visual Display of Quantitative Information.* Cheshire, CT.: Graphic Press.

Venables, W.N. and Ripley, B.D. (1998). *Modern Applied Statistics with S-Plus.* Second Edition (3rd corrected printing). New York: Springer.

Venter, J.C. *et al.* (2001). The sequence of the human genome. *Science* 291, 1304-1351.

Weinstein, J.N., Kurt, W.K., Michael, R.G., Vellarkad, N.V., Rubinstein, L.V., Monks, A., Scudiero, D.A., Welch, L., Koutsoukos, A.D., Chiausa, A.J. and Paull, K.D. (1992). Neural computing in cancer drug development: Predicting mechanism of action. *Science* 258, 447-451.

Yang, Y.H., Dudoit, S., Luu, P., and Speed, T.P. (2001). Normalization for cDNA microarray data. In *Microarrays: Optical Technologies and Informatics*, Bittner, M.L., Chen, Y., Dorsel, A.N. and Dougherty, E.R. (Editors). *Proc. SPIE* 4266, 141-152.

Yanofsky, C. and Kolter, R. (1982). Attenuation in amino acid biosynthesis operons. *Ann. Rev. Genetics* 16, 113-134.

MEAN-VARIANCE EDGE PLOTS FOR ENSEMBLE VOTING SCHEMES

PAMELA J. DAVY AND VIRGINIA L. WHEWAY
School of Mathematics and Applied Statistics
University of Wollongong, Wollongong, NSW 2522, Australia
E-mail: pam_davy@uow.edu.au, vlw02@uow.edu.au

Boosting, bagging and other ensemble classifiers are all based on the idea of fitting multiple classification rules and then applying a voting scheme to determine a consensus. In some methods, the individual classifiers are trained sequentially, so that the occurrence of misclassification errors in earlier iterations can be used to adapt later iterations. It is therefore of interest to consider the sequences of correct and incorrect classifications for individual observations. The edge, or in other words the voting weight assigned to incorrect classes, can be evaluated for each observation after each iteration. A scatter plot based on the mean and variance of the edge over all iterations turns out to be a useful diagnostic tool.

Keywords: boosting, ensemble voting.

1 Introduction

Consider a classification problem for which the final classification rule is obtained by applying a voting scheme to a sequence of m individual classifiers. Examples of base classifiers include decision trees and neural networks. The usual method of obtaining variation among individual classifiers is to modify the training by reweighting, resampling or relabelling responses. An alternative approach is to randomise some aspects of the fitting algorithm, a typical example being the random initialisation of weights in a neural network. In some schemes such as bagging (see Breiman, 1996), the individual classifiers can be fitted in parallel, so that there is no inherent order in the m iterations. Other schemes such as boosting (see Breiman, 1996) involve adaptive modifications to the training set based on earlier iterations. Although this rules out the possibility of parallel computation, adaptive schemes generally result in better performance, as discussed by Bauer and Kohavi (1999).

The individual classifiers are sometimes given equal voting weights, or can be assigned different weights based on classification performance within the training set. The ensemble classification rule usually selects the class with the largest weighted sum of votes, but other variations are also possible.

Although the final classification rule and its performance on unseen data

are arguably the most important outcomes of a particular application, from a diagnostic point of view it is also of interest to track the evolution of the ensemble classifier over successive iterations and to identify any problem observations within the training set. Despite the best efforts expended in the data cleaning phase of data mining, outliers and anomalies are a reality in most real data sets. Adaptive classification schemes can try very hard to accommodate such anomalies, often at the expense of performance within the remainder of the dataset.

2 Edge behaviour of a fixed observation

The *edge* of an observation in the training set used for an ensemble classification problem is the total voting weight assigned to incorrect classifications (see Breiman, 1995). For a particular observation, let $\mathbf{Z} = (Z_1, \ldots, Z_m)'$ be the vector of indicator variables identifying misclassification for the m individual classifiers. After k iterations, the edge of this observation can be expressed in the form $\sum_{j=1}^{k} c_{kj} Z_j$, where c_{kj} is the voting weight given to the jth classifier ($\sum_{j=1}^{k} c_{kj} = 1$). In matrix notation, we may express the vector of m edge statistics as $\mathbf{CZ}$ where

$$\mathbf{C} = \begin{pmatrix} c_{11} & 0 & \ldots & 0 \\ c_{21} & c_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ c_{m1} & c_{m2} & \ldots & c_{mm} \end{pmatrix}$$

Note that $\mathbf{C1} = \mathbf{1}$.

The mean edge of this particular observation over the m iterations is given by $\mu'\mathbf{Z}$ and the variance is $\mathbf{Z}'\mathbf{\Sigma}\mathbf{Z}$ where

$$\mu = \frac{1}{m}\mathbf{C}'\mathbf{1} \tag{1}$$

and

$$\mathbf{\Sigma} = \frac{1}{m}\mathbf{C}'\left(\mathbf{I} - \frac{1}{m}\mathbf{J}\right)\mathbf{C}. \tag{2}$$

Here $\mathbf{I}$ is an $m \times m$ identity matrix and $\mathbf{J}$ is an $m \times m$ matrix of ones.

Since $\mathbf{C1} = \mathbf{1}$, it follows that

$$\begin{aligned} \mu'\mathbf{1} &= \frac{1}{m}\mathbf{1}'\mathbf{C1} \\ &= \frac{1}{m}\mathbf{1}'\mathbf{1} \\ &= 1 \end{aligned} \tag{3}$$

and

$$\begin{aligned}\boldsymbol{\Sigma}\mathbf{1} &= \frac{1}{m}\mathbf{C}'\left(\mathbf{I} - \frac{1}{m}\mathbf{J}\right)\mathbf{1} \\ &= \frac{1}{m}\mathbf{C}'(\mathbf{1} - \mathbf{1}) \\ &= \mathbf{0}\end{aligned} \tag{4}$$

For any binary vector $\mathbf{Z}$,

$$\begin{aligned}\mu'(\mathbf{1} - \mathbf{Z}) &= \mu'\mathbf{1} - \mu'\mathbf{Z} \\ &= 1 - \mu'\mathbf{Z}\end{aligned} \tag{5}$$

and

$$\begin{aligned}(\mathbf{1} - \mathbf{Z})'\boldsymbol{\Sigma}(\mathbf{1} - \mathbf{Z}) &= \mathbf{1}'\boldsymbol{\Sigma}\mathbf{1} - 2\mathbf{1}'\boldsymbol{\Sigma}\mathbf{Z} + \mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z} \\ &= \mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z}.\end{aligned} \tag{6}$$

For a two-class classification problem, identities (5) and (6) have a simple interpretation in terms of "mirror" observations. Consider two observations with an identical sequence of predictions for the m classifiers, but with opposite class labels. This would be the case if the observations were identical except for the response variable, but could also occur if the predictor variables differed only in details considered unimportant by the boosting algorithm. Whenever a classifier correctly predicts one of these observations, it misclassifies the other. The edge statistics for one observation can therefore found by subtracting the edge statistics of the other observation from 1, and likewise for the mean edge. However the variance of the edge will be identical for the two observations. On a scatter plot of variance versus mean, these two observations will appear as mirror images about mean edge = 0.5.

Wheway (2001) has constructed mean-variance edge plots for a variety of datasets, using 10 boosting iterations and C4.5 (see Quinlan, 1993) as the base classifier. Initially all observations are given equal weights. In the $(j+1)$th iteration, the weight of each observation misclassified in iteration j is multiplied by the weighted odds ratio of correct classification. Attention is thereby focussed on difficult observations. The final ensemble classifier is found by allocating voting weights proportional to the weighted log odds ratios of correct classification. In particular, Figures (1) and (2) show the resulting plots for the *colic* and *hearth* datasets from the UCI (see Blake and Marz, 1988) repository. The labels inside the plot are observation identification numbers. These plots, like others not shown here, display a characteristic pattern. The majority of observations lie in the low mean/low variance corner of the plot, and display a positive relationship between mean and variance. A minority of

points display either high variance and/or high mean, with a clear negative relationship between mean and variance. A finer scale banded structure is also evident.

From a diagnostic point of view, high mean edge indicates intractable observations, while high edge variance indicates instability as iterations continue. The presence of observations with high mean edge is mostly due to the nature of the underlying data, while high edge variance is due to the sensitivity of the base classifier to changes induced by the boosting algorithm. The recurrence of the same type of pattern in the mean-variance edge plots of different datasets may seem surprising at first, and the reasons will be explored in the next section.

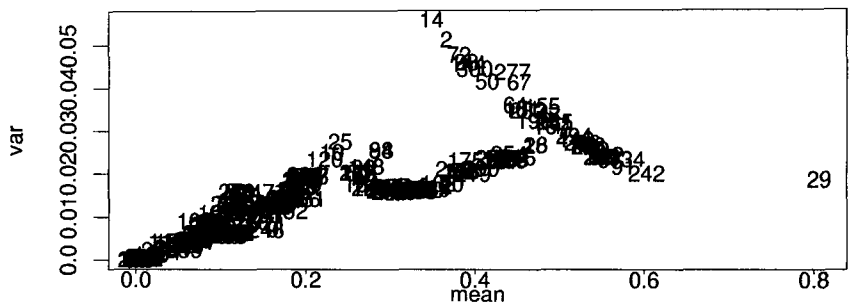


Figure 1. Mean-variance edge plot for the *colic* data set

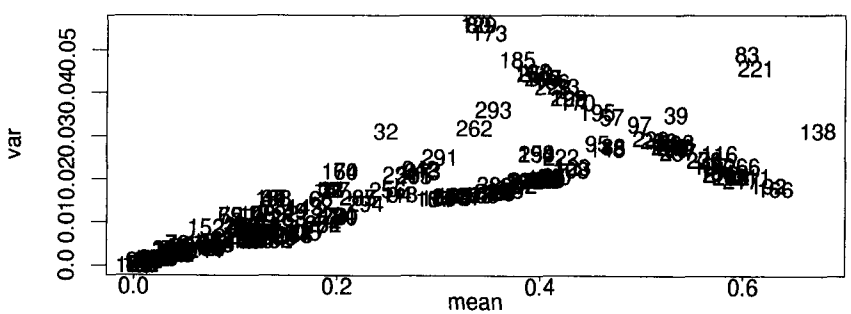


Figure 2. Mean-variance edge plot for the *hearth* data set

3 Mean-variance state space

For m iterations, there are 2^m possible binary vectors $\mathbf{Z}$. In any particular application, only a subset of these possibilities will actually be realised. However in order to obtain a better understanding of the underlying ensemble classification algorithm, it is of interest to plot the entire state space. In other words, each possible pair of mean edge and edge variance values can be plotted, irrespective of whether such values were actually observed within the dataset and irrespective of whether Figure (3) shows such a state space plot for equal voting weights, in which case

$$\mathbf{C} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ \frac{1}{2} & \frac{1}{2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \frac{1}{m} & \frac{1}{m} & \dots & \frac{1}{m} \end{pmatrix}.$$

For unequal voting weights, the corresponding plot (not shown here) shows similar qualitative behaviour. Note the symmetry of the state space plot, which is a consequence of mirror pairs as discussed in the previous section. Other features include the "forbidden zone" above the horizontal axis, the negative trend on the left side of the plot, and the fine-scale pattern of curved bands.

Of course there are many other ways of displaying 2^m binary vectors in two dimensions, including a $2^a \times 2^{m-a}$ rectangular lattice. The advantage of this particular representation is that it leads to direct interpretations relating to the underlying dataset and the behaviour of the ensemble algorithm. By comparing the state space plot to Figures (1) and (2), it can be observed that the boosting algorithm avoids regions of high mean edge.

Further understanding can be obtained by plotting subsets of the state space.

For example, Figures (4) and (5) show the regions of state space occupied by observations which are misclassified in the first and final iterations respectively. Clearly the observations misclassified on the first iteration tend to have higher mean edge than observations misclassifed at later iterations. This is because the initial misclassification contributes to all of the m edge statistics, while a misclassification at iteration j only contributes to $m+1-j$ edge statistics. A misclassification at the final iteration only contributes to the final edge statistic, so the mean-variance edge plot for observations misclassified in the final iteration looks very similar, although less dense, to the plot of the entire state space shown in Figure (3). Note that Figures (3), (4) and (5) all depict possible states rather than realised states for a particular

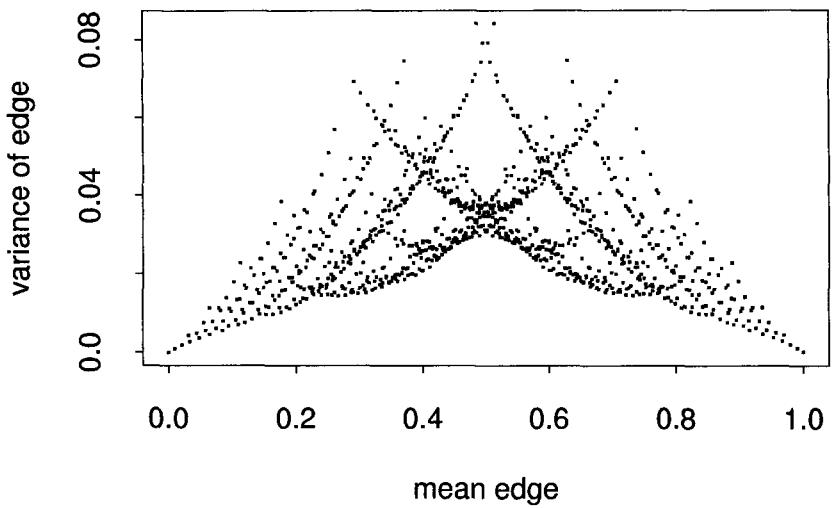


Figure 3. Complete state space for mean-variance edge plot

dataset, and therefore show more points in the high mean edge region of the plot than would be expected in particular applications.

In order to understand the basic structure of the state space plot, it is useful to consider the variance of the edge as a function of Z_1, the misclassification indicator for the initial classifier. Let μ_* denote the vector obtained by omitting the first element of μ, and $\mathbf{\Sigma}_*$ denote the matrix obtained by omitting the first row and column of $\mathbf{\Sigma}$. Also let $\mathbf{Y}' = (\, Z_2 - Z_1 \quad \ldots \quad Z_m - Z_1 \,)$. As $\mathbf{\Sigma 1} = \mathbf{0}$, we may express the variance as

$$\begin{aligned} \mathbf{Z}'\mathbf{\Sigma}\mathbf{Z} &= (\mathbf{Z} - Z_1\mathbf{1})'\mathbf{\Sigma}(\mathbf{Z} - Z_1\mathbf{1}) \\ &= \mathbf{Y}'\mathbf{\Sigma}_*\mathbf{Y} \end{aligned} \tag{7}$$

The ratio of quadratic forms $\mathbf{Y}'\mu_*\mu_*'\mathbf{Y}/\mathbf{Y}'\mathbf{\Sigma}_*\mathbf{Y}$ achieves a maximum value of

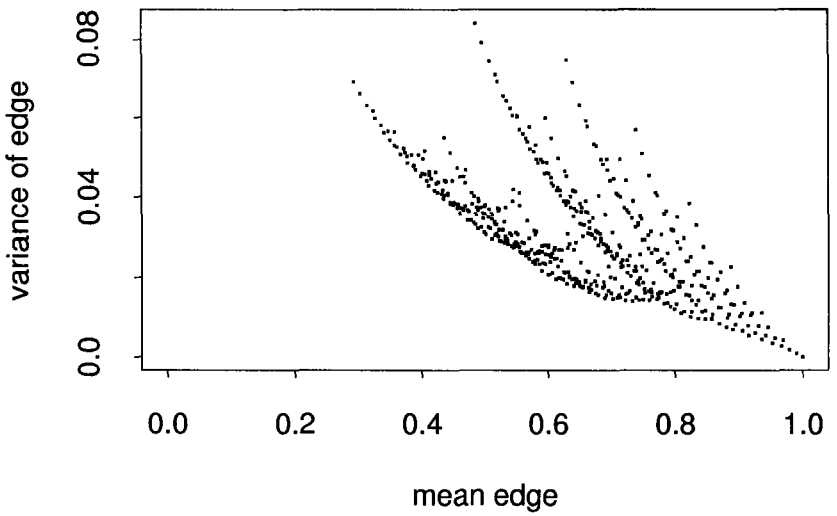


Figure 4. Mean-variance state space for observations initially misclassified ($m = 10, Z_1 = 1$)

$\mu_*'\boldsymbol{\Sigma}^{-1}\mu_*$ when $\mathbf{Y} = \boldsymbol{\Sigma}_*^{-1}\mu_*$. Therefore

$$\mathbf{Y}'\boldsymbol{\Sigma}_*\mathbf{Y} \geq \frac{\mathbf{Y}'\mu_*\mu_*'\mathbf{Y}}{\mu_*'\boldsymbol{\Sigma}_*^{-1}\mu_*} \tag{8}$$

But $\mu'_*\mathbf{Y} = \mu'\mathbf{Z} - Z_1$ and

$$\begin{aligned}\mu_*'\boldsymbol{\Sigma}_*^{-1}\mu_* &= \frac{1}{m}\mathbf{1}'_{m-1}\mathbf{C}_*\mathbf{C}_*^{-1}(\mathbf{I}_{m-1} - \frac{1}{m}\mathbf{J}_{m-1})^{-1}(\mathbf{C}_*')^{-1}\mathbf{C}_*'\mathbf{1}_{m-1} \\ &= \frac{1}{m}\mathbf{1}'_{m-1}(\mathbf{I}_{m-1} + \mathbf{J}_{m-1})\mathbf{1}_{m-1} \\ &= m - 1\end{aligned}$$

Hence

$$\mathbf{Z}'\boldsymbol{\Sigma}\mathbf{Z} \geq (\mu'\mathbf{Z} - Z_1)^2/(m-1) \tag{9}$$

In other words, the variance is bounded below by the parabola $\text{mean}^2/(m-1)$ for $Z_1 = 0$ and by the parabola $(\text{mean} - 1)^2/(m-1)$ for $Z_1 = 1$. This explains

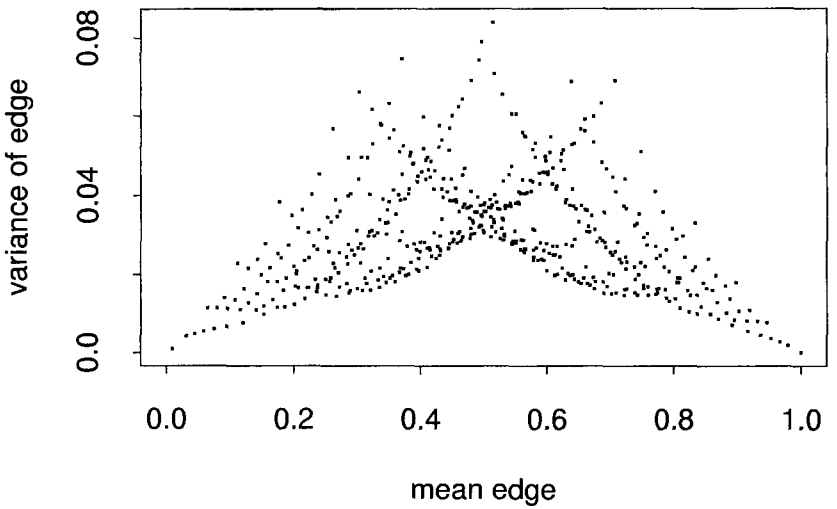


Figure 5. Mean-variance state space for observations misclassified by final classifier ($m = 10, Z_m = 1$

the positive and negative trends between mean and variance on the left and right sides of the plot respectively.

4 Conclusion

For each observation in the training set, the edge statistic can be recomputed as each new iteration of an ensemble classification algorithm is completed. A scatter plot of the mean and variance of the edge over all iterations is useful both for identification of anomolous data points and for understanding the nature of an adaptive algorithm such as boosting. The removal of right-hand mirror points from the training set has been found by Wheway (2001) to result in simpler classifiers (i.e. decision trees with fewer nodes) without degradation of generalisation error.

If a test set is available, the sequence of missclassification indicator vari-

ables can be computed for each test observation, and the mean and variance of the edge can be found using the same voting weights used in the training set. Test points can then be plotted on the same mean-variance edge plot as the training points, using a different colour or marker. One avenue of future research involves the comparison of mean-variance edge plots for different ensemble classifiers applied to the same dataset. The underlying state space will vary as the voting weights change, and different algorithms will tend to concentrate realised test and training points within different regions of the available state space.

REFERENCES

Bauer, E. and Kohavi, R. (1999). An empirical comparison of voting classification algorithms: bagging, boosting, and variants, *Machine Learning* 36, 105-142.

Blake, C.L. and Merz, C.J. (1998). UCI repository of machine learning databases, *http://www1.ics.uci.edu/ mlearn/MLRepository.html.*

Breiman, L. (1995). Arcing the edge. *Technical Report 486*, Department of Statistics, University of California, Berkeley.

Breiman, L. (1996). Bagging Predictors, *Machine Learning* 26, 123-140.

Quinlan, J.R. (1993). *Programs for Machine Learning.* San Francisco: Morgan Kaufman.

Schapire, R.E., Freund, Y., Bartlett, P. and Lee, W.S. (1988), Boosting the margin: a new explanation for the effectiveness of voting methods, *Annals of Statistics* 26, 1651-1686.

Wheway, V.L. (2001) Using boosting to simplify classification models, *Proceedings of the 2001 IEEE International Conference on Data Mining*, 558-565. IEEE Computer Society: Los Alamitos, California.

APPROXIMATING THE CAUCHY PRINCIPAL VALUE INTEGRAL VIA HERMITE-HADAMARD TYPE INEQUALITIES

S. S. DRAGOMIR
School of Communications and Informatics
Victoria University of Technology
PO Box 14428, Melbourne City MC, 8001, Victoria, Australia.
E-mail: sever@matilda.vu.edu.au

Numerical evaluation of the Cauchy principal value (CPV) integral, is encountered in many areas of Applied Mathematics such as the crack problems in plane elasticity, the singular eigenfunction method in neutron transport, in airfoil theory, in electromagnetic scattering in waveform encoding of binary signals, in visualization of cardiovascular velocity maps and in demultiplexing of interferometrically interrogated fiber Bragg Gratings etc. In the present paper, by the use of Hermite-Hadamard inequality for convex functions, we establish some error bounds for the approximation of the CPV integral of a differentiable function whose derivative is convex. Some numerical experiments are performed as well.

Keywords: Cauchy principal value, Hilbert transform, convex functions.

1 Introduction

Let $\Omega = (a, b)$ be an open interval of real numbers. The usual $\mathcal{L}^1-$space with respect to the Lebesgue measure λ restricted to the open interval Ω will be denoted by $\mathcal{L}^1(\Omega)$.

We define a linear operator T (see Okada and Elliott, 1994) from the vector space $\mathcal{L}^1(\Omega)$ into the vector space of all $\lambda-$measurable functions on Ω as follows. Let $f \in \mathcal{L}^1(\Omega)$. The Cauchy principle value integral

$$\frac{1}{\pi} PV \int_a^b \frac{f(\tau)}{\tau - t} d\tau = \lim_{\varepsilon \downarrow 0} \left[\int_a^{t-\varepsilon} + \int_{t+\varepsilon}^b \right] \frac{f(\tau)}{\pi(\tau - t)} d\tau \tag{1}$$

exists for $\lambda-$almost every $t \in \Omega$.

We denote the left-hand side of (1) by $(Tf)(a, b; t)$ for each $t \in \Omega$ for which $(Tf)(a, b; t)$ exists. The so-defined function Tf, which we call the *finite Hilbert Transform* of f, is defined $\lambda-$almost everywhere on Ω and is $\lambda-$measurable; (see, for example, Butzer and Nessel, 1997, Theorem 8.1.5). The resulting linear operator T will be called the *finite Hilbert transform operator.*

Now, if we assume that the mapping $f : (a,b) \to \mathbf{R}$ is convex on (a,b), then it is locally Lipschitzian on (a,b) and then the finite Hilbert transform of f exists in every point $t \in (a,b)$.

The following result holds (see Dragomir *et al.*, 2001).

Theorem 1 *Let $f : (a,b) \to \mathbf{R}$ be a differentiable convex function on (a,b). Then we have the inequality*

$$\begin{aligned} &\frac{1}{\pi}\left[f(t) - f(a) + f'(t)(b-t) + f(t)\ln\left(\frac{b-t}{t-a}\right)\right] \\ &\leq (Tf)(a,b;t) \\ &\leq \frac{1}{\pi}\left[f(t)\ln\left(\frac{b-t}{t-a}\right) + f(b) - f(t) + f'(t)(t-a)\right] \end{aligned}$$

for all $t \in (a,b)$.

In this paper, by the use of the well known Hermite-Hadamard integral inequality for convex functions, we point out some inequalities for the finite Hilbert transform of functions whose first derivatives are convex. Some numerical experiments for particular examples of such functions are performed as well.

2 The main results

Firstly, we recall that the following integral inequality

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(t)dt \leq \frac{f(a)+f(b)}{2}$$

where $f : [a,b] \to \mathbf{R}$ is convex, is well known in the literature as the Hermite-Hadamard inequality.

The following lemma holds.

Lemma 1 *Let $g : [a,b] \to \mathbf{R}$ be a convex function on $[a,b]$ and $t, \tau \in [a,b]$ with $t \neq \tau$. If $0 = \lambda_0 < \lambda_1 < \ldots < \lambda_{n-1} < \lambda_n = 1$, then we have the inequality:*

$$\begin{aligned} &\sum_{i=0}^{n-1}(\lambda_{i+1} - \lambda_i)\, g\left[\left(1 - \frac{\lambda_i + \lambda_{i+1}}{2}\right)t + \frac{\lambda_i + \lambda_{i+1}}{2}\cdot\tau\right] \qquad (2) \\ &\leq \frac{1}{\tau - t}\int_t^\tau g(u)\,du \end{aligned}$$

$$\leq \frac{1}{2}\sum_{i=0}^{n-1}(\lambda_{i+1}-\lambda_i)\left\{g\left[(1-\lambda_i)t+\lambda_i\tau\right]+g\left[(1-\lambda_{i+1})t+\lambda_{i+1}\tau\right]\right\}.$$

Proof: Consider the partitioning of $[t,\tau]$ (if $t<\tau$) or $[\tau,t]$ (if $\tau<t$) given by

$$I_n: x_i=(1-\lambda_i)t+\lambda_i\tau,\ \ (i=0,1,\ldots,n-1).$$

Then, obviously,

$$\frac{x_i+x_{i+1}}{2}=\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot\tau,\ \ (i=0,1,\ldots,n-1)$$

and

$$x_{i+1}-x_i=(\tau-t)(\lambda_{i+1}-\lambda_i),\ \ (i=0,1,\ldots,n-1).$$

Applying the Hermite-Hadamard inequality on $[x_i,x_{i+1}]$ $(i=0,1,\ldots,n-1)$, we may write that

$$\begin{aligned}
&g\left[\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot\tau\right]\\
&\leq\frac{1}{(\tau-t)(\lambda_{i+1}-\lambda_i)}\int_{x_i}^{x_{i+1}}g(u)\,du\\
&\leq\frac{1}{2}\left\{g\left[(1-\lambda_i)t+\lambda_i\tau\right]+g\left[(1-\lambda_{i+1})t+\lambda_{i+1}\tau\right]\right\}
\end{aligned}$$

for any $i=0,1,\ldots,n-1$. If we multiply by $\lambda_{i+1}-\lambda_i>0$ and sum over i from 0 to $n-1$, we deduce the desired inequality (2).

The following theorem holds.

Theorem 2 *Assume that $f:(a,b)\to\mathbf{R}$ has its derivative convex on (a,b). Then for all $n\geq 1$, and $0=\lambda_0<\lambda_1<\ldots<\lambda_{n-1}<\lambda_n=1$, we have the inequality*

$$\begin{aligned}
&\frac{f(t)}{\pi}\ln\left(\frac{b-t}{t-a}\right)+\frac{b-a}{\pi}\sum_{i=0}^{n-1}(\lambda_{i+1}-\lambda_i) \qquad (3)\\
&\quad\times\left[f;\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot b,\right.\\
&\qquad\left.\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot a\right]\leq(Tf)(a,b;t)
\end{aligned}$$

$$\leq\frac{1}{2\pi}\left\{\lambda_1(b-a)f'(t)+(1-\lambda_{n-1})\left[f(b)-f(a)\right]\right\}$$

$$+\frac{b-a}{2\pi}\sum_{i=0}^{n-1}(\lambda_{i+1}-\lambda_i)\,[f;(1-\lambda_i)\,t+\lambda_i b,(1-\lambda_i)\,t+\lambda_i a]$$
$$+\frac{f(t)}{\pi}\ln\left(\frac{b-t}{t-a}\right)$$

for any $t\in(a,b)$.

Proof: If we write the inequality (2) for f', then we have

$$\sum_{i=0}^{n-1}(\lambda_{i+1}-\lambda_i)\,f'\left[\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot\tau\right] \tag{4}$$
$$\leq\frac{f(\tau)-f(t)}{\tau-t}$$
$$\leq\frac{1}{2}\sum_{i=0}^{n-1}(\lambda_{i+1}-\lambda_i)\left\{f'\left[(1-\lambda_i)\,t+\lambda_i\tau\right]+f'\left[(1-\lambda_{i+1})\,t+\lambda_{i+1}\tau\right]\right\}$$
$$=\frac{1}{2}\left[\lambda_1 f'(t)+\sum_{i=1}^{n-1}(\lambda_{i+1}-\lambda_i)\,f'\left[(1-\lambda_i)\,t+\lambda_i\tau\right]\right.$$
$$\left.+\sum_{i=0}^{n-2}(\lambda_{i+1}-\lambda_i)\,f'\left[(1-\lambda_{i+1})\,t+\lambda_{i+1}\tau\right]+(1-\lambda_{n-1})\,f'(\tau)\right]$$
$$=\frac{1}{2}\left[\lambda_1 f'(t)+\sum_{i=1}^{n-1}(\lambda_{i+1}-\lambda_i)\,f'\left[(1-\lambda_i)\,t+\lambda_i\tau\right]+(1-\lambda_{n-1})\,f'(\tau)\right].$$

Applying the PV over t, i.e., $\lim\limits_{\varepsilon\to0+}\left(\int_a^{t-\varepsilon}+\int_{t+\varepsilon}^b\right)$ to the inequality (4), we deduce

$$\sum_{i=0}^{n-1}(\lambda_{i+1}-\lambda_i)\,PV\int_a^b f'\left[\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot\tau\right]d\tau \tag{5}$$
$$\leq PV\int_a^b\frac{f(\tau)-f(t)}{\tau-t}d\tau$$
$$\leq\frac{1}{2}\left[\lambda_1(b-a)\,f'(t)+\sum_{i=1}^{n-1}(\lambda_{i+1}-\lambda_i)\,PV\int_a^b f'\left[(1-\lambda_i)\,t+\lambda_i\tau\right]d\tau\right.$$
$$\left.+(1-\lambda_{n-1})\,(f(b)-f(a))\right].$$

Since

$$
\begin{aligned}
&PV\int_a^b f'\left[\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot\tau\right]d\tau\\
&=\frac{2}{\lambda_i+\lambda_{i+1}}\left(f\left[\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot b\right]\right.\\
&\quad\left.-f\left[\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot a\right]\right)\\
&=(b-a)\left[f;\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot b,\right.\\
&\qquad\left.\left(1-\frac{\lambda_i+\lambda_{i+1}}{2}\right)t+\frac{\lambda_i+\lambda_{i+1}}{2}\cdot a\right]
\end{aligned}
$$

and

$$
PV\int_a^b f'\left[(1-\lambda_i)t+\lambda_i\tau\right]d\tau=(b-a)\left[f;(1-\lambda_i)t+\lambda_i b,(1-\lambda_i)t+\lambda_i a\right],
$$

then by (5) we deduce the desired inequality (3).

The following corollary also holds.

Corollary 1 *Assume that* $f:(a,b)\to\mathbf{R}$ *fulfills the hypothesis of the above theorem. Then for* $n\geq 1$ *we have:*

$$
\begin{aligned}
&\frac{f(t)}{\pi}\ln\left(\frac{b-t}{t-a}\right)+\\
&\frac{b-a}{\pi}\left[\frac{1}{2^{n-1}}\left[f;\left(1-\frac{1}{2^n}\right)t+\frac{1}{2^n}b,\left(1-\frac{1}{2^n}\right)t+\frac{1}{2^n}a\right]\right]\\
&+\frac{b-a}{\pi}\sum_{i=1}^{n-1}\frac{1}{2^{n-1}}\left[f;\left(1-\frac{3}{2^{n-i}}\right)t+\frac{3}{2^{n-i}}b,\left(1-\frac{3}{2^{n-i}}\right)t+\frac{3}{2^{n-i}}a\right]\\
&\leq(Tf)(a,b;t)\\
&\leq\frac{1}{2\pi}\left\{\frac{(b-a)f'(t)}{2^{n-1}}+\frac{1}{2}\left[f(b)-f(a)\right]\right\}\\
&+\frac{b-a}{2\pi}\left(\frac{1}{2^{n-2}}-1\right)\left[f;\left(1-\frac{1}{2^{n-1}}\right)t+\frac{1}{2^{n-1}}b,\left(1-\frac{1}{2^{n-1}}\right)t+\frac{1}{2^{n-1}}a\right]\\
&+3\cdot\frac{b-a}{2\pi}\sum_{i=2}^{n-1}\frac{1}{2^{n-i+1}}\left[f;\left(1-\frac{1}{2^{n-i}}\right)t+\frac{1}{2^{n-i}}b,\left(1-\frac{1}{2^{n-i}}\right)t+\frac{1}{2^{n-i}}a\right]\\
&+\frac{f(t)}{\pi}\ln\left(\frac{b-t}{t-a}\right),
\end{aligned}
$$

for any $t \in (a,b)$.

The proof follows by Theorem 2 applied for $\lambda_0 = 0$, $\lambda_i = \frac{2^i}{2^n}$, $i = 1, \ldots, n$. We omit the details.

3 Numerical experiments

For an equidistant partitioning of $[a,b]$, let us define the following bounds

$$L_n(f,a,b;t) := \frac{b-a}{n\pi} \sum_{i=1}^{n-1} \left[f; t - \left(i + \frac{1}{2}\right) \cdot \frac{t-a}{n}, t + \left(i + \frac{1}{2}\right) \cdot \frac{b-t}{n} \right] + \frac{f(t)}{\pi} \ln\left(\frac{b-t}{t-a}\right)$$

called the *lower bound* and

$$U_n(f,a,b;t) := \frac{f(b) - f(a) + f'(t)(b-a)}{2n\pi} + \frac{b-a}{n\pi} \sum_{i=1}^{n-1} \left[f; t - i \cdot \frac{t-a}{n}, t + i \cdot \frac{b-t}{n} \right] + \frac{f(t)}{\pi} \ln\left(\frac{b-t}{t-a}\right)$$

called the *upper bound* for the Finite Hilbert Transform $(Tf)(a,b;t)$. We also define the *left error* $LEr_n(f,a,b;t)$

$$LEr_n(f,a,b;t) := (Tf)(a,b;t) - L_n(f,a,b;t) \geq 0$$

and the *right error* $REr_n(f,a,b;t)$

$$REr_n(f,a,b;t) := U_n(f,a,b;t) - (Tf)(a,b;t) \geq 0$$

and will investigate them numerically for different functions f and natural numbers n.

If we consider the function $f : [-1,1] \to R, f(x) := \exp(x)$, then the exact finite Hilbert transform provided by Maple 6, is

$$(Tf)(-1,1;t) = (exp(t)Ei(1-t) - exp(t)Ei(-t-1))/\pi, t \in [-1,1].$$

If we plot in the same system of co-ordinates $LEr_n(f,a,b;t)$ and $REr_n(f,a,b;t)$, for $n = 100$, then we observe that the distance between the exact Hilbert transform and its lower bound is smaller than the distance between the same Hilbert transform and its upper bound (see Figure 1). A theoretical investigation on this fact is conducted in Dragomir (2002).

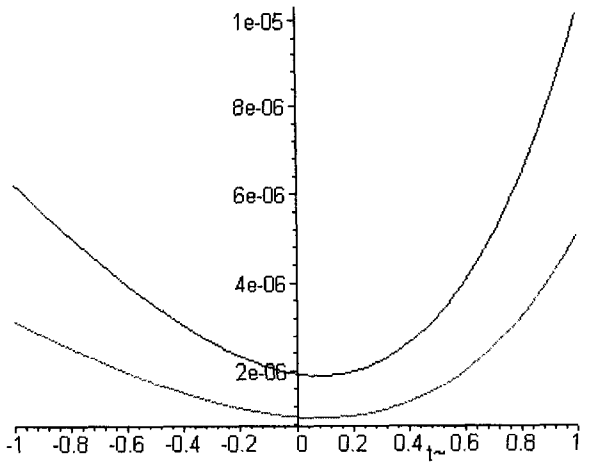


Figure 1.

REFERENCES

Butzer, P.L., and Nessel, R.J. (1977), *Fourier Analysis and Approximation Theory*, Vol. 1, Basel: Birkhaüser Verlag.

Dragomir, N.M., Dragomir, S.S., and Farrell, P. (2001), Some inequalities for the finite Hilbert transform, 113-122. In *Theory of Inequalities and Applications*, Cho, Y.J., Dragomir, S.S. and Kim, J.Y. (Editors). New York: Nova Science Publishers.

Dragomir, S.S. (2002), The application of Bullen's inequality for convex functions to the finite Hilbert transform, in preparation.

Okada, S. and Elliott, D. (1994), Hölder continuous functions and the finite Hilbert transform, *Mathematische Nach.*, 163, 250-272.

NONHOMOGENEOUS SUSCEPTIBLES AND INFECTIVES AMONG NEEDLE SHARING IVDUS

JOE. GANI
School of Mathematical Sciences
Australian National University
Canberra ACT 0200, Australia
E-mail: gani@maths.anu.edu.au

This paper extends some earlier random allocation models where either susceptibles or infectives, but not both, among needle sharing IVDUs are nonhomogeneous. We begin with a brief review of earlier results, illustrated by an example of vaccination. We then consider the case where there are two types of susceptibles and two types of infectives. This leads to unwieldy algebra; a model with successive waves of infectives exchanging needles with two types of susceptibles is outlined.

Keywords: infective, susceptible, needle sharing, random allocation model.

1 Introduction

Gani (1991), (1993) was able to develop some random allocation models for needle sharing among intravenous drug users (IVDUs); in these, both susceptibles and infectives were homogeneous. Gani and Yakowitz (1993) used these models to describe the spread of HIV among IVDUs. More recently, Gani (2001), extended the models to the case where the infectives are homogeneous, but there are two or more types of susceptible. In later work, two types of infective exchanged needles with homogeneous susceptibles. The present paper extends these models to a situation in which there are two types of infective and two types of susceptible. This model can be generalized further, but leads to unwieldy algebra. The paper concludes with a model in which successive waves of infectives exchange needles with vaccinated and unvaccinated susceptibles.

In Gani (2001), it was shown that if there are i infectives exchanging needles with n_1 susceptibles of type 1 with probability $p > 0$, and n_2 of type 2 with probability $q = 1 - p > 0$, then the joint probability generating function (pgf) of the newly generated infectives of type 1 and 2 respectively is

$$\varphi_{i;n_1 n_2}(u,v) = \sum_{j=0}^{i} \binom{i}{j} p^i q^{i-j} \varphi_{j\,n_1}(u) \varphi_{i-j\,n_2}(v). \tag{1}$$

The pgfs on the right hand side satisfy the respective difference-differential equations

$$\varphi_{j+1\,n_1}(u) = \frac{u(1-u)}{n_1}\frac{d\varphi_{j\,n_1}}{du} + u\varphi_{j\,n_1},$$
$$\varphi_{i-j+1\,n_2}(v) = \frac{v(1-v)}{n_2}\frac{d\varphi_{i-j\,n_2}}{dv} + v\varphi_{i-j\,n_2}. \tag{2}$$

where these pgfs can be expressed explicitly for $1 \le r \le j+1$ in the form

$$\varphi_{j+1\,n_1}(u) = u^{j+1}a_{j+1}(j+1)\left[\left(1-\frac{1}{n_1}\right)\cdots\left(1-\frac{j}{n_1}\right)\right] + \cdots$$
$$\cdots + \frac{u^r}{n_1^{j+1-r}}a_r(j+1)\left[\left(1-\frac{1}{n_1}\right)\cdots\left(1-\frac{r-1}{n_1}\right)\right] + \cdots$$
$$\cdots + \frac{u}{n_1^j}a_1(j+1), \tag{3}$$

with $a_{j+1}(j+1) = a_1(j+1) = 1$, and $a_r(j+1) = ra_r(j) + a_{r-1}(j)$. These coefficients have been derived explicitly in terms of products of known matrices. A similar expression holds for $\varphi_{i-j+1\,n_2}(v)$. While (3) holds on the assumption that $n_1 > j$, the equation is equally valid for $n_1 \le j$. It was also shown in the same paper that for the newly generated infectives $Y_{1;j\,n_1}$ and $Y_{2;i-j\,n_2}$ of respective types 1 and 2, the expectations are given by

$$m_1 = E(Y_{1;j\,n_1}) = n_1\left(1-\left(1-\frac{1}{n_1}\right)^j\right),$$
$$m_2 = E(Y_{2;i-j\,n_2}) = n_2\left(1-\left(1-\frac{1}{n_2}\right)^{i-j}\right). \tag{4}$$

Their variances are

$$E(Y_{1;j\,n_1}^2) - m_1^2 = n_1(n_1-1)\left(1-\frac{2}{n_1}\right)^j + n_1\left(1-\frac{1}{n_1}\right)^j - n_1^2\left(1-\frac{1}{n_1}\right)^{2j},$$
$$E(Y_{2;i-j\,n_2}^2) - m_2^2 = n_2(n_2-1)\left(1-\frac{2}{n_2}\right)^{i-j} + n_2\left(1-\frac{1}{n_2}\right)^{i-j}$$
$$-n_2^2\left(1-\frac{1}{n_2}\right)^{2(i-j)}. \tag{5}$$

We now give an example of this case, where the two types of susceptible are unvaccinated and vaccinated individuals, in order to illustrate the effect of vaccination on the spread of infection.

2 The effect of vaccination

Let us consider the special case of n susceptibles of two types, of which n_1 are unvaccinated against a particular disease (hepatitis, say) and $n - n_1$ are vaccinated. We evaluate the effect of vaccination after an exchange of needles between the i infectives and the n susceptibles. We assume that the i initial infectives exchange needles at random with the susceptibles, so that the probabilities of exchange with the unvaccinated and vaccinated susceptibles are respectively

$$p = \frac{n_1}{n} \qquad \text{and} \qquad q = \frac{n - n_1}{n}.$$

Then, from Equation (1), on the assumption that the vaccinated individuals are immune to the disease, setting $v = 1$, we see that the pgf of the newly generated infectives is

$$\varphi_{i\,n_1}(u) = \sum_{j=0}^{i} \binom{i}{j} \left(\frac{n_1}{n}\right)^j \left(1 - \frac{n_1}{n}\right)^{i-j} \varphi_{j\,n_1}(u). \tag{6}$$

The pgf $\varphi_{j\,n_1}(u)$ satisfies the first difference-differential equation in (2), with the expectation and variance of $Y_{1;j\,n_1}$, the newly generated infectives, as in (4) and (5).

It follows from (4) and (6) that the total expected number of newly generated infectives is given by

$$\sum_{j=0}^{i} \binom{i}{j} \left(\frac{n_1}{n}\right)^j \left(1 - \frac{n_1}{n}\right)^{i-j} n_1 \left(1 - \left(1 - \frac{1}{n_1}\right)^j\right) = n_1 \left(1 - \left(1 - \frac{1}{n}\right)^i\right), \tag{7}$$

while the variance of the total number of new infectives is

$$n_1(n_1 - 1)\left(1 - \frac{2}{n}\right)^i + n_1\left(1 - \frac{1}{n}\right)^i - n_1^2\left(1 - \frac{2}{n}\left(1 - \frac{1}{2n_1}\right)\right)^i. \tag{8}$$

We note that when $i = cn$, where c is some constant, and both i and n are large, we can derive the following asymptotic results for the expectation and variance respectively:

$$n_1(1 - e^{-c}) \qquad \text{and} \qquad n_1(n_1 - 1)e^{-2c} + n_1 e^{-c} - n_1^2 e^{-2c(1 - \frac{1}{2n_1})}. \tag{9}$$

It is clearly desirable to keep the number of new infectives to a minimum; thus if we require the average proportion of new infectives among the

n susceptibles to be no greater than some value a (0.05, say), then

$$n_1\left(1-\left(1-\frac{1}{n}\right)^i\right) \leq an,$$

or for large i and n

$$n\left(1-e^{-c}\right) \leq an.$$

To illustrate this, we consider in Table 1 the exact and asymptotic results for $n = 1000$ and $i = cn$, where $c = 0.1, 0.3, 0.5, 0.7, 0.9$, and $a = 0.05$.

Table 1. Proportions of unvaccinated not to be exceeded

c	0.1	0.3	0.5	0.7	0.9
Exact	0.5252	0.1928	0.1270	0.0993	0.0842
Asymptotic	0.5254	0.1929	0.1271	0.0993	0.0843

We note that these sets of figures agree up to the third decimal point; thus to keep the expected number of new infectives below 5% of the initial number of susceptibles, we need to vaccinate at least the following percentage of $n = 1000$.

c	0.1	0.3	0.5	0.7	0.9
% to be vaccinated	47.5	80.7	87.3	90.1	91.6

3 Two types of infectives and two types of susceptibles

Let us now consider the case where there are two types of infectives numbering i_1 and i_2 respectively. The i_1 will have respective probabilities p_{11}, p_{12} of preference for the n_1 susceptibles of type 1 and n_2 of type 2, with the i_2 having respective probabilities p_{21}, p_{22} of preference for the n_1 susceptibles of type 1 and n_2 of type 2, with $p_{11} + p_{12} = p_{21} + p_{22} = 1$. Then from (1), the pgf of the numbers s_{11} , s_{12} of new infectives among the suceptibles of types 1 and 2 due to $i_1 > 0$ infectives of type 1 and $i_2 = 0$ infectives of type 2 is

$$\varphi_{i_1 0;n_1 n_2}(u_1, u_2) = \sum_{j=0}^{i_1} \binom{i_1}{j} p_{11}^j p_{12}^{i_1-j} \varphi_{j\,n_1}(u_1)\varphi_{i_1-j\,n_2}(u_2), \qquad (10)$$

where $\varphi_{j\,n_1}(u_1)$ and $\varphi_{i_1-j\,n_2}(u_2)$ satisfy the difference-differential equations (2) with u_1 replacing u in the first, and u_2 replacing v in the second.

Similarly, for $i_1 = 0$ and $i_2 > 0$ infectives, the pgf of the numbers s_{21}, s_{22} of new infectives among the susceptibles of types 1 and 2 is

$$\varphi_{0\,i_2;n_1n_2}(v_1, v_2) = \sum_{j=0}^{i_2} \binom{i_2}{j} p_{21}^{j} p_{22}^{i_2-j} \varphi_{j\,n_1}(v_1)\varphi_{i_2-j\,n_2}(v_2), \tag{11}$$

where $\varphi_{j\,n_1}(v_1)$ and $\varphi_{i_2-j\,n_2}(v_2)$ satisfy the difference-differential equations (2) with v_1 replacing u in the first, and v_2 replacing v in the second.

As simple examples, we can readily calculate the following pgfs:

$$\begin{aligned}
&\varphi_{1\,0;n_1n_2}(u_1, u_2) = p_{11}u_1 + p_{12}u_2, \qquad \varphi_{0\,1;n_1n_2}(v_1, v_2) = p_{21}v_1 + p_{22}v_2,\\
&\varphi_{2\,0;n_1n_2}(u_1, u_2) = \frac{p_{11}^2}{n_1}u_1(1-u_1) + \frac{p_{12}^2}{n_2}u_2(1-u_2) + (p_{11}u_1 + p_{12}u_2)^2,\\
&\varphi_{0\,2;n_1n_2}(v_1, v_2) = \frac{p_{21}^2}{n_1}v_1(1-v_1) + \frac{p_{22}^2}{n_2}v_2(1-v_2) + (p_{21}v_1 + p_{22}v_2)^2. \qquad (12)
\end{aligned}$$

The formulae become rather unwieldy for larger values of i_1, i_2. We shall see that both the pgfs (10) and (11) are required to build up the general pgf

$$\varphi_{i_1i_2;n_1n_2}(u_1, u_2, v_1, v_2) \qquad i_1 > 0, \;\; i_2 > 0, \tag{13}$$

of new infectives s_{11} , s_{12} among the initial susceptibles n_1 of type 1 and n_2 of type 2 due to the $i_1 > 0$ infectives, and s_{21}, s_{22} due to the $i_2 > 0$ infectives. We proceed to show how this can be done.

We first define $p_{s_{11}s_{12}s_{21}s_{22}}(i_1, i_2; n_1, n_2)$ as the probability

$$P\left\{\begin{array}{l|c} \begin{array}{l} s_{11},\ s_{21} \text{ new infectives among} \\ \quad \text{the } n_1 \text{ susceptibles, and} \\ s_{12},\ s_{22} \text{ new infectives among} \\ \quad \text{the } n_2 \text{ susceptibles due} \\ \text{to type 1 and type 2 infectives} \\ \quad \text{respectively} \end{array} & \begin{array}{c} i_1,\ i_2 \text{ infectives and} \\ n_1,\ n_2 \text{ susceptibles initially} \end{array} \end{array}\right\}. \tag{14}$$

We shall assume that if a susceptible is first attacked by a type 1 infective, a later attack by either a type 1 or type 2 infective will not modify the initial type 1 infection. Likewise, a susceptible first attacked by a type 2 infective will not be modified by a later infective attack of type 1 or 2.

Note that the probabilities

$$p_{s_{11}s_{12}00}(i_1, i_2 = 0; n_1, n_2), \qquad p_{00s_{21}s_{22}}(i_1 = 0, i_2; n_1, n_2)$$

have the respective pgfs

$$\varphi_{i_1 0;n_1n_2}(u_1, u_2), \qquad \varphi_{0i_2;n_1n_2}(v_1, v_2)$$

of (10) and (11). We now derive a recursive equation for the probability (14).

It is clear that if one increases the number of infectives i_1 or i_2 by 1 with respective probabilities p and q, where $0 < p < 1$, $p + q = 1$, then one can readily derive the recursive equation for nonzero s_{11}, s_{12}, s_{21}, s_{22} as

$$\begin{aligned}
p_{s_{11}s_{12}s_{21}s_{22}}\ &(i_1, i_2; n_1, n_2) = \\
&pp_{s_{11}s_{12}s_{21}s_{22}}(i_1 - 1, i_2; n_1, n_2)\Big(p_{11}\frac{s_{11} + s_{21}}{n_1} + p_{12}\frac{s_{12} + s_{22}}{n_2}\Big) \\
&+qp_{s_{11}s_{12}s_{21}s_{22}}(i_1, i_2 - 1; n_1, n_2)\Big(p_{21}\frac{s_{11} + s_{21}}{n_1} + p_{22}\frac{s_{12} + s_{22}}{n_2}\Big) \\
&+pp_{s_{11}-1\,s_{12}s_{21}s_{22}}(i_1 - 1, i_2; n_1, n_2)p_{11}\Big(1 - \frac{s_{11} + s_{21} - 1}{n_1}\Big) \\
&+pp_{s_{11}s_{12}-1\,s_{21}s_{22}}(i_1 - 1, i_2; n_1, n_2)p_{12}\Big(1 - \frac{s_{12} + s_{22} - 1}{n_2}\Big) \\
&+qp_{s_{11}s_{12}s_{21}-1\,s_{22}}(i_1, i_2 - 1; n_1, n_2)p_{21}\Big(1 - \frac{s_{11} + s_{21} - 1}{n_1}\Big) \\
&+qp_{s_{11}s_{12}s_{21}s_{22}-1}(i_1, i_2 - 1; n_1, n_2)p_{22}\Big(1 - \frac{s_{12} + s_{22} - 1}{n_2}\Big). \qquad (15)
\end{aligned}$$

This leads directly to the difference–differential equation

$$\begin{aligned}
\varphi_{i_1 i_2;n_1 n_2}&(u_1, u_2, v_1, v_2) = \\
&p(p_{11}u_1 + p_{12}u_2)\varphi_{i_1-1\,i_2;n_1 n_2} + q(p_{21}v_1 + p_{22}v_2)\varphi_{i_1 i_2-1\,;n_1 n_2} \\
&+p\Big[\frac{p_{11}u_1}{n_1}(1 - u_1)\frac{\partial\varphi_{i_1-1\,i_2;n_1 n_2}}{\partial u_1} + \frac{p_{11}v_1}{n_1}(1 - u_1)\frac{\partial\varphi_{i_1-1\,i_2;n_1 n_2}}{\partial v_1} \\
&+\frac{p_{12}u_2}{n_2}(1 - u_2)\frac{\partial\varphi_{i_1-1\,i_2;n_1 n_2}}{\partial u_2} + \frac{p_{12}v_2}{n_2}(1 - u_2)\frac{\partial\varphi_{i_1-1\,i_2;n_1 n_2}}{\partial v_2}\Big] \\
&+q\Big[\frac{p_{21}v_1}{n_1}(1 - v_1)\frac{\partial\varphi_{i_1 i_2-1;n_1 n_2}}{\partial v_1} + \frac{p_{21}u_1}{n_1}(1 - v_1)\frac{\partial\varphi_{i_1,i_2-1;n_1 n_2}}{\partial u_1} \\
&+\frac{p_{22}v_2}{n_2}(1 - v_2)\frac{\partial\varphi_{i_1 i_2-1;n_1 n_2}}{\partial v_2} + \frac{p_{22}u_2}{n_2}(1 - v_2)\frac{\partial\varphi_{i_1 i_2-1;n_1 n_2}}{\partial u_2}\Big]. \qquad (16)
\end{aligned}$$

which holds for all $i_1, i_2 > 0$. We see that when $i_1 = 1$, $i_2 \geq 1$,

$$\varphi_{1\,i_2;n_1 n_2}(u_1, u_2, v_1, v_2)$$

will depend on $\varphi_{0\,i_2;n_1 n_2}(u_1, u_2, v_1, v_2)$ and $\varphi_{1\,i_2-1;n_1 n_2}(u_1, u_2, v_1, v_2)$. If $i_2 = 1$ as well, then the latter pgf will be $\varphi_{1\,0;n_1 n_2}(u_1, u_2, v_1, v_2)$; these are given in equation (12).

To illustrate the use of these equations, we derive

$$\varphi_{1\,1;n_1 n_2}(u_1, u_2, v_1, v_2) = p_{11}p_{21}\Big(1 - \frac{1}{n_1}\Big)u_1 v_1 + p_{12}p_{22}\Big(1 - \frac{1}{n_2}\Big)u_2 v_2$$

$$+p_{11}p_{22}u_1v_2 + p_{12}p_{21}u_2v_1$$
$$+p_{11}p_{21}\frac{pv_1+qu_1}{n_1} + p_{12}p_{22}\frac{pv_2+qu_2}{n_2}, \qquad (17)$$

and

$$\varphi_{1\,2;n_1n_2}\ (u_1,u_2,v_1,v_2) =$$
$$p(p_{11}u_1+p_{12}u_2)\Big\{p_{21}^2 v_1\tfrac{1-v_1}{n_1} + p_{22}^2 v_2\tfrac{1-v_2}{n_2} + (p_{21}v_1+p_{22}v_2)^2\Big\}$$
$$+q(p_{21}v_1+p_{22}v_2)\Big\{p_{11}p_{21}\Big(1-\frac{1}{n_1}\Big)u_1v_1 + p_{12}p_{22}\Big(1-\tfrac{1}{n_2}\Big)u_2v_2$$
$$+p_{11}p_{22}u_1v_2 + p_{12}p_{21}u_2v_1 + p_{11}p_{21}\frac{pv_1+qu_1}{n_1}$$
$$+p_{12}p_{22}\frac{pv_2+qu_2}{n_2}\Big\}$$
$$+p\Big\{p_{11}v_1\frac{1-u_1}{n_1}\Big[p_{21}^2\frac{1-2v_1}{n_1} + 2p_{21}(p_{21}v_1+p_{22}v_2)\Big]$$
$$+p_{12}v_2\frac{1-u_2}{n_2}\Big[p_{22}^2\frac{1-2v_2}{n_2} + 2p_{22}(p_{21}v_1+p_{22}v_2)\Big]\Big\}$$
$$+q\Big\{p_{21}v_1\frac{1-v_1}{n_1}\Big[p_{11}p_{21}\Big(1-\frac{1}{n_1}\Big)u_1 + p_{12}p_{21}u_2 + p\frac{p_{11}p_{21}}{n_1}\Big]$$
$$+p_{21}u_1\frac{1-v_1}{n_1}\Big[p_{11}p_{21}\Big(1-\frac{1}{n_1}\Big)v_1 + p_{11}p_{21}v_2 + q\frac{p_{11}p_{21}}{n_1}\Big]$$
$$+p_{22}v_2\frac{1-v_2}{n_2}\Big[p_{12}p_{22}\Big(1-\frac{1}{n_2}\Big)u_2 + p_{11}p_{22}u_1 + p\frac{p_{12}p_{22}}{n_2}\Big]$$
$$+p_{22}u_2\frac{1-v_2}{n_2}\Big[p_{12}p_{22}\Big(1-\frac{1}{n_2}\Big)v_2 + p_{12}p_{21}v_1 + q\frac{p_{12}p_{22}}{n_2}\Big]\Big\} \qquad (18)$$

The pgf $\varphi_{2\,1;n_1n_2}(u_1,u_2,v_1,v_2)$ can be found similarly, or by symmetry. The pgfs for higher values of i_1 ,i_2 can also be derived, but rapidly become algebraically unwieldy, though they remain manageable for relatively small values of i_1 ,i_2.

Likewise, it is possible to consider several types of infectives, as well as several types of susceptibles, but the equations then become intractable.

4 Two waves of infectives

The model in §3 is clearly difficult to handle with any ease. We shall therefore consider a simpler model extending the vaccination model of §2 in which

n_1 unvaccinated and $n - n_1$ vaccinated susceptibles are subject to two waves of infectives. In the first wave, i_1 infectives carrying a benign form of the infection exchange needles with the susceptibles. This has the effect of "vaccinating" the newly generated infectives against the second wave in which i_2 infectives, carrying the virulent form of the infection, exchange needles with the n initial susceptibles. We evaluate the expected number of newly infected susceptibles now carrying the virulent infection.

We have already seen in §2 that the expected number of newly generated infectives, now immunised against the more virulent infection of type 2 is

$$n_1\left(1-\left(1-\frac{1}{n}\right)^{i_1}\right)$$

Thus the remaining expected number of "unvaccinated" susceptibles will be

$$z_1 = n_1 - n_1\left(1-\left(1-\frac{1}{n}\right)^{i_1}\right) = n_1\left(1-\frac{1}{n}\right)^{i_1} \tag{19}$$

In exactly the same way, we can argue that when z_1 unvaccinated and $n - z_1$ "vaccinated" individuals exchange needles with i_2 virulent infectives, the expected number of newly generated infectives will be

$$z_1\left(1-\left(1-\frac{1}{n}\right)^{i_2}\right) = n_1\left(1-\frac{1}{n}\right)^{i_1}\left(1-\left(1-\frac{1}{n}\right)^{i_2}\right). \tag{20}$$

Hence the expected number of remaining unvaccinated susceptibles will be the difference between (19) and (20), or

$$n_1\left(1-\frac{1}{n}\right)^{i_1+i_2}, \tag{21}$$

precisely the same as if there had been $i = i_1 + i_2$ homogeneous infectives. Note that this is an approximate result, in terms of expected values, since z_1 may not turn out to be an integer.

To illustrate the effect of these two waves of infectives, we give an example where $n = 1000$, the initial number of unvaccinated individuals is $n_1 = 500$, and the numbers of infectives of types 1 and 2 are $i_1 = i_2 = 100$, 300 and 500. We calculate (21) to obtain:

$i_1 + i_2$	200	600	1000
Expected number of remaining unvaccinated susceptibles	409.32	274.32	183.84

REFERENCES

Gani, J. (1991). Generating function methods in a random allocation problem of epidemics. *Bull. Inst. Combinatorics Applicns* 3, 43-50.

Gani, J. (1993). Random allocation methods in an epidemic model. *Contribution to Stochastic Processes: A Festschrift in Honour of Gopinath Kallianpur*, Cambanis, S., Ghosh, J.K., Karandikar, R.L. and Sen, P.K. (editors). New York: Springer-Verlag, 97-106.

Gani, J. (2001). Needle sharing infections among heterogeneous IVDUs. To appear in *Monatsh. Math.*

Gani, J. and Yakowitz, S. (1993). Modelling the spread of HIV among intravenous drug users. *IMA J. Maths. Appl. Med. Biology* 10, 51-65.

LONG MEMORY PROCESSES - AN ECONOMIST'S VIEWPOINT

CLIVE W.J. GRANGER
Department of Economics
University of California
San Diego La Jolla
CA 92093-0508 USA
E-mail: cgranger@ucsd.edu

Fractionally integrated processes are a class of linear processes which lay between stationary autoregressive and unit root series in terms of their properties. Recent work on transformations of random walks have not been completely extended to I(d) series, but some results are available and are presented. It is noted that if the input shocks to the models have a positive mean, the process will have a non-linear trend in mean. Most economic "examples" of I(d) processes are positive series and so should have such trends, but as these are not observed, it follows that long-memory models hardly ever occur in economic. An alternative is stationary models with breaks.

Keywords: Unit root, fractionally integrated processes, non-linearity, breaking processes, trends in mean.

1 Introduction and some definitions

It is convenient to start with a stationary process Y_t, with zero mean, (constant) variance $= \sigma^2$ and autocorrelation $\rho_j^Y = corr(Y_t, Y_{t-j})$ and spectrum

$$f_Y(w) = \frac{1}{\pi} + \frac{2}{\pi}\Sigma\rho_j^Y \exp(ijw). \tag{1}$$

The process is called long memory if

$$\lim_{w \to 0} f(w) = \infty. \tag{2}$$

The definition can be extended to certain linear non-stationary processes in the following way: Suppose that X_t is a stationary series, mean zero, constant variance σ_x^2 and spectrum $f_x(w)$ with the property $f_x(0) = C$, $C > 0$. Let $g(B)$ be a polynomial in the backward operator B, which is called a filter, with the property that

$$g(1) = 0$$

so that if $g(B)X_t$ has a spectrum that is zero at zero frequency. If Y_t is generated by

$$g(B)Y_t = X_t \tag{3}$$

where X_t is stationary, with positive spectrum at zero frequencies, then Y_t can be called long-memory as the spectrum of Y_t can be written

$$f_Y(w) = \frac{f_x(w)}{|g(e^{iw})|^2} \tag{4}$$

as shown in Zygmond (1959). Note that $f_Y(W)$ has the long memory property because of the unit root property of $g(B)$. It should also be noted that a consequence of (1) and (2) is that

$$\lim_{k\to\infty} \sum_{j=1}^{k} \rho_j^Y = \infty \tag{5}$$

so that the sequence of ρ_j^Y is diverging.

From the Wold representation, ignoring any deterministic components, one has

$$X_t = \sum_{j=0}^{\infty} c_j^x \varepsilon_{t-j} \tag{6}$$

where ε_{t-j} is a zero mean, white noise (that is, serially uncorrelated) process. It follows from (3) that

$$Y_t = \sum_{j=0}^{\infty} c_j^Y \varepsilon_{t-j} \tag{7}$$

where the sequence c_j^Y declines, if at all, slowly to achieve (5).

Although there are many possible long memory processes, two have received virtually all the attention in the literature:

a *The unit root process*

$$g(B) = (1 - B) \tag{8}$$

corresponding to $f_Y(w) = cw^{-2}$ for small w and $c_j^Y =$ constant, for all j. Y_t is denoted as I(1) and integrated of order one, in this case. The I(2) process, with $g(B) = (1 - B)^2$ and $c_j^Y = cj$ has received some attention, where now $f_Y(w) = cw^{-4}$ for small w.

b *The fractional unit root process*

$$g(B) = (1 - B)^d, \qquad 0 < d < 1 \tag{9}$$

corresponding to $f_Y(w) = cw^{-2d}$, for small w. The notation $Y_t = I(d)$ is used. The properties of this process are noted below.

The phrase long-memory has little implication except that there is naturally interest in processes that have autocorrelations, and thus relationship, over long horizons. The concept of persistence is better formulated in the moving average representations (7). It tells how an old shock is remembered or still influences current values. In the I(2) model, old shocks increase in importance and dominate new shocks, which makes it less appealing for economists. For an economist, not only is the mathematics and statistical properties of a model interesting but also the interpretation in terms of real activities and the potential usefulness of the model for decision makers. This is a theme to which I will return.

It is important to distinguish between generating mechanisms that produce long memory processes and the properties that such processes have. Equations (3) and (7) are examples of generating mechanisms and equation (5) and the spectrum being unbounded at zero frequency are examples of properties. However, it should be noted that the properties do not determine the generating mechanism as several alternative generating mechanisms produce similar properties. Thus, for example, it has been common to link long memory with I(d) processes on a one-to-one form, but this is not correct as will be indicated later.

2 Fractionally integrated processes

I will consider just the properties of the process

$$(1 - B)^d Y_t = \varepsilon_t \tag{10}$$

where ε_t is a zero mean, uncorrelated process (or Martingale difference). More generally, ε_t is replaced by a stationary series X_t with a spectrum that is bounded above and is positive at the zero frequency. In this more general case, the properties of the process are quantitatively similar but more complicated to state. Details can be found in Ballie and King (1996) and Beran (1998). Using power series expansions of $(1 - B)^d$ and assuming that the process started at time $t = 0$, gives the representations:

1. *Moving Average*

$$Y_t = \sum_{j=0}^{t} c_j \varepsilon_{t-j} \tag{11}$$

where $c_j = (\Gamma(d))^{-1} j^{d-1}$ for j large.

2. *Autoregressive*

$$Y_t = \sum_{i}^{t} a_k Y_{t-k} + \varepsilon_t \tag{12}$$

where $a_k \approx (\Gamma(-d))^{-1} k^{-d-1}$. The spectrum has the form

$$f_Y(w) = (\frac{\sigma^2}{2\pi})|1 - e^{-iw}|^{-2d}$$

and for w small $f_Y(w) \approx cw^{-2d}$. The variance of the process is

$$\text{var } (Y_t) = \begin{cases} ct^{2d-1} & \text{if } d > 0.5 \\ c \log(t) & \text{if } d = 0.5 \end{cases} \tag{13}$$

both for t large, where throughout c is some constant, but var $(Y_t) \rightarrow$ finite value as $t \uparrow$ if $0 < d < 0.5$. The values of d that are ≥ 0.5 are called the non-stationary region in the literature. This is only correct if the mean of the inputs ε_t is zero. If $E[\varepsilon_t] = m \neq 0$, in (10), then

$$E[Y_t] = mt^{d-1} \tag{14}$$

and so the first moment produces non-stationarity.

If $m = 0$ and $0 < d < 0.5$, autocorrelations will exist and take the form

$$\rho_k = \text{ corr } (Y_t, Y_{t-k}) \approx ck^{2d-1}. \tag{15}$$

They are seen to decline with k, but at slower than the exponential rate observed with a stationary autoregressive process.

The class of long-memory processes so far considered, the I(d) processes, by no means exhausts the possibilities. Granger and Ding (1996) discuss some others, including using the filter

$$g(B) = [L(1 - B)^{-1}]^{\varepsilon} \tag{16}$$

where $L(x)$ is a slowly-varying function, so that for any $k \neq 0$,

$$\lim_{k \rightarrow \infty} \frac{L(kx)}{L(x)} = 1.$$

An example is $\log x$. The process generated by

$$Y_t = [L(1-B)^{-1}]^{-\varepsilon}\varepsilon_t \tag{17}$$

is called a generalized integrated process of order O^+, denoted GI(O^+). In the case where $L(x) = \log x$, the spectrum is proportional to $(\log w)^{2q}$ for small w. It follows that this process is dominated by I(d) for any $d > 0$, however small, in the sense that the spectrum of I(d) goes to infinity faster than that for the GI(O^+) process. Very little attention has been given to this process as its autocorrelations are small and it will be very difficult to detect. As any divergent series inserted into the moving average (7) will produce a long memory series, the range of possible, simple models is seen to be large.

The substantial interest in the I(d), $0 < d < 1$, processes seems to arise because of the mathematical properties of this class which lie between stationary and unit root processes and because of associated statistical problems. The low frequency part of the spectrum is particularly important and new information accumulates slowly, partly because of the long-memory, so that estimation can be slow to converge and tests have to be carefully constructed. Peter Robinson, of the London School of Economics, has many papers in these areas.

A particularly interesting case, which has been little studied, is I($\frac{1}{2}$) as it represents the bifurcation mark for moving from stationarity to non-stationarity in variance. For $d < \frac{1}{2}$ the variance tends to a constant, for $d > \frac{1}{2}$ it tends to t^{2d-1} and for $d = \frac{1}{2}$ to $\log t$. Tanaka (1999) consider the $X_t \sim J\frac{1}{2}$ process with input innovations that have zero mean, are stationary, and with finite variance and shows that

$$\frac{1}{\sqrt{\log T}} Y_t \to N(0, c_1) \tag{18}$$

and

$$p\lim \frac{1}{T}\Sigma_{t=2}^{T}(y_t - y_{t-1})^2 = c_2 \tag{19}$$

where c_1, c_2 are finite constants. If the innovations are iid with variance σ^2, then $c_1 = \frac{\sigma^2}{\pi}$ and $c_2 = 4\frac{\sigma^2}{\pi}$.

3 Generalizations

In moving from standard ARMA models to those involving unit roots is the first small step in moving from an assumption of stationarity to starting consideration of non-stationary models. This is all done with a basic assumption

of linearity of the system. Econometricians developed tests for I(0) versus I(1), with an untested background assumption of linearity. Within this framework, many economic series were deemed to be I(1) or, strictly, the null hypothesis of I(1) was not rejected against an alternative if I(0).

Confusion soon began when non-linear questions arose such as: Can a bounded variable, such as a positive variable (an interest rate, for example) be I(1)? Similarly, if X_t is I(1), will X_t^2 be I(1)? Will it be the same for other instantaneous transforms?

It should be noted that **some** aspects of the I(0) versus I(1) classification have a natural linear orientation. The most important is that the difference of an I(1) is I(0) as the difference is a linear operator, $Y_t - Y_{t-1}$. It follows that I(1) can be distinguished as being a process whose difference is I(0). A question remains, what is I(0)? It might also be noted that the properties of I(0), I(1) (and I(d)) processes are usually stated in terms of their linear properties; i.e., the means, variances, auto-covariances, and autocorrelations that arise when the process is regressed linearly on a set of past values of interest. Although these are valuable measures of the properties of a process, they are less helpful in describing a nonlinear process.

Amongst linear processes, a general definition of an I(0) process is an ARMA

$$a(B)Y_t = c(B)\varepsilon_t$$

such that $a(1) \neq 0$, $c(1) \neq 0$ so that the spectrum of Y_t at zero frequency is bounded above and is positive. This definition of I(0) does not easily translate to nonlinear univariate models. A more general definition, proposed by Davidson (1999) is: X_t is I(0) if its accumulation

$$Y_t = \sum_{j=0}^{t} X_{t-j} \qquad t = 0, \cdots, T$$

obeys the Functional Central Limit Theorem. That is, if a new variable is defined by

$$Q_T(r) = \frac{1}{\sigma\sqrt{T}} Y_{t-j} \qquad \text{if } \tfrac{j-1}{T} \leq r \leq \tfrac{j}{T}$$

(with $Q_T(1) = \frac{1}{\sigma\sqrt{T}} Y_T$, $Q_T(0) = 0$, $\sigma = var(X_t)$) so that $r < 1$, then

$$X_T(r) \to W(r) \qquad \text{as } T \uparrow \infty$$

where $W(r)$ is a Weiner process in continuous space $[0, 1)$.

Although theoretically this definition appears to be generally acceptable and to have a number of advantages there remains the substantial practical

problem that there exists no test for whether or not an accumulation Y_t obeys the functional central limit theorem. [I will not discuss the equally potential disturbing problem of how one can define a linear time series introduced by the results of Bickel and Bühlmann (1996).]

4 Transformation of random walks

If a random walk X_t is the classical form of long memory or process with persistence, then a simple transform, such as $Y_t = T(X_t) + \varepsilon_t$ may well also have long memory of some form. Park and Phillips (1999, 2001) consider various aspects of such transformations, including estimation questions for any parameters involved. The first paper considers sums of the form

$$\frac{1}{N}\sum_{t=1}^{n} T(X_t) \qquad \text{as } n \uparrow \infty \tag{20}$$

where N is a suitable normalization. The results depend on the property of the function $T(x)$ and require the use of mathematics different from that used with linear random walks. Three classes of functions are considered:

(i) Class I (integrable)

$$\int_0^t T(n)dx \qquad \text{finite for any } t; \tag{21}$$

(ii) Class H (homogenous)

$$T(\lambda x) = v(\lambda)H(x) + R(x, \lambda) \tag{22}$$

where $H(x)$ is integrable;

(iii) Class E (exponential)

$$T(x) = E(x) + R(x)$$

where $E(x)$ is monotonic.

In (ii) and (iii) the remainder term R is dominated by the first term for $|x|$ large. (These are not the precise definitions used.) It should be noted that some functions do not fall into these classes.

The new concept needed is that of local time, $L(t, s)$ which is the probability of a Brownian motion $W(r)$, $0 \leq r \leq 1$, falling into the small region around s given by

$$L(t, s) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t I_{\{|W(r)-s|<\varepsilon\}} dr \tag{23}$$

so that, in the limit $W(r)$ is near s. One get the useful property

$$\int_0^t T(W(r))dr = \int_{-\infty}^{\infty} T(s)L(t,s)ds \tag{24}$$

if T is locally integrable. An example of the results obtained is:

If $T(x) \sim H$, then

$$\lim_{n\to\infty} \frac{1}{nv(n)} \sum_{t=1}^{n} T(X_t) \stackrel{d}{=} \int_{-\infty}^{\infty} H(s)L(t,s)ds \tag{25}$$

where, in the integral, $H(s)$ is not random but $L(t,s)$ is. If $T(x) \sim E$, the equivalent result is found to be path-specific.

The results, when applied to the question of estimation of α by OLS in the model

$$Y_t = \alpha T(X_t) + u_t$$

where u_t is stationary, gives specific results for bias which are stochastic and have a specific normalization.

As might be expected, it is shown that to obtain mathematical results even in fairly simple simulations becomes considerably more complicated when functions of random walks are considered.

An alternative approach is taken by Karlson, Mykelbust, and Tjostheim (1999). They consider a class of relationships

$$Z_t = f(X_t) + W_t$$

where X_t is persistent so that $f(X_t)$ is also persistent and thus Z_t also. W_t is stationary, such as a moving average. Here persistence is defined as null-recurrent. In a Markov chain some states can occur once and then there is a zero probability that it will re-occur, then this state is null-recurrent. For example, a random walk with continuous iid inputs, the probability of returning to any particular value is zero, unlike a stationary series with mean reversion, which will return to states near the mean often.

However, even though a persistent process, such as a random walk does not recur, it may come close occasionally to a previous state. Denote s to be the recurrence time, from the initial state to the process being close to it again, in some sense.

Suppose

$$Prob(s > n) = \frac{n^{-\beta}}{c}.$$

For a random walk $L = \frac{1}{2}$, but L will take other values for different processes. The authors show that the function $f(X_t)$ can be well estimated using a non-parametric kernel of the form

$$K_{h_n}(v) = h_n^{-1} K(u/k_n)$$

where, for example, $K(u) = 0.75(1 - u^2)$, $-1 \leq u \leq 1$. The kernel provides a window through which the underlying curve can be viewed and then estimated. The width is determined by h_n which is taken to be $-1 \geq k^{\beta/2}$, where _ is given above. It is shown that if

$$\hat{\rho} = \frac{\sum_{t=0}^{n} K_{h_n}(X_t) Z_t}{\sum_{t=0}^{n} K_{hn}(X_t)}$$

then

$$\left[h_n \sum_{t=0}^{n} K_{h_n}(X_t) \right] \left[\hat{f}(x) - f(x) \right]^d \rightarrow N(0, \sigma_w^2 \int k^2(u) du) \qquad \text{as } n \rightarrow \infty$$

where σ_w^2 is the variance of the shock and the final term is a constant depending on the shape of the kernel. The result suggests that a non-parametric technique should work well in the estimation of a nonlinear relationship between persistent variables.

5 Transformations of fractionally integrated processes

Although considerable progress has been made with functions of unit root and other persistent processes, only incomplete theoretical results, complimented by simulation results, are available for transformations of I(d) processes. The results are for Gaussian processes and the available theorem is:

Theorem 1: If $X_t \approx I(d)$, $0 < d < \frac{1}{2}$ and if $g(.)$ is a univariate transformation that can be written as the finite sum of Hermitic polynomials $H_j(.)$;

$$g(x) = g + \sum_{j=J}^{k} g_j H_j(x)$$

with $I \leq J < k < \infty$ and $g_J \neq 0$. Here J is called the Hermitic rank. Then $g(x_t)$ is a long memory process is a long memory process I(d) with $\bar{d} = \max(0, (d - 0.5)J + 0.5)$.

Some illustrations:

This table shows $\bar{d}$ values for transformed series

$g(x)$ and rank	Original $d = 0.2$	Original $d = 0.4$
X (Rank 1)	0.2	0.4
X^2 (Rank 2)	0	0.3
X^3 (Rank 1)	0.2	0.4
X^4 (Rank 2)	0	0.3
$X^3 - 3X$ (Rank 3)	0	0.2
$X^4 - 6X^2$ (Rank 4)	0	0.1

The theorem and simulations supporting the above theoretical results can be found in Dittman and Granger (2002). The theorem was originally shown in a different from by Taqqu (1979) who also established a functional central limit theorem for $g(x_t)$.

For I(d), $0.5 < d < 1$ mainly simulation results are available in Dittman and Granger (2002) but it is shown that X^2 is I(d) asymptotically (at least an empirically equivalent form of I(d)). This suggests that this is at least approximately true for the other transforms shown in the above table where the estimated d value for the transformed series is slightly below the d of the original series for $d = 0.6$ and 0.8, although there does appear to be a slippage as the rank increases. For example:

	Estimated d	
	Original $d = 0.6$	Original $d = 0.8$
X^3 (Rank 1)	0.55	0.78
$X^3 - 3X$ (Rank 3)	0.54	0.77
$X^4 - 6X^2$ (Rank 4)	0.50	0.75

It is clear that the Hermitic rank is having very much less influence for $d > 0.5$ than it did for the $d < 0.5$ case.

Simulations were also conducted for a few transcendental transformations, $\sin(x)$, $\cos(x)$, $\exp(x)$, and the logistic $= (1 + \exp(X^{-1}))^{-1}$. The first three did not give monotonically increasing estimates of d as the d of the input series increased from 0 to 1, whereas d for the logistic was close to the input d in all cases except for $d = 1$ where it was rather low (around 0.8).

6 The economists' viewpoint - empty boxes

I believe that it is relevant to ask the questions: why do we, or should we, study long-memory processes? There is surely the intellectual curiosity in working on unsolved questions, particularly if they prove to be mathematically interesting. To econometricians, the area is potentially important, we expect

the economy to produce series that are both non-stationary and non-linear, and this is an immense and largely unexplored area. The papers by Karlson et al. (1999) and by Park and Phillips (2001) and some others, are virtually the first attempts to enter this difficult but promising section of the research map.

Parts of this research involve the introduction of specific techniques and models, such as the fractionally integrated or I(d), which is then investigated in terms of its properties, estimation methods, and how to relate to other models. However, as an economist, there is a pragmatic criterion of considerable importance that needs to be considered - are there any examples in the actual economy where this model occurs? One can always argue that the model has some interesting, sensible, properties but if it never happens in reality it becomes merely a toy, something to play with in one's spare time. Situations or theories on topics which never occur are said to belong in an empty box, an idea which is described in an entry with that name in the *New Palgrave Dictionary of Economics,* (McMillan, 1987).

A case can be made that the I(d) process is an empty box in economics. It should be noted that to estimate d sufficiently well so that the capturing values of 0 and 1 can be rejected requires a large sample, too large for most time series in economics apart from those from speculative markets. By far the strongest evidence, in terms of autocorrelations and spectral shapes as well as estimates of d, come from consideration of absolute returns, possibly raised to some power, $|r_t|^\theta$. Using stock market returns, commodity prices, exchange rates, and interest rates, significant d values were often found and long, positive, and significant autocorrelation sequences indicated. All of the second-moment evidence clearly suggested an I(d) process, but ignores the first-moment. If X_t is a positive series and is I(d), it will have a trend in mean of order ct^d, where c is a constant. For a sample of several thousand, such a trend will be visible but none of the absolute return series in finance contain such a trend. The conclusion is that these best-examples cannot be fractionally integrated. As there are few, if any, other compelling examples it seems that for economists I(d) belongs in the empty box. I have no opinion about other series but the evidence for positive series and lack of trend in mean needs to be considered.

An alternative model which produces the same covariance long-memory results is a short-memory stationary series with occasional random breaks in mean (random in location, size, and sign). The unit root model still remains in active duty in economic and econometrics, although its use with positive series has not been discussed explicitly, to the best of my knowledge.

REFERENCES

Baillie, R.F. and King, M.L. (1996). Fractional differencing and long memory process. Special issue of the *Journal of Econometrics* 73, 1-24.

Beran, J. (1998). *Statistics for Long-Memory Processes.* New York: Chapman and Hall.

Bickel, P.J. and Bühlmann, P. (1996). What is a linear process? *Proceeding of the U.S. National Academy of Science* 93, 12128-12131.

Davidson, J. (1997). Definitions of I(1), I(0). Unpublished working paper, Department of Economics, University of Cardiff.

Dittman, I. and Granger, C.W.J. (2002). Properties of nonlinear transformations of fractionally integrated processes. To appear *Journal of Econometrics.*

Granger, C.W.J. and Ding, Z. (1996). Varieties of long memory models. *Journal of Econometrics* 73, 61-78.

Karlson, H.A., Myklebust, T. and Tjostheim, D. (1999). Nonlinear estimation in a nonlinear cointegration type model. Working paper, *Department of Statistics, University of Bergen, Norway.*

Park, J.Y. and Phillips, P.C.B. (1999). Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory* 4, 468-497.

Park, J.Y. and Philips, P.C.B. (2001). Nonlinear regressions with integrated time series. *Econometrica* 69, 117-162.

Tanaka, K. (1999). The non-stationary fractional unit root. *Econometric Theory* 15, 549-582.

Taqqu, M. (1979). Convergence of integrated processes of arbitrary hermitic rank. *Probability Theory and Related Fields* 50, 53-83.

Zygmund, A. (1959). *Trigonometric Series*, 2nd edition. Cambridge University Press.

AN INDIFFERENCE ZONE APPROACH TO TESTING FOR A TWO-COMPONENT NORMAL MIXTURE

MICHELE HAYNES

School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD 4000, Australia
email: m.haynes@fsc.qut.edu.au

KERRIE MENGERSEN

School of Mathematical and Physical Sciences, University of Newcastle, NSW 2308, Australia
email: kerrie.mengersen@newcastle.edu.au

We focus on a test for a two-component Normal mixture model f by assessing its "closeness", in terms of L2 distance, to a Normal distribution f_N. If the L2 estimate is in a specified "indifference zone", the two distributions are asserted to be "close enough" and by parsimony f_N is preferred over f. This approach is applicable in both Bayesian and frequentist frameworks and is easily generalised to a test for more than two components in a mixture. Results are extended to model selection using predictive densities under a Bayesian approach and are illustrated by example.

Keywords: Bayesian approach, indifference zone, L2 distance, model selection, Normal mixture distribution, predictive distributions.

1 Introduction

Testing for the number of components, k, in a mixture of the form $\Sigma_{i=1}^{k} p_i f_i(x)$ is difficult because of the geometry of the parameter space and the number of expressions that contribute to a null hypothesis. Popular approaches to testing include likelihood methods, Bayes factors and reversible jump MCMC (Richardson and Green, 1997). An alternative test based on the distributional distance between two models was proposed by Mengersen and Robert (1996). Here, the null hypothesis H_0 of no mixture is accepted if $d(f, f*) \leq \alpha$, where $d(f, f^*)$ is the Kullback-Leibler (KL) distance between a two-component mixture f and a Normal distribution f^*, and α is a pre-specified "indifference zone" (IZ, Gupta and Panchapakesan, 1979). The IZ may be considered to be the maximum distortion that we are prepared to allow in the shape of a Normal distribution, due to an additional component, before we select a two-component Normal mixture as a more adequate distribution.

A noted difficulty with the approach of Mengersen and Robert (1996), however, was the failure to derive a closed form expression for the KL distance which led to difficulties in developing the IZ. In the present paper we propose a remedy to this problem by considering the use of the L2 instead of the KL distance measure. The derived test is applicable in both frequentist and Bayesian paradigms. Although it is developed in the context of a two-component Normal mixture, it is easily extended to situations of general k and non-Normality. In a Bayesian setting, we investigate an adaptation of the test based on posterior predictive distributions which are arguably more interesting in this context.

2 Testing for a mixture

Without loss of generality, consider a standardised two-component Normal mixture $f = pN(0,1) + (1-p)N(\theta, \sigma^2)$, $\theta \geq 0, \sigma \geq 1, 0.5 \leq p \leq 1$. Denote by $f_N = N(\epsilon, \omega^2)$ the closest Normal distribution in that ϵ and ω^2 are derived to minimise the L2 distance given by

$$\begin{aligned} L2(f, f_N) = \frac{1}{\sqrt{2\pi}}[\frac{1}{\sqrt{2}}(\frac{1}{\omega} + p^2 + \frac{(1-p)^2}{\sigma}) \\ - \frac{2p}{\sqrt{\omega^2+1}} exp[\frac{-\epsilon^2}{2(\omega^2+1)}] \\ - \frac{2(1-p)}{\sqrt{\omega^2+\sigma^2}} exp[\frac{-(\epsilon-\theta)^2}{2(\omega^2+\sigma^2)}] \\ + \frac{2p(1-p)}{\sqrt{1+\sigma^2}} exp[\frac{-\theta^2}{2(1+\sigma^2)}]]. \end{aligned} \tag{1}$$

The behaviour of $L2(f, f_N)$ over the region defined by $-3 \leq \theta \leq 3$ and $0.1 \leq \sigma \leq 20$ is illustrated in Figure 1 for $p = 0.5$. Similar behaviour is observed for other values of p, although as p increases, there is a slow down in the rate of increase in $L2(f, f_N)$ with increasing $|\theta|$ and σ.

It remains now to define an IZ based on $L2(f, f_N)$. Below we consider two approaches. Whereas the first is appropriate in both Bayesian and frequentist settings and provides a binary decision rule for H_0, the second builds on the Bayesian paradigm to provide a posterior probability of H_0.

2.1 Approach 1

In order to test H_0, it is necessary to construct a reference measure. By considering the L2 distances (2) between the standardised mixture and the

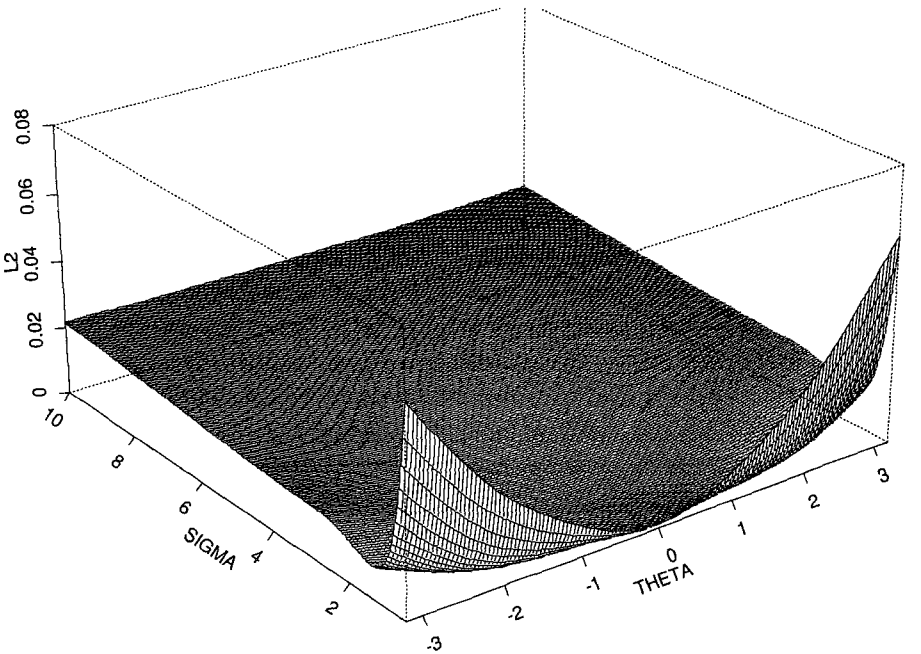


Figure 1. L2 distance between 2-component Normal mixture and closest normal with p=0.5

standard normal distribution, f_{SN}, as the parameters p, θ and σ vary, the boundary of the IZ could represent a measure of the maximum amount of distortion that we are prepared to allow in the shape of the standard normal distribution due to an additional component, before we accept f as the more adequate distribution.

$$L2(f, f_SN) = \int (f - f_N)^2 dx$$
$$= \frac{1}{\sqrt{2\pi}}(p-1)^2 \left[\frac{1}{\sqrt{2}} \left(\frac{1+\sigma}{\sigma} \right) - \frac{2}{\sqrt{1+\sigma^2}} e^{-\theta^2/2(1+\sigma^2)} \right]. \quad (2)$$

Here we consider the volume V under the $L2(f, f_{SN})$ region from $\min(L2(f, f_N))$ where $(p = 1, \theta = 0, \sigma = 1)$ to an arbitrary $(p_0, \theta_0, \sigma_0)$, defined

by

$$V(p_0, \theta_0, \sigma_0) = \int_{p_0}^{1} \int_{0}^{\theta_0} \int_{1}^{\sigma_0} L2(f, f_N) d\sigma d\theta dp \ . \tag{3}$$

Under H_0, interest is focused on small $L2$ values. For fixed p, the maximum region of interest in $L2$ may thus be defined as extending from $\min(L2(f, f_N))$ to values of θ and σ corresponding to the point of inflexion in V . For all p and fixed σ, the value of θ at the inflexion in V occurs at $\theta_0 = \sqrt{1+\sigma^2}$. It can be shown that at this point of inflexion $\sigma_0 = 1.69$. Hence from (3), $\max(V) = V(0.5, 1.96, 1.69) = 0.0081$.

In the simplest case, the test of the null hypothsis can be defined in terms of V, with the IZ arbitrarily nominated as $\alpha_V = cV$, $0 < c < 1$. The test may then be described as follows.

Step 1 Decide on IZ by nominating α_V.

Step 2 Estimate $f = \hat{p}N(0,1) + (1-\hat{p})N(\hat{\mu}, \hat{\sigma}^2)$ and the closest Normal $f_N = N(\hat{\epsilon}, \hat{\omega}^2)$. Compute $L2(f, f_N)$.

Step 3 Compute the corresponding $V^* = V(\hat{p}, \hat{\theta}, \hat{\sigma})$ from (2) and (3).

Step 4 If $V^* < \alpha$, $L2$ is in the IZ so the null hypothesis of no mixture is accepted and f_N is chosen. Otherwise f is chosen.

Note that V^* does not have an analytic solution but can be computed numerically in the same procedure that estimates $\hat{\epsilon}$ and $\hat{\omega}$. Other decision rules can be constructed from V, according to appropriate loss functions.

2.2 Approach 2

For the mixture f, consider a joint prior distribution, p_P say, on (p, θ, σ). A prior distribution, $L2_P$ say, may be constructed by drawing samples from p_P, deriving f_N corresponding to the realised f and calculating $L2(f, f_N)$. Imposition of a prior on H_0 of 0.5, which conforms to the usual application of Bayes Factors, provides an obvious choice for α as $\alpha_P = \textit{median } L2_P$.

The posterior distribution $\pi = \pi(p, \theta, \sigma | X, f)$ may be estimated numerically using Markov chain Monte Carlo (MCMC) for example. After appropriate burn-in, at each iteration, a realisation of (p, θ, σ) is drawn from π, from which the posterior f, f_N and $L2$ may be derived. The result is a sequence of posterior estimates $L2^{(1)}, L2^{(2)}, ..., L2^{(i)}, ..., L2^{(m)}$, where m is arbitrarily large.

A test of H_0 may be described as follows.

Step 1 Determine α_P based on p_P as described above.

Step 2 Simulate posterior estimates of $L2^{(1)}, L2^{(2)}, ..., L2^{(i)}, ..., L2^{(m)}$ as described above.

Step 3 Estimate $\Pr(L2(f, f_N) \leq \alpha_P)$, by $\Sigma_{i=1}^{m} I(L2^{(i)} \leq \alpha_P)$, where I is the indicator function.

Note that the outcome from this test is the posterior probability of being in IZ. This is argued to be more informative than Approach 1 in making choices between f and f_N. The price of this added information is computational, both in the derivation of α_P and the computation of π.

3 Testing predictive distributions

In a Bayesian context, predictive distributions may be more comparable than posterior distributions for model assessment (Box, 1980; Berger, 1985). Gelfand et al. (1992) consider issues of model adequacy and model choice for predictive distributions using cross-validation methods. Gelfand and Dey (1994) provide a definition for a general predictive density and compare different forms of Bayes factors that can be used for model choice.

Although the $L2$ distance between the predictive densities g and g_N say, of a Normal distribution and a two-component Normal mixture, respectively, can be written down, it does not behave in such a fashion as to allow the construction of a closest predictive Normal distribution as in Section 2. For example, the mean of g_N may not be centred about the weighted mean of the two components of g when the sample mean is close to the prior mean.

Two alternative approaches are considered here. First, with conjugate priors on the parameters of the two-component mixture model, the predictive density g is a four-component mixture of t-distributions. This might reasonably be approximated by a Normal density g', say. In this case we can proceed as in Section 2.1 or 2.2 by constructing the closest Normal density g'_N to g' and the corresponding $L2(g', g'_N)$.

A second approach avoids this approximation and such a heavy reliance on the choice of the prior distribution on g. Here we follow McCullogh and Rossi (1992)who proposed a method for computing a Bayes Factor which compares the expected values of the likelihood with respect to a prior over the restricted and unrestricted parameter space, respectively. The method proceeds by imposing a prior p_g on g and then inducing a prior $p_g N$ on g_N by projecting draws from the space of p_g onto the restricted space of $p_g N$. Here, the minimum $L2(g, g_N)$ is defined as the projection function.

This approach is highly computational and indeed leads naturally into the more elegant Reversible Jump MCMC algorithm described by Richardson and Green (1997).

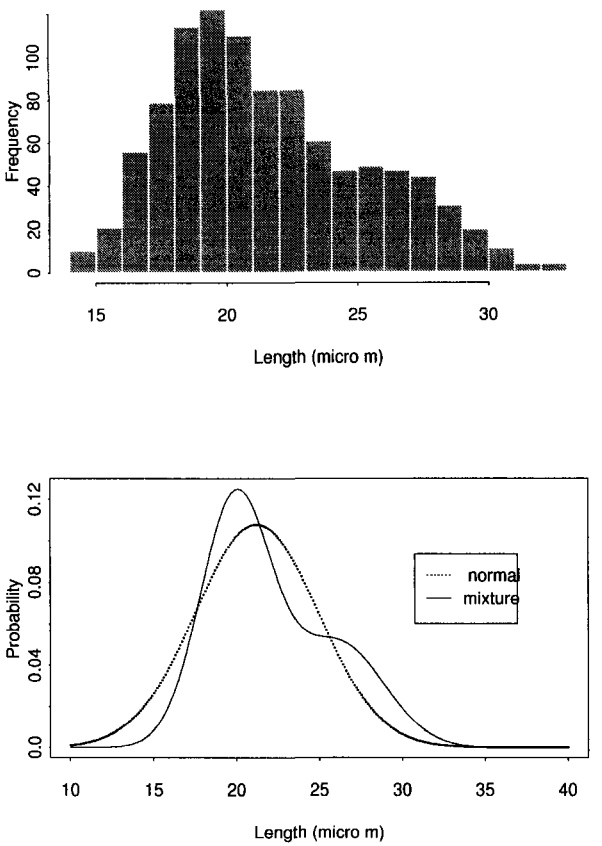


Figure 2. (Top) Histogram for protozoon data (Bottom) Estimated mixture density 0.65N(19.96, 4.62) + 0.35N(26.16, 7.62) and closest Normal N(21.21, 13.71)

4 Example

Consider a dataset introduced and studied by Pearson (see Everitt and Hand, 1981) which contains the length in micrometers of 1000 parasitic protozoon called trypanasome. The data comprise measurements from two strains of protozoon. It is of interest to assess whether the data are best described by a two-component Normal mixture or can be adequately summarised by a Normal distribution. Figure 2 depicts a histogram of the data, the estimated mixture density $0.65N(19.96, 4.62) + 0.35N(26.16, 7.62)$ and its closest Normal density $N(21, 21, 13.71)$ in terms of $L2$ distance. The respective standardised distributions are $f = 0.65N(0, 1) + 0.35N(2.89, 1.65)$ and $f_N = N(0.58, 2.97)$, with an $L2$ distance of 0.0104.

Using Approach 2 in Section 2, values for (p, θ, σ) in f were generated from priors $p \sim U[0.5, 1], \theta \sim U[0, 3], \sigma \sim U[1, 3]$. The upper limits of the intervals are reasonably arbitrary in this example, although sensitivity analyses (not shown here) indicated a high degree of robustness in decision-making. Note that if the upper limits are increased, α will also increase.

With 10,000 simulations from the prior distribution, the induced value of $\alpha_P = \text{median}(L2_P)$ was 0.0018. Since $0.0104 > 0.0018$, the observed $L2(f, f_N)$ is not in the IZ. We conclude that the data are not reasonably described by a Normal distribution and hence we choose the two-component Normal mixture.

REFERENCES

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis.* Second edition. New York: Springer-Verlag.

Box, G. (1980). Sampling and Bayes' inference in scientific modelling and robustness. *J. Royal Statist. Soc.* A 143, 383-430.

Everitt, B. and Hand, D. (1981). *Finite Mixture Distributions.* London: Chapman and Hall.

Gelfand, A., Dey, D. and Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods. In *Bayesian Statistics 4*, Bernardo, J., Berger, J., Dawid, A., Smith, A. (editors), 147-167, Oxford University Press.

Gelfand, A. and Dey, D. (1994). Bayesian model choice: asymptotics and exact calculations. *J. Royal Statist. Soc.* B 56, 501-514.

Gupta, S. and Panchapakesan, S. (1979). *Multiple Decision Procedures: Theory and Methodology of Selecting and Ranking Populations.* New York: John Wiley.

McCullogh, R. and Rossi, P. (1992). Bayes factors for nonlinear hypotheses and likelihood distributions. *Biometrika* 79, 663-676.

Mengersen, K. and Robert, C.P. (1996). Testing for mixtures via entropy distance and Gibbs sampling. In *Bayesian Statistics 5*, Berger, J., Bernardo, J., Dawid, A., Lindley, D., and Smith, A. (editors), 255-276, Oxford University Press.

Richardson, S. and Green, P. (1997). On Bayesian analysis of mixtures with an unknown number of components. *J. Royal Statist. Soc.* B 59, 731-792.

RISK MODELS FOR INDIVIDUAL DECISION MAKING

MALCOLM HUDSON
Department of Statistics, DEFS
Macquarie University, NSW 2109, Australia
E-mail: Malcolm.Hudson@mq.edu.au

ALLAN LANGLANDS AND JOHN SIMES
NHMRC Clinical Trials Centre
University of Sydney, NSW 2050, Australia

A decision analysis framework for risk models for health state transitions and valuations is described. Emphasis is placed on the sensitivity of choices among treatments to the absolute level of risk and the risk profile of the individual. The uses of Clinical Trials databases to inform decision aids is considered. We employ Markov models with Q-TWiST valuation of health states histories and examine the requirements for effective valuations of treatment alternatives for the individual.

Keywords: decision analysis, breast cancer, survival, trade-offs, quality of life

1 Introduction

The better choice of treatment for a health condition depends on the health experience that follows it, together with the individual valuation of that health state history. Since any future course of health states or health events is uncertain we cannot know exactly what history will eventuate. It may be useful for a new individual making a choice among treatment options to be presented with simulations of future health states selected from health state histories that have occurred in similar individuals on the same treatment in clinical trials. These realizations determine corresponding simulated histories for the new individual. The treatment option valued best on average might be preferred (leaving aside cost considerations).

Even in life threatening conditions, factors other than survival gains may be important in the decision. This decision may depend on individual circumstances, age and risk profile attributes. The risk profile determines a net valuation accounting for benefits, costs and harms. Risk profile attributes determine:

- how trade-offs between different dimensions in outcomes are valued by doctor and patient; e.g. willingness to trade-off quality of life or other

treatment effects for increased survival.

- the relative valuations of alternative health states;
- willingness to entertain the possibility of prolonged low quality states in anticipation of longer term benefits in health;
- the discounting applied to the future as against the present.

Uncertainty as to future health states leads to uncertainty in the valuation which is best allowed for in repeated simulations of future histories (or realizations) applicable to the new individual. Different valuations may lead to different treatment preferences among patients.

In two examples considered below, clinical trials data is available and further population data (e.g. cohort data) may be gathered to provide parameters necessary for decision making informing appropriate treatment choices for a new individual. Our goal is to develop a decision framework that uses survival, quality of life and risk profile data from clinical trials to assist choice among treatment options. The method might find use in providing valuations from the perspective of an individual or of the community.

With event rate data (occurrence or non-occurrence of nominated events during some fixed time interval), crucial parameters in the decision are the absolute risks for *this* individual and relative risks of treatment outcome (the increases in event rates of a 'benefit' such as a cure, and the increase in event rate of a 'harm' such as a toxicity of treatment) may be determined in the trial. In many cases a risk ratio or relative risk is estimated. Reduction in the risk of an adverse event, a benefit, can be offset by potential increased occurrence of harms, such as side-effects and treatment toxicity, creating a tradeoff between potential benefits and harms.

Example Prophylactic warfarin to prevent stroke

Prophylactic warfarin is a treatment option for preventing stroke among individuals at high risk of stroke (e.g. following an initial stroke). Controlled clinical trials can provide event rates in groups prescribed warfarin relative to *event rates* in untreated individuals. While warfarin carries a benefit in reducing fatal stroke, it also carries harms in (i) increasing the event rates of non-fatal but disabling stroke; (ii) introducing a risk of death due to fatal bleeds, a side effect of warfarin. When the rates of non-fatal stroke are in proportion to rates of fatal strokes we can expect a net benefit provided fatal stroke is considered to be a worse outcome than disabling stroke *and* that provided the risk of bleeding is not too great. The risk of bleeding is, however,

largely unrelated to the absolute risk of stroke. A decision as to the net benefit for an *individual* of warfarin treatment will depend on the tradeoff between benefits and harms. Decision analysis (see e.g. Kirkwood, 1997, Chapter 4) will allocate utilities to each outcome (fatal stroke, disabling stroke, fatal bleed, recovery) and using known risk probabilities the expected utility is calculated for each treatment option:

$$E(U) = \sum u_i \, p_i$$

where the sum is over all outcomes i. This calculation requires not only eliciting the utility relativities $\{u_i\}$ of the individual but also the probabilities of each potential outcome for this individual. A key point is that the comparison of expected utilities will depend on *relative risk* of stroke (assessed from clinical trials of warfarin) and the *absolute risk* of stroke for this individual (if untreated).

Table 1 provides an example of this calculation. The category "No stroke" refers to neither stroke nor bleed. In the Table $p_3 = 1 - p_1 - p_2$, $p'_3 = 1 - p'_1 - p'_2 - p'_4$ and f is the risk ratio of total or disabling stroke for warfarin relative to no treatment. We assume no change with Warfarin treatment in the distribution of outcomes among individuals experiencing fatal or disabling stroke.

Table 1. Utility calculations for Warfarin therapy following stroke

Outcome	Utility	Probability		Utility increment
i	U_i	Untreated	Warfarin	$U_i \Delta p_i$
Fatal stroke	0	p_1	$p'_1 = f\, p_1$	0
Disabling stroke	0.25	p_2	$p'_2 = f\, p_2$	-0.25 $p_2(1-f)$
No stroke	1	p_3	p'_3	$(p_1+p_2)(1-f) - p'_4$
Fatal bleed	0	0	p'_4	0

Suppose Warfarin has a risk ratio $f = 0.33$ and that the baseline risk, without Warfarin, is 6 fatal strokes and 12 disabling strokes per thousand. Then the events rates with Warfarin are 2 fatal strokes and 4 disabling strokes per thousand. With 10 fatal bleeds per thousand, the two alternative treatments have almost equal expected utility (or average quality of life valuation). For this event rate of fatal bleeds, Warfarin would be preferred for an individual at higher risk of stroke than the average in this population. On the other hand, if the individual's risk is less than average withholding treatment would be preferred.

Note that we do *not* require an assumption that the new individual is a representative member of the trial patient population. It can be demonstrably untrue that a population of individuals similar on all measured prognostic risk factors to the control population of the clinical trial has the same event rates experienced in the trial. Some trials show a similar effect to the well-known 'healthy worker effect' in epidemiology. Accrual rates suggest trial participants are *healthier* than a comparable individual in the general population. In other trials, where treatment may carry severe side effects for uncertain benefit, entry to the trial may be highly selective toward those with poor prognosis. While our analysis does not assume event rates from the trials are indicative, an assumption of homogeneity in the risk reduction within a comparable trial group is required, i.e., the risk ratio (comparing treatment with control) for the new individual should remain unchanged from the risk ratio of a participant in the trial (or the risk ratio determined from meta-analysis of related trials).

In many context there will be prognostic factors that adjust the absolute risk of an outcome, and therefore the appropriate decision. For example, in the treatment of stroke there are many other factors that play an important role in deciding the treatment, e.g. age. A number of drugs are known to interact with warfarin. There are also disease conditions where use of warfarin might prove harmful, e.g. liver disorders, vitamin K deficiency, thrombocythaemia. Clinical trials can generally provide prognostic indices which may be used to adjust baseline risk (as well as the estimates of relative effects). Prognostic factors will *not* normally affect the essential methodology or vitiate the assumption made above. Analysis needs only uniformity (homogeneity) of treatment effect as provided in relative risks, relative hazards or risk ratios. That is, we require absence of effect modification in subgroups. Treatment decisions based on these estimates may be expected to apply well under conditions for which warfarin therapy is considered good practice (e.g. those defining trial eligibility).

If the assumption of homogeneity is unsatisfactory, the relative risk used in the decision analysis must be adjusted according to the subgroup of the new individual.

The following section introduces a second example of a more complex decision analysis, and methodology that applies to health state progression.

2 Management decisions in early breast cancer

Consider now a management problem in the treatment of breast cancer. This decision analysis is not only important in its own right, e.g. Langlands (1983),

NIH Consensus (1983), but argues the importance to decision making of clinical trials including data on quality of life and risk taking behaviour.

Management of early breast cancer

Women diagnosed with early breast cancer are entitled to choose between mastectomy and breast conservation if they wish to be involved in the management decision.

Mastectomy (often referred to as total mastectomy) involves removal of all breast tissue and some of the breast skin carrying the nipple areola complex. At the same time and through the same incision the glands in the axilla are removed. At present, studies are in progress to try to determine whether it is necessary to remove the glands in every case. Women undergoing mastectomy will find an average of up to seven days in hospital, during which time the wound will require drainage. Once discharged a woman can expect to be back to more or less normal activity in 4-6 weeks.

When the breast is conserved only the tumor with a small margin of normal tissue is removed through a small incision directly over the tumor. The main criterion for this procedure is that doctors agree that this excision will be possible and will leave a good cosmetic result. For the procedure to have the highest chance of success the pathologist must confirm that the excision has removed all the tumor with a (safety) margin of normal tissue. Occasionally it is necessary to undergo a second excision to obtain that margin. In breast conservation the remainder of the breast tissue has to be treated to equate with its removal in mastectomy. The remainder of the breast is treated by radiation.

The course of radiation therapy following breast conservation usually takes 6-7 weeks. A daily visit to the hospital is required for scanning and treatment on 30-35 occasions. Treatment of the breast takes less than 15 minutes and side-effects are non life-threatening.

Follow-up of the conserved breast is done clinically and by mammography. Early detection of recurrence allows its treatment by mastectomy without survival being compromised. The risk of recurrence, leading to mastectomy is one case per hundred women per annum. This means that a woman choosing breast conservation has a 90% chance of avoiding mastectomy over the next ten years.

Management decisions as to form of surgery will involve survival comparisons and further considerations summarised in Table 2.

Once the breast has been treated, whether by mastectomy or conservation, a second treatment decision is necessary. That decision is based on the pathological examination of the breast tumor and the lymph glands from the

Table 2. Breast cancer management considerations in early stage disease

Outcome	Considerations
Mastectomy	Short overall treatment time
	No necessity for radiation unless stage 2 disease
	Body image and sexuality may be strong considerations
	Cosmetic result worse than for a conserved breast
Conservation	Body image and sexuality less adversely affected
	Longer overall treatment time and inconvenience
	Necessity for radiation therapy
	Necessity for regular mammographic checks for recurrence
	Occasional requirement for a second operation

axilla. It is unaffected by the woman's choice of mastectomy or conservation.

This second decision is whether further treatment by chemotherapy and/or hormone therapy is necessary. Where there is major involvement of the glands from the armpit additional radiation may be advised equally for mastectomy and breast conservation cases. See references contained in Jefford et al. (2001) and Shapiro and Recht (2001).

Role of decision analysis in breast cancer management

Informed decision making finds application in assisting clinicians to advise premenopausal women whose tumour offers the prospect of breast conservation on their decision as to treatment (Liberati, 1987).

With treatment options as described above – mastectomy or breast conserving surgery – expected survival would appear to be very similar following breast conservation, though reduced by the small attendant risk of cancer recurrence in the breast not being detected sufficiently early. Apart from this small risk, the comparison of benefits will depend on the risk of a second operation, and balance benefits in body image against duration and convenience of treatment as well. Note highly subjective attitudes, such as those to gambling and risk taking, will be significant.

The second stage decision as to whether to initiate adjuvant chemotherapy and/or hormonal therapy requires knowledge of toxicity, disease free survival, relapse durations and quality of life. A well established natural history and clinical progression applies in breast cancer, following key events (nodal involvement, tumour recurrence, etc.) in the disease process. See Carter et al. (1989), Koscielny et al. (1989).

Data on which to base the decision analysis is available in a number of

current studies being conducted by ANZ BCTG and NHMRC Clinical Trials Centre examining (i) axillary nodes vs. sentinal node disection in surgery, (ii) conserving breast surgery with and without radiotherapy, or (iii) use of adjuvant chemotherapy [ANZ BCTG study 20]. When risk and quality of life data is supplemented by attitudinal (utility) data from each individual in trials, it is possible to value different outcomes, derive net clinical benefit using quality adjusted survival and other measures, and to evaluate the decision model formulation.

Health state transitions

A significant aspect of this decision problem is that it involves disease progression through a number of states. Methods for valuations of complex outcomes involving transitions between health states are reviewed next.

3 Statistical Methods

QALY and Q-TWiST When outcomes develop over time, a common approach to valuation requires specifying a health state process, where health states (such as TOX, TWiST and REL, referring to periods of toxicity from treatment, time without symptoms or toxicity, and period following relapse) may define distributions of durations in health states. These correspond to the absolute risks of the Warfarin example. Treatment effects will be summarised in hazard ratios for relative benefits or harms of treatments, corresponding to relative risks.

The simple but effective statistical summary of Glasziou et al. (1990) known as quality adjusted life years (QALY) weighs years spent in different health states according to coefficients specifying relative quality of life for each state. For instance, in advanced breast cancer, ratings of outcomes such as pain, appetite, nausea and vomiting may contribute to assessing the quality of life coefficients for TOX and REL. It is conceptually and statistically convenient to calculate QALY from areas under survival curves with defined times from randomization, e.g. duration of toxicity, recurrence free survival, total survival. The method is easily adapted to censored survival data. It also provides a simple discounted analysis (see Eliciting Utilities), requiring only a corresponding transformation of the time axis.

In the Q-TWiST analysis of Gelber et al (1995) an accelerated failure time model for state durations is used to generalise this approach. Accelerated failure time modelling permits inclusion of other covariate information, such as patient attributes, in survival models. An important difference between Q-TWiST and QALY analysis is the need, in Q-TWiST, to specify the dis-

tribution (conditional on the sequence and durations of preceeding states) of the duration of each succeeding health state. In contrast, analysis of clinical trials data normally proceeds through modelling time from randomization to event. When censoring occurs the difference introduces new difficulties.

Eliciting utilities involving valuing future events

Valuations of relative quality expected in different health states are necessary to the decision. After these weightings have been considered, there remains the need to assess risk taking attitudes affecting gambling.

It is necessary for effective decision analysis that weightings elicited accurately reflect the patient's response to outcomes they will experience in a health state and that these relativities remain unchanged over the history that follows treatment. Clinical trial data can be very effective here in letting us explore the relationship between initial and later weightings.

Beyond the relative preferences for a progression of health states that may eventuate, a Q-TWiST or other analysis will need to elicit risk taking attributes of the individual and the discount that they apply to the future as against the present.

There are significant issues (McNeil et al., 1982; Martin et al., 2000) in the phrasing and framing of questions used to elicit patient preferences. With care, preference data may be obtained in trials using appropriate instruments, e.g. Lumley et al. (2001), or even outside trials by examining alternative scenarios, e.g. McNeil et al. (1981). These methods assist in eliciting utilities attached by patients to states. This then provides the quality of life weightings for calculating benefits of alternative treatments.

Monte-Carlo Evaluation of Utilities in Health States

It is not always convenient, or appropriate, to apply QALY or Q-TWiST analyses. An important example not covered by such theory is one where the weights specifying quality of life valuations cannot be assumed constant throughout a health state. It is therefore useful to consider another general approach.

We propose a methodology to evaluate the predictive distribution of the utility $U(Y)$ specified in any model. The expected utility $E(U)$ is obtained as the average valuation of repeated realizations of Y, whose components represent both survival durations and weights, from their joint distribution. In the simple case with fixed weights, the components of Y represent the *history*, a sequence of health states and corresponding survival durations. The joint distribution is adjusted for the individual's baseline risk (or the distribution of uncertainty about this). Individual weights and time varying weights can be included in each realized history and its valuation.

Besides the advantages of providing simulations of outcomes, the general applicability will be helpful where analytic utility calculations are not available. The utility function is itself complex, a non-linear function of survival time and weights, and models involving correlation between survivals and weights are realistic. The Monte-Carlo substitution of an average for expected value is then useful.

Simulating health state histories

Consider a survival model with baseline hazard function, $h(t;\beta_0)$, depending on a parameter β_0, for time to some health state transition (outcome). A relative hazard $\exp(\beta_1)$ adjusts this baseline in the treated group. More generally, when covariates, z, adjust baseline hazard, the Cox model, $h(t;\beta) = h_0(t)\exp(\beta^T z)$, appears suitable for many studies, or Accelerated Failure Time models may be considered. Estimates of relative hazards are commonly available in clinical trials. For simplicity of exposition we specify only two covariates: z_1 an indicator of treatment and z_2 a (trial derived) prognostic index (PI). The PI is used to predict the adjustment to relative hazard appropriate for known prognostic risk factors.

The Monte-Carlo (MC) generation of a health state history we propose begins by fixing the baseline risk parameter β_0, thereby specifying the hazard function $h(t;\beta_0)$ for the population. Then:

i. Sample the hazard ratios for treatment and for a unit change in prognostic index from a bivariate distribution specifying uncertainty in each effect. The adjusted hazard is then $h(t;\beta_0)\exp(\beta^T z)$, where $z = (z_1, z_2)$ has components z_1 indicating new, versus standard, treatment and z_2, the standardised prognostic index.

ii. Sample a hazard ratio adjustment (such as for frailty) and adjustments to each coefficient of this hazard to provide the risk for the new individual. Coefficients are sampled from prespecified distributions. The hazard function for the sampled duration before outcome becomes

$$h(t;\beta_0)\exp(\beta^T z)\exp(\xi_0 + \xi^T z),$$

where the three components of $\exp(\xi)$ allow for random effects introducing *individual* frailty: individual variation from baseline hazard; an individual adjustment to hazard ratio (otherwise uniform) applying with treatment; and, individual adjustment of the hazard associated with unit change in PI.

iii. Simulate survival duration Y consistent with the hazard in [ii] for treatment of interest z_1 and the individual's PI, z_2.

iv. Adopt a model of observable QoL experience, W, (conditional on time to outcome, treatment group, PI and the utility parameters elicited from the individual for this health state). Using this model, generate a realization W of observable QoL and any other risk profile parameters required for valuation of the health state history. For the realization generated, a valuation $U(Y, W)$ of the corresponding individual history Y may now be calculated. The utility calculated can either use the individual's elicited utility preferences or use values adjusted for changes over time in these preferences as occurs in trials data.

Hence to estimate expected utility of survival in each health state, repeating this sequence we obtain a sample of outcomes Y, W (survival times and summaries of QoL experience). Then an overall valuation is provided by

$$\frac{1}{N}\sum_{t=1}^{N} U_t \to E\left\{U(Y;W)\right\}.$$

The Monte-Carlo method first samples the health state realization, providing survival times for given health states, then the quality experience during each health state for that health state history. These stages utilise parameters (and their standard errors) established in clinical trials. The sampling also involves other parameters β_0 and those specifying the distribution of ξ, which cannot be established in the trials. These require pre-specified values or distribution (e.g. a betting odds distribution for this patient).

Discussion

This approach provides a framework for Monte Carlo computation of utility in health state modelling. Adaption may be necessary to capture all features of the application, particularly in the following areas:

1. Multi-state modelling such as a semi-Markov process of health state transitions will be necessary to generate full *histories*. Then the subject's weights for each state's quality of life must be elicited. The expected valuation is the sum of increments valuing separate health state durations, so few additional difficulties are introduced.

2. Uncertainty in utility and survival characteristics of the individual will affect both the utility function itself (e.g. risk parameters for the individual) and the expected utility calculation. For example, while it may be reasonable to set $\xi_1 = \xi_2 = 0$, and $E\xi_0 = 0$, the variability of individual relative hazard, $\exp(\xi_0)$, requires specifying a variance $\sigma^2_{\xi_0}$.

3. While regression adjustment for covariates (e.g. using a PI) provides one method to place the new individual at an appropriate level of risk relative to a standard, another approach might define a window defining comparable individuals from the trial. A k-nearest-neighbour or classification method might then be considered. Corresponding sampled histories from matched individuals in the Clinical Trial database could simulate the future experience of this new individual.

4. Comparisons of utilities are required. The MC method will need to be conducted in a way suitable to determine the boundary of equal utility (threshold analysis) in the parameter domain specifying utility and survival characteristics of the individual.

4 Conclusion

There are many questions raised by the structure of the modelling employed in this paper. It may serve to provoke other views.

First there is the question of whether it is feasible to elicit reliable information from patients capturing their valuation of health states yet to occur. If this is not possible it may still be helpful to those presenting options (e.g. clinicians) and to patients making the decision to present information not only on health state transitions and statistics but also on other patients' decisions, experiences, valuations and the way these change over the course of treatment. Information on how decisions made may vary with other factors such as age of patient are also useful to consider.

Certainly, it appears that information on a key question, 'What is my risk?' may not be presented to patients because of guidelines that emphasise evidence on relative risks. The question as to what reduction in survival would be sufficient to alter one's decision is also important.

For further issues in the decision methodology itself the reader is referred to the Discussion of Cole et al. (1994).

It is then interesting to consider whether trials or other protocols can provide representative decision making parameters for a general population. Trials admit patients under very specific entry criteria and patients attitude to risk may influence their availability to participate. When trials are not feasible, there is the option of decision analysis based on consumer research in the general population. Consumer research has determined very strong preferences in studies of community approval of renal transplants, for example. This approach may be criticised because of the lack of engagement with the medical condition among those participating.

Decision aids developed may provide clinicians with tools to chart the expected benefits of proposed treatments involving a number of uncertain stages of treatment. An effective follow up would provide the patient with material providing scenarios about the future course of the illness. Clinician and patient might use such scenarios to adjudge the uncertainties involved and the competing benefits and risks. This approach may not require explicit individual valuation of the various scenarios presented, but introduce the pertinent information.

Clinicians may also require education in presentation of treatment options so as not to bias valid treatment options, nor to recommend decisions through guidelines merely used as a recipe. Appropriate training would emphasise how to take account of information rather than prescribing the option to be taken. Presentation of tradeoffs in risks and benefits is an important component in such education.

We have introduced a strategy using clinical histories from patients suffering from an identified disease condition to judge the outcome for a a new patient who might suffer from a similar disease condition. The effectiveness of such an approach needs study. While our procedure offers the prospect to adjudge (and maybe quantify) overall quality of life for a patient after a particular disease treatment, the process has not yet been tested to understand its feasibility and practicability. Such testing is an ongoing project being strongly pursued within our group. Research is needed in this area to provide decision models for specific treatment management issues and contexts, where survival outcomes and quality of life valuations are available; determine what information is practical to collect and what is necessary; model observed utility data gathered in clinical trials; and, test such models in new cohorts of patients.

Acknowledgments

The authors thank Val Gebski, Thomas Lumley and a reviewer for stimulating contributions to this work.

REFERENCES

ANZ Breast Cancer Trials Group and the Breast Cancer Institute of Australia (2001). *2001 Annual Scientific Report.* ANZ BCTG.

Carter, C.L., Allen, C. and Henson, D.E. (1989). Relation of tumor size, lymph node status, and survival in 24, 740 breast cancer cases. *Cancer* 63, 181-187.

Cole, B.F., Gelber, R.D. and Anderson, K.M. (1994). Parametric approaches to quality-adjusted survival ananlysis. *Biometrics* 50, 621-631.

Gelber, R.D., Cole, B.F., Gelber, S. and Goldhirsch, A. (1995). Comparing treatments using quality-adjusted survival: The Q-TWiST method. *The American Statistician* 49, 161-169.

Glasziou, P., Simes, R. and Gelber, R. (1990). Quality adjusted survival analysis. *Statistics in Medicine* 9, 1259-1276.

Kirkwood, C.W. (1997). *Strategic Decision Making: Multiobjective Decision Analysis with Spreadshhets*, Chapter 4. Duxbury Press.

Koscielny, S., Le, M.G. and Tubiana, M. (1989). The natural history of human breast cancer, the relationship between involvement of axillary lymph nodes and the initiation of distant metastases. *Br. J. Cancer* 59, 775-782.

Jefford, M. *Breast Cancer Therapy. Information Guide, Vol. 1: Adjuvant Therapy* (2001). Pharmacia Australia.

Langlands, A.O. (1983). Decision making in clinical oncology. *Cancer Topics*, 114–115.

Liberati, A., Patterson, W.B., Biener, L. and McNeil, B.J. (1987). Determinants of physicians' preferences for alternative treatments in women with early breast cancer. *Tumori* 73, 601-609.

Lumley, T., Simes, R., Gebski, V. and Hudson, H. (2001). Combining components of quality of life to increase precision and evaluate tradeoffs. *Statistics in Medicine*, 20, 3231-3249.

Martin, A.J., Simes, R., Lumley, T. and Glasziou, P. (2000). A comparison of standard gamble, time trade-off, and adjusted time trade-off scores. *International Journal of Technology Assessment in Health Care* 16, 137-147.

McNeil, B., Pauker, S., Sox, H. and Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *The New England Journal of Medicine* 306, 1259-1262.

McNeil, B.J., Weichselbaum, R. and Pauker, S.G. (1981). Tradeoffs between Quality and Quantity of Life in Laryngeal Cancer. *The New England Journal of Medicine* 305, 982-987.

N. I. of Health Consensus Development Panel (2001). NIH consensus development conference statement: Adjuvant therapy for breast cancer, Nov. 1-3, 2000. *J. National Cancer Institute* 93, 979-989.

Shapiro, C.L. and Recht, A. (2001). Side effects of adjuvant treatment of breast cancer. *The New England Journal of Medicine* 344, 1997-2008.

ADAPTIVE M-ESTIMATION PROCEDURE: ADEQUACY AND ALTERNATIVE APPROACH

NETHAL K. JAJO
Australian Careers Business College
149-151 George St. Liverpool NSW 2170, Australia
E-mail: n_jajo@yahoo.com.au

KENAN M. MATAWIE
School of Quantitative Methods and Mathematical Sciences
University of Western Sydney
Locked Bag 1779 Penrith South DC NSW 1797, Australia
E-mail: k.matawie@uws.edu.au

XIZHI WU
Renmin University of China,
No 175 Hai Dian Road, HaiDian District Beijing, 100872, China
E-mail: xwu@public3.bta.net.cn

Moberg, et al. (1980) adaptive M-estimation procedure depends heavily on the measures of skewness and kurtosis. The sensitivity of Moberg et al. procedure to the skewness measure was shown using a particular bimodal distribution for errors in linear regression. Properties of this particular distribution have been exploited and investigated; for this type of errors distribution we developed an alternative procedure for choosing the appropriate classification function. Results and performance of the procedure are presented via simulation.

Keywords: bimodal distribution, linear regression, adaptive M-estimator, skewness.

1 Introduction

The multiple linear regression problem assumes a model of the form;

$$Y = X\beta + \epsilon \tag{1}$$

where Y is an $n \times 1$ vector of random variables, X an $n \times k$ matrix of known constants X_{ij}, β is a $k \times 1$ vector of unknown parameters which are to be estimated, and ϵ is $n \times 1$ vector of independently and identically distributed random errors. M-estimation method is one of the robust methods that can provide ample protection from the influence of outlying observations. To apply M-estimation to the regression case, for any estimates b, denote the residual

values by

$$r_i(b) = Y_i - \sum_{j=1}^{k} X_{ij} b_j, \qquad i = 1, 2, ..., n, \tag{2}$$

and the M-estimates of β are defined to be the solutions of the following k equations;

$$\sum_{i=1}^{n} \psi(\frac{r_i(b)}{s(b)}) X_{ij} = 0, \qquad j = 1, 2, \ ..., k, \tag{3}$$

where ψ is one of the appropriate functions (see Andrews et al. 1974), that could be chosen depending on the data's distribution, $s(b)$ is the scale estimate that could be determined simultaneously with b.

Adaptive M-estimation procedures have been suggested by many authors such as Kelly (1992 and 1996), Maronna and Yohai (1981), Crisp and Burridge (1993), Hennig (1995), Yang and Van Ness (1995), Kent and Tyler (1996), Barnett and Lewis (1994), and Mckean, et al. (1993). Most of these papers focused on the determination and limitation of some special form of ψ function, while others discussed the uniqueness and robustness of the estimators. In this paper we will concentrate on Moberg et al. (1980) adaptive M-estimators procedure suggested for selecting the ψ function depending on the measures of skewness and Kurtosis of the errors distribution rather than the true data distribution.

Certainly, Moberg et al. classification was suitable for unimodal where the selection of ψ function depends on the degree of skewness of the errors distribution. Jajo and Hussain (1989) intuitively speculated the use of a reverse scheme of Moberg et al. (1980) classification in the case of bimodal distribution which we believed it needed more study, attention and validation.

To explore the problem more precisely we will use a common bimodal distribution (Slippage model) for errors. In Section 2, the properties of bimodal distribution are given, and the sensitivity of the Moberg et al. (1980) approach to skewness measure is shown. In Section 3, the proposed alternative Moberg et al. classification is presented. Section 4 contains the simulation technique that illustrates the robustness of this rule. Conclusions will be given in Section 5.

2 Bimodal distributions

To illustrate our alternative procedure let us consider a simple linear regression, this procedure can be extended to multiple linear regression in a standard

way. Let

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i, \qquad i = 1, 2, ..., n, \tag{4}$$

be a simple linear regression model with ϵ_i as iid random errors having distribution $N(0,1)$ with probability P and $N(k,1)$ with probability $(1-P)$, where $0.5 < P < 1$ and $k \neq 0$. The pdf of ϵ is given by

$$f_\epsilon(\epsilon) = \frac{1}{\sqrt{2\pi}} \left[P \exp\left(\frac{-\epsilon^2}{2}\right) + (1-P) \exp\left(\frac{-(\epsilon - k)^2}{2}\right) \right]. \tag{5}$$

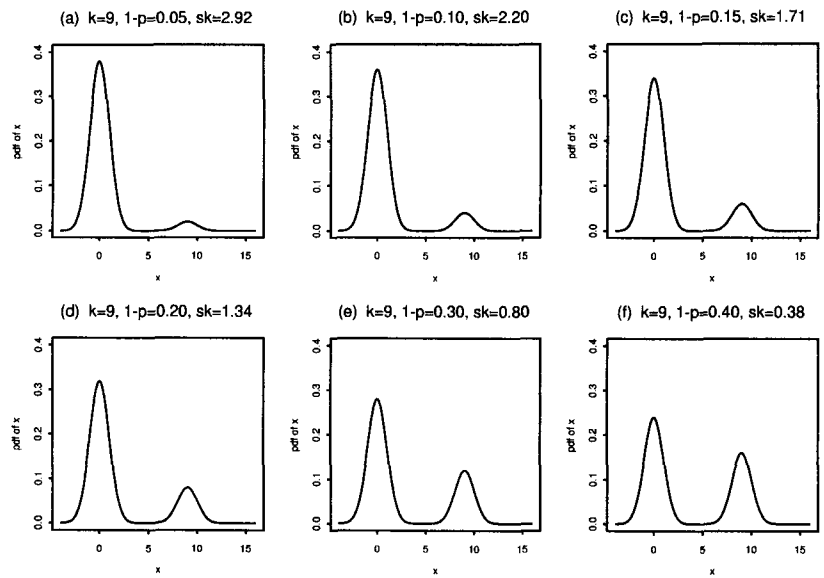


Figure 1. The relationship between skewness and tailweight

Figure 1 provides some pdf of ϵ for various P values of the tailweight. The moment generating function for ϵ can be found as

$$M_\epsilon(t) = P \exp\left(\frac{t^2}{2}\right) + (1-P) \exp\left(kt + \frac{t^2}{2}\right). \tag{6}$$

Proposition 1: For ϵ as defined in (2.5),

1. $E\epsilon = (1-P)k$;

2. $E\epsilon^2 = (1-P)k^2 + 1;$

3. $E\epsilon^3 = k(1-P)(k^2+3).$

This proposition can be proved as $E\epsilon^r = \frac{d}{dt^r}(M_\epsilon(0))^r = M^{(r)}(0)$, where $r = 1, 2$ and 3.

Proposition 2: For ϵ as defined in (2.5), with $\mu_2 = E(\epsilon - E\epsilon)^2$ and $\mu_3 = E(\epsilon - E\epsilon)^3$,

1. $\mu_2 = Pk^2(1-P) + 1;$

2. $\mu_3 = Pk^3(1-P)(2P-1).$

Using $\mu_2 = E(\epsilon - E\epsilon)^2 = E\epsilon^2 - (E\epsilon)^2, \mu_3 = E(\epsilon - E\epsilon)^3 = E\epsilon^3 - 3E\epsilon E\epsilon^2 + 2(E\epsilon)^3$ and Proposition 1, we can easily prove this proposition.

Using Proposition 2 we can conclude that the common skewness measure of the pretense distribution is:

$$sk = \frac{\mu_3}{\mu_2^{3/2}} = \frac{Pk^3(1-P)(2P-1)}{(Pk^2(1-P)+1)^{3/2}}$$

For $P > 0.5$ and $k > 0$ the distribution of ϵ in (2.4) is skewed to the right.

Moberg *et al*, (1978 and 1980) suggested adaptive procedures, see section 3, based on the estimates or the empirical measures of the third and forth moments of the data, Q_3 and Q_4 defined below. These measures were used to select one of the five classification functions(ψ)for the M-estimation procedure. Q_3 and Q_4 are also a traditional estimators for skewness and tailweight of the data.

$$Q_3 = \frac{\bar{U}(0.05) - \bar{M}(.5)}{\bar{M}(.5) - \bar{L}(.05)}; \quad Q_4 = \frac{\bar{U}(.05) - \bar{L}(.05)}{\bar{U}(.5) - \bar{L}(.5)},$$

where $\bar{L}(\gamma)$, $\bar{M}(\gamma)$ and $\bar{U}(\gamma)$ are the arithmetic means of the smallest, middle and largest $N\gamma$ of the ordered residuals $z_1, z_2, \ldots, z_N$ obtained from (1.2) with $k-1$ zero residuals deleted and $N = n - k + 1$ is the number of these ordered residuals.

In bimodal distribution, the case we are concerned with, the commonly used skewness measure sk is a misleading and improper measure as it decreases in value when the area of the tail $(1-P)$ increases. For example, for $k = 9$ and $\sigma = 1$, when $1 - P$ increases from 0.05 to 0.40, sk decreases from 2.92 to 0.38 (see Figure 1). Clearly sk, and therefore its empirical measure (Q_3), used by Moberg et al. (1980) approach for choosing ψ function is critical and not suitable for the bimodal case.

In the above situation the skewness measure need to be based on the number of outliers and their distance from the main body of the data, i.e. the values of P and k in ϵ's distribution. So, we suggest that Q_3 should be replaced by an alternative measure of skewness as in the following form

$$ALQ_3 = Cn^{\alpha}k(1-P),$$

where $C = 1$ and $0 < \alpha \leq 0.5$ ($\alpha = 0.45$ is usually recommended).

The usefulness and efficient performance of the alternative measure of skewness ALQ_3 compared to Q_3 was clearly noticeable in our example in section 4.

3 Alternative regression procedure

The procedure we use depends on examining the residuals by calculating their Q_4 and ALQ_3 and identifying the position of (Q_4, ALQ_3) point on a partitioned positive quarter of Q_4-ALQ_3 plane to determine which ψ function should be selected. Figure 2 is a partitioned quarter for classification when $n = 10$ and 20. It is quite similar for a general n.

Our classification scheme can be described by the five lines that divides the positive quarter of Q_4-ALQ_3 plane into five areas. These lines, shown below, are similar to those given by Moberg *et al.* (1978 and 1980) with some modifications.

L1: Q_4= 5.46 - (21.6/N) - (1.7 - 9/N)ALQ_3
L2: Q_4= 2.09 - 2/N
L3: Q_4= 3.09 - 9/N
L4: Q_4= (0.75 - 1/N)ALQ_3 + (1.65 - 2/N)
L5: Q_4= (0.5 - 1/N)ALQ_3 + (1.25 - 1/N).

The selection of the ψ function is determined by the area where the point (Q_4, ALQ_3) is located among the five areas (shown in Figure 2). Specifically if the point (Q_4, ALQ_3) is in:

Area I, choose class 1: $\psi(u) = u^3$;
Area II, choose class 2: $\psi(u) = u$;
Area III, choose class 3: $\psi(u) = 2.25u/(u^2 + 0.5s^2)$;
Area IV, choose class 4: $\psi(u) = 5.1u/[(u + 0.74s)^2 + 0.87s^2]$;
Area V, choose class 5: $\psi(u) = 3.3u/[(u - 0.54s)^2 + 0.23s^2]$.

Classes 1 and 2 corresponds to symmetric distributions while classes 3, 4 and 5 are for distributions with light, moderate and heavy skewness respectively. The proposed scale estimate is $s = [F_N^{-1}(.75) - F_N^{-1}(.25)]/1.35$, where F_N is the empirical cdf based on the ordered residuals. s can also be computed

using the interquartile range (IQR) as mentioned in Swallow and Kianifard (1996).

The effectiveness of our adaptive rule is that we can select a suitable $\psi(\cdot)$ for either case of bimodal or unimodal distribution without having to know the true distribution of the data. We also found that the larger the sample size, the better the performance of our suggested adaptive procedure. Simulation example in the following section demonstrates the excellent performance of our technique.

4 Simulation example

Based on the model $Y_i = 2+x_i+\epsilon_i$, $i = 1, 2, ..., n$, 100 samples of $n = 10$ and 20 observations were generated by simulation. The values of x's are fixed for each sample which is generated from uniform $U(1, n + 10)$ while ϵ's are generated from $N(0, 1)$ distribution with probability P and $N(k, 1)$ distribution , $k \neq 0$, with probability $(1 - P)$.

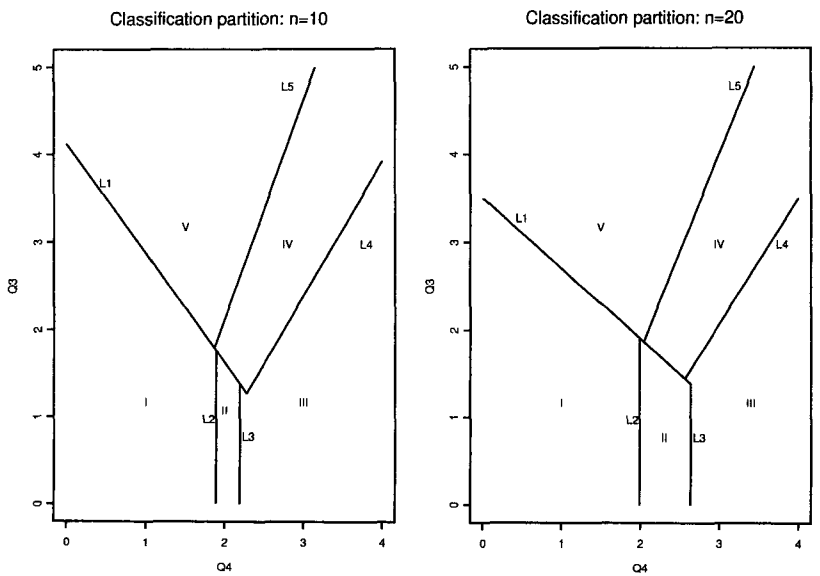


Figure 2. Classification scheme

Table 1 gives the relative sum of squares for errors for each of the 100 simulations when $n = 10$ and 20 with k=9 and 15 and various values of P.

Table 1. The efficiency and statistical measures for different criteria

n	k	P	Efficiency						Measures		
			C_3		C_4		C_5		Q_4	Q_3	$ALQ3$
			β_0	β_1	β_0	β_1	β_0	β_1			
10	9	0.9	1.95	2.21	2.41	2.48	2.82	2.72	3.83	5.15	2.53
		0.8	1.24	1.45	1.66	1.92	2.02	2.18	2.37	5.15	5.07
		0.7	1.05	1.10	1.40	1.43	1.74	1.67	1.67	8.13	7.60
	15	0.9	2.06	2.67	2.46	3.15	2.82	3.58	4.18	8.63	4.22
		0.8	1.37	1.75	1.56	2.00	1.76	2.26	2.41	8.56	8.45
		0.7	1.09	1.20	1.28	1.38	1.47	1.55	1.72	13.6	12.7
20	9	0.95	9.29	8.96	6.29	11.9	3.55	9.30	6.01	4.91	1.73
		0.90	27.8	19.3	19.5	21.7	10.5	17.5	4.06	4.79	3.46
		0.85	7.18	7.33	7.61	8.20	7.11	8.69	3.04	4.61	5.19
	15	0.95	3.68	3.52	2.51	4.66	1.41	3.65	6.99	7.59	2.88
		0.90	9.82	7.01	7.44	8.15	4.04	6.67	4.35	7.93	5.77
		0.85	5.80	4.78	6.05	5.64	4.79	5.80	2.77	7.66	8.66

n	k	P	Selecting criteria			Class of high effi.	
			DR ALQ_3	DR Q_3	MC		
						β_0	β_1
10	9	0.9	3	4	3	5	5
		0.8	5	5	3	5	5
		0.7	5	5	1	5	5
	15	0.9	4	5	3	5	5
		0.8	5	5	3	5	5
		0.7	5	5	1	5	5
20	9	0.95	3	3	3	3	4
		0.90	4	4	3	3	4
		0.85	5	4	3	4	5
	15	0.95	3	3	3	3	4
		0.90	4	5	3	3	4
		0.85	5	5	3	4	5

The efficiency of β_0 and β_1 are defined respectively by

$$\frac{\sum_{i=1}^{100}(\hat{\beta_{0i}}-2)^2}{\sum_{i=1}^{100}(\tilde{\beta_{0i}}-2)^2} \quad \text{and} \quad \frac{\sum_{i=1}^{100}(\hat{\beta_{1i}}-1)^2}{\sum_{i=1}^{100}(\tilde{\beta_{1i}}-1)^2},$$

where $\hat{\beta}_0$ and $\hat{\beta}_1$ are the ordinary least squares estimates of β_0 and β_1 respectively, while $\tilde{\beta}_0$ and $\tilde{\beta}_1$ are the estimates of the same parameters obtained by using Moberg et al. (1980) adaptive M-estimation method. Since the model

we used is not symmetric and it is skewed to the right, so the procedure should concentrate on one of the classes 3, 4 or 5. The results are given in Table 1 under the efficiency columns C_3, C_4 and C_5. To obtain the adaptive M-estimations, we need to solve the equations in (1.3) iteratively. Gaussian-Newton iterative method (Agee and Turner, 1979) was applied using Theil estimates for $n = 10$ and Modified Kilda for $n = 20$ as suggested starting values for β_0 and β_1 (Hussain and Sprent 1983).

The chosen class of ψ function is usually determined by the highest efficiency of the parameter estimates. For example, the highest values of the estimated efficiency in the presence of one outlier when $n = 10$, $k = 9$ and $P = 0.9$ are 2.82 and 2.72 for estimating β_0 and β_1 respectively. These two values are shown in the efficiency column C_5 of Table 1, hence class 5 is classified as being of high efficiency, and recorded in the last column of Table 1 as a selected class of high efficiency for that particular case.

Table 1 summarizes the information of the simulation study for different n, k, and P. It also gives the statistical measures Q_4, Q_3 and ALQ_3 and their determined class, which are based on the location of the points (Q_4, ALQ_3) and (Q_4, Q_3) on their respective planes as shown in Figure 2.

The following key words were used in Table 1 to refer to different selecting criteria:

$\mathrm{DR}(Q_4, ALQ_3)$ = Dropping the points (Q_4, ALQ_3) on Q_4-ALQ_3 plane.

$\mathrm{DR}(Q_4, Q_3)$ = Dropping the points (Q_4, Q_3) on Q_4-Q_3 plane.

MC = Moberg et al. (1980) criteria.

To compare these criteria we present the following bar plots (Figure 3) showing the probability of success for each procedure with different number of outliers.

$$\text{The probability of success} = \begin{cases} 1, & \text{if the high efficiency class is the chosen one;} \\ 0.33, & \text{if the chosen class is either 3, 4, or 5;} \\ 0, & \text{otherwise.} \end{cases}$$

From Table 1 and Figure 3 we notice that for both n=10 and n=20 with various values of k and P, the suggested measure ALQ_3 has always the highest probability of success while using MC has the lowest, especially as P decreases. When P =0.7 the probability of success is zero and there are no observations in the bar plot. This example demonstrates the adequacy and superiority of ALQ_3 over the MC method of using Q_3, it also shows the accuracy obtained by dropping the points on $Q_4 - Q_3$ plane rather than using Moberg et al. (1980) inequalities.

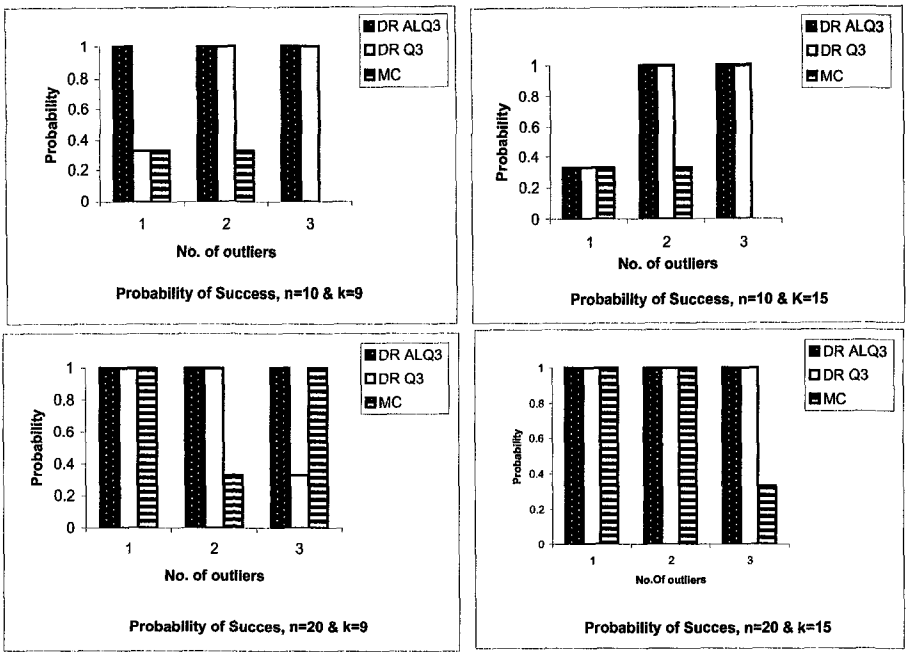


Figure 3. Plots of the probability of success. Note: 1, 2 and 3 (on the x-axis) indicates the number of the outliers determined by using the corresponding P values.

5 Conclusion

The principle aim of this work was to devise an adaptive regression estimation procedure for bimodal distributions. The adaptive procedure described is competitive with other robust procedures given in the literature and it is superior when the underlying distribution is bimodal. As mentioned in Section 2 the results in the previous section can be extended to a more general form of linear models in which same results and conclusions can be obtained.

REFERENCES

Agee, W. S. and Turner, R. H. (1979). *Application of Robust Regression to Trajectory Data Reduction.* R. L. Launer and G. N. Wilkinson (eds.). London: Academic Press.

Andrews, D. F. (1974). A robust method for multiple linear regression. *Technometrics* 16, 523-531.

Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data.* 3rd Edition. New York: John Wiley.

Crisp, A. and Burridge, J. (1993). A note on the uniqueness of M-estimators in robust regression. *The Canadian Journal of Statistics* 21, 205-208.

Hennig, C. (1995). Efficient high-breakdown-point estimators in robust regression: which function to choose?. *Statistics & Decisions* 13, 221-241.

Hussain, S. S. and Sprent, P. (1983). Non-Parametric regression. *J. A. Statist. Soc. A. Part 2* 146, 182-191.

Jajo, N. K. and Hussain, S. S. (1989). Robust estimation in linear regression model. *Journal of Management and Economic* 10, 1-15.

Kelly, G. E. (1992). Robust regression estimators–the choice of tuning constants. *Statistician* 41, 303-314.

Kelly, G. E. (1996). Adaptive choice of tuning constant for robust regression estimators. *Statistician* 45, 35-40.

Kent, J. and Tyler, D. (1996). Constrained M-estimation for multivariate location and scatter. *The Annals of Statistics* 24, 1346-1370.

Maronna, A. and Yohai, V. (1981). Asymptotic behavior of general M-estimates for regression and scale with random carriers. *Z. Wahrsch. Verw. Geb.* 58, 7-20.

Mckean, J. W., Sheather, S. J. and Hettmansperger, T. P. (1993). The use and interpretation of residuals based on robust estimation. *Journal of the American Statistical Association* 88, 1254-1263.

Moberg, T. F., Ramberg, J. S. and Randels, R. H. (1978). An adaptive M-estimator and its application to a selection problem. *Technometrics* 20, 255-263.

Moberg, T. F., Ramberg, J. S. and Randels, R. H. (1980). An adaptive multiple regression procedures based on M-estimators. *Technometrics* 22, 213-223.

Swallow, W. H. and Kianifard, F. (1996). Using robust scale estimates in detecting multiple outliers in linear regression. *Bimetrics* 52, 545-556.

Yang, J. J. and Van Ness, J. W. (1995). Breakdown points for redescending M-estimates for location. *Commun. Statist. Theory and Methods* 7, 1769-1787.

BAYESIAN INFERENCE FOR CATEGORICAL DATA WITH MISCLASSIFICATION ERRORS

MASAHIRO KURODA
Department of Computer Science and Mathematics
Kurashiki University of Science and the Arts
Nishinoura 2640 Turajima-cho Kurasihi-shi
Okayama 712-8505, Japan
E-mail: kuroda@soft.kusa.ac.jp

ZHI GENG
Department of Statistics
Peking University
Beijing 100871, China,
E-mail: zgeng@math.pku.edu.cn

In epidemiological studies, observed data are often collected subject to misclassification errors. In this paper, we discuss the Bayesian estimation for contingency tables with misclassification errors. Employing the exact Bayesian computations to obtain posterior means as estimates, we are faced with computational difficulties. In order to find the posterior distribution, we apply the data augmentation(DA) algorithm to misclassified categorical data.

Keywords: Bayesian estimation, contingency table, data augmentation algorithm, misclassification, posterior means.

1 Introduction

In epidemiological studies, observed data are often collected subject to misclassification errors. Such misclassification errors cause bias of estimation and reduce efficiency in the analysis of contingency tables. Many investigators have discussed how to adjust for the effects of misclassification, see, for example, the review papers Chen (1989) and Walter and Irwig (1987). Espeland and Odoroff (1985) discussed the maximum likelihood estimation for a recursive system of log-linear models based on double sampling schemes by the EM algorithm. Espeland and Hui (1987) applied the Fisher scoring algorithm to evaluate variances and covariances of estimates.

From the Bayesian viewpoint, Geng and Asano (1989) considered estimation methods for misclassified categorical data making use of prior information and double sampling schemes, and obtained posterior means as estimates.

Viana (1994) applied Bayesian computations based on the matrix of misclassification probabilities to small-sample multinomial data. Evans et al (1996) discussed the implementation of the Gauss-Jacobi quadrature and the Gibbs sampling algorithm for the posterior analysis of binary response data with misclassification.

In this paper, we present a Bayesian approach that utilizes prior knowledge about misclassification and incorporates this prior knowledge with observations subject to misclassification. Although the EM algorithm or the Fisher scoring algorithm are often applied to estimating model parameters, these algorithms can not evaluate posterior distributions on the model parameters. Furthermore, these algorithms do not apply to our estimation problem because of the unidentifiability of the model parameters. However, our Bayesian approach, assuming a prior distribution on the model parameters, can overcome these problems. In order to find the posterior distribution of model parameters and calculate posterior means as estimates of them, we use the data augmentation(DA) algorithm of Tanner and Wong (1987).

In Section 2, we discuss the Bayesian computation to find a posterior distribution given misclassified observed data. In Section 3, we give the DA algorithm to approximate the posterior distribution, because of difficulties with the calculation of the posterior distribution. Section 4 presents two numerical experiments to examine the performance of the DA algorithm.

2 Misclassified observed data and Bayesian computation

Let X and Y be categorical variables having I and J categories, respectively, and let Y' be a misclassified variable of Y having K categories. We assume that two types of misclassified data are observed: (i) data for X and Y', denoted as $n = \{n_{i+k} \mid i \in \{1,\ldots,I\}, k \in \{1,\ldots,K\}\}$, and (ii) data for Y and Y', denoted as $m = \{m_{+jk} \mid j \in \{1,\ldots,J\}, k \in \{1,\ldots,K\}\}$, where the symbol "+" means the sum over corresponding variables, for example, $n_{i+k} = \sum_j n_{ijk}$. Let p_{ijk} denote a probability for $(X,Y,Y') = (i,j,k)$ and $\theta_{XYY'} = \{p_{ijk} \mid i \in \{1,\ldots,I\}, j \in \{1,\ldots,J\}, k \in \{1,\ldots,K\}\}$ denote a set of probabilities.

In this paper, the goal is to find the posterior distribution of model parameters $\theta_{XY} = \{p_{ij+} \mid i \in \{1,\ldots,I\}, j \in \{1,\ldots,J\}\}$ which are the marginal probabilities of X and Y, and obtain the posterior means of θ_{XY} as estimates.

Assume that n and m have independently multinomial distributions with parameters $\theta_{XY'} = \{p_{i+k} \mid i \in \{1,\ldots,I\}, k \in \{1,\ldots,K\}\}$ and $\theta_{YY'} = \{p_{+jk} \mid j \in \{1,\ldots,J\}, k \in \{1,\ldots,K\}\}$, respectively, that is,

$$f(n \mid \theta_{XY'}) = \frac{n_{+++}!}{\prod_{i,k} n_{i+k}!} \prod_{i,k} p_{i+k}^{n_{i+k}}, \quad f(m \mid \theta_{YY'}) = \frac{m_{+++}!}{\prod_{j,k} m_{+jk}!} \prod_{i,k} p_{+jk}^{m_{+jk}}, \quad (1)$$

and that the prior distribution of $\theta_{XYY'}$ is a Dirichlet distribution which has the density function

$$\pi(\theta_{XYY'} \mid \alpha_{XYY'}) = \frac{\Gamma[\alpha_{+++}]}{\prod_{i,j,k} \Gamma[\alpha_{ijk}]} \prod_{i,j,k} p_{ijk}^{\alpha_{ijk}-1}, \quad (2)$$

where $\alpha_{XYY'} = \{\alpha_{ijk} \mid i \in \{1, \ldots, I\}, j \in \{1, \ldots, J\}, k \in \{1, \ldots, K\}\}$ is a set of hyper-parameters of the prior distribution of $\theta_{XYY'}$.

From the equations (1) and (2), we obtain the mixture posterior distribution given n and m. The posterior density is given by

$$\begin{aligned}
&\pi(\theta_{XYY'} \mid n, m) \\
&\propto f(n \mid \theta_{XY'}) f(m \mid \theta_{YY'}) \pi(\theta_{XYY'} \mid \alpha_{XYY'}) = \prod_{i,k} p_{i+k}^{n_{i+k}} \prod_{j,k} p_{+jk}^{m_{+jk}} \prod_{i,j,k} p_{ijk}^{\alpha_{ijk}-1} \\
&= \prod_k \left\{ \prod_i \sum_{\Omega(n)} \frac{n_{i+k}!}{\prod_j \tilde{n}_{ijk}!} \prod_j \sum_{\Omega(m)} \frac{m_{+jk}!}{\prod_i \tilde{m}_{ijk}!} p_{ijk}^{\alpha_{ijk}+\tilde{n}_{ijk}+\tilde{m}_{ijk}-1} \right\}, \quad (3)
\end{aligned}$$

where $\sum_{\Omega(n)}$ denotes the sum over all possible $\{\tilde{n}_{ijk}\}$ under the conditions $\tilde{n}_{ijk} \geq 0$ for all i, j and k, and $\sum_j \tilde{n}_{ijk} = n_{i+k}$, and $\sum_{\Omega(m)}$ denotes the sum over all possible $\{\tilde{m}_{ijk}\}$ under the conditions $\tilde{m}_{ijk} \geq 0$ for all i, j and k, and $\sum_i \tilde{m}_{ijk} = m_{+jk}$.

However, the posterior density (3) is a very complicated function. It is extremely difficult to calculate exactly the posterior distribution and these calculation may take a long time when the observed data set is moderately large.

For cases where the incomplete-data posterior density is as complicated as equation (3) and the complete-data posterior density is relative easy to handle and draw from, the DA algorithm is very suitable.

In the next section, we present the DA algorithm to approximate the posterior distribution and estimate posterior means of model parameters θ_{XY}.

3 DA algorithm for misclassified data

The DA algorithm consists of iterating between the *imputation*-step and the *posterior*-step.

For this misclassified multinomial model, the DA algorithm is given by the following iterative scheme:

Imputation-step: Repeat the following two steps for $l = 1, \ldots, L$ to obtain the imputed data of n and m such as $\tilde{n} = \{\tilde{n}_{ijk} \mid i \in \{1, \ldots, I\}, j \in \{1, \ldots, J\}, k \in \{1, \ldots, K\}, \sum_j \tilde{n}_{ijk} = n_{i+k}\}$ and $\tilde{m} = \{\tilde{m}_{ijk} \mid i \in \{1, \ldots, I\}, j \in \{1, \ldots, J\}, k \in \{1, \ldots, K\}, \sum_i \tilde{m}_{ijk} = m_{+jk}\}$.

1. Generate cell probabilities $\{p^*_{ijk}\}$ from the current estimate of the posterior distribution,
2. Generate the imputed data $\tilde{n}^{(l)}$ and $\tilde{m}^{(l)}$ from the predictive distributions which have the conditional multinomial distributions, given n and m, with densities

$$f(\tilde{n} \mid \{p^*_{j|i,k}\}, n) = \prod_{i,k} \frac{n_{i+k}!}{\prod_j \tilde{n}_{ijk}!} \prod_{i,j,k} p^{*\tilde{n}_{ijk}}_{j|i,k},$$

$$f(\tilde{m} \mid \{p^*_{i|j,k}\}, m) = \prod_{j,k} \frac{m_{i+k}!}{\prod_i \tilde{m}_{ijk}!} \prod_{i,j,k} p^{*\tilde{m}_{ijk}}_{i|j,k},$$

where $p^*_{j|i,k} = p^*_{ijk}/p^*_{i+k}$ and $p^*_{i|j,k} = p^*_{ijk}/p^*_{+jk}$.

Posterior-step: Update the current approximation of the posterior distribution of $\theta_{XYY'}$, given these imputed data $\tilde{n}^{(l)}$ and $\tilde{m}^{(l)}$, for $l = 1, \ldots, L$, by the Monte Carlo method,

$$\pi(\theta_{XYY'} \mid n, m) = \frac{1}{L} \sum_{l=1}^{L} \pi(\theta_{XYY'} \mid \tilde{n}^{(l)}, \tilde{m}^{(l)}).$$

Until the approximated distribution converges to a stationary distribution, the imputation-step and the posterior-step are iterated.

In obtaining a stationary distribution, the values of L may be increased to improve the accuracy with respect to the Monte Carlo method.

We can then find the posterior distribution of θ_{XY} and easily obtain their posterior means and variances. Furthermore, it is possible to calculate the highest posterior density(HPD) region that is the Bayesian analogue of the confidence intervals.

4 Numerical experiments

We provide two numerical experiments to examine the performance of the DA algorithm described in the previous section.

Comparison of estimates and exact values

In the following numerical experiment, we compare the estimates obtained by the DA algorithm with the posterior means using the exact Bayesian calculation given by Geng and Asano (1987) regarding the data from Diamond and Lilienfeld (1962). Their data reported a case-control study concerning the circumcision status of male partners of woman with and without cervical cancer.

The study sample was categorized by cervical cancer status, X (Case and Control), and self-reported circumcision status, Y' (Yes or No), in the left side of Table 1.

In order to gain the information on the degree of misclassification of circumcision status, the supplemental sample concerning the relationship between actual circumcision status, Y (Yes or No), and Y' was gathered from the separate population shown by in the center of Table 1.

Espeland and Hui (1987) described that, for the misclassified multinomial model, the conditional independence model between X and Y given Y' from the class of hierarchical log-linear models was appropriate, since no observed data for X, Y and Y' were obtained.

Furthermore, for the conditional independence model, Geng and Asano (1989) gave the hypothetical prior information shown in the right side of Table 1.

Table 1. Data from Diamond and Lilienfeld (1962) and hypothetical prior information $\alpha_{XYY'}$

Y'	Y	Study Sample		Supple. Sample	Hypothetical prior	
		X		X	X	
		Case	Control	Unknown	Case	Control
Yes	Yes	5	14	37	80	10
	No			19	20	40
No	Yes	95	86	47	40	20
	No			89	10	80

We use their prior information as hyper-parameters $\alpha_{XYY'} = \{\alpha_{ijk} \mid i \in \{\text{Case, Control}\}, j \in \{\text{Yes, No}\}, k \in \{\text{Yes, No}\}\}$ and then obtain a posterior distribution and estimates of $\theta_{XY} = \{p_{ij+} \mid i \in \{\text{Case, Control}\}, j \in \{\text{Yes, No}\}\}$ for X and Y by posterior means.

In this numerical experiment, we evaluate the accuracy of the estimates using the DA algorithm in comparison with the exact posterior means given

by Geng and Asano (1989). Table 2 shows the exact values, and the posterior means, the standard deviations(SDs) and the posterior 95 % credible intervals(CIs) of θ_{XY} obtained from simulated 10,000 samples after a *burn-in* of 1,000 samples. It can be seen that the estimates have approximately three-digit accuracy for the exact values. From the numerical results, we can see that the DA algorithm works quite well in estimating the posterior means.

Table 2. Posterior means and SDs and 95% CIs using the DA algorithm and the exact posterior means given by Geng and Asano(1989)

X	Y	Exact Bayes	DA	
		Posterior means	Posterior means ± SD	CI (lower-upper)
Case	Yes	0.3794	0.3786 ± 0.0127	0.3512 - 0.4017
	No	0.1116	0.1134 ± 0.0159	0.0838 - 0.1460
Control	Yes	0.0921	0.0927 ± 0.0107	0.0737 - 0.1142
	No	0.4169	0.4152 ± 0.0113	0.3916 - 0.4364

Performance of the DA algorithm

In this experiment, we examined the performance of the DA algorithm in comparison with the EM algorithm and the Fisher scoring algorithm. We apply the DA algorithm to the *double sampling* data from Hochberg (1977). The data were the highway safety research data relating the seat-belt usages to driver injuries. The main sample was of 80,084 accidents that were recorded by police subject to misclassification errors. The subsample was of 1,796 accidents that were recorded by both *imprecise* police reports and *precise* hospital interviews. Then, by the double sampling design, the subsample was randomly selected from the main sample. Thus, the subsample and the main sample have independent and identical distributions.

The main sample and the subsample in Table 3 were categorized by four variables X, X', Y and Y', where X and Y denote precise personal survey for seat-belt usages and driver injuries, and X' and Y' denote imprecise police reports for them.

In this experiment, we estimated model parameters under the saturated multinomial model, because our purpose is to investigate whether the DA algorithm is applicable to estimate model parameters, but not to analyze the misclassified observed data.

For these data we assume that the main sample data and the subsample

data have independent and identical multinomial distributions with

$$\theta_{XX'YY'} = \{p_{ijkl} \mid i \in \{\text{Yes, No}\}, j \in \{\text{Yes, No}\}, \\ k \in \{\text{Yes, No}\}, l \in \{\text{Yes, No}\}\},$$

where $p_{ijkl} = \Pr(X = i, X' = j, Y = k, Y' = l)$ and the prior distribution for $\theta_{XX'YY'}$ has the Dirichlet distribution with hyper-parameters $\alpha_{XX'YY'} = \{\alpha_{ijkl} \mid i \in \{\text{Yes, No}\}, j \in \{\text{Yes, No}\}, k \in \{\text{Yes, No}\}, l \in \{\text{Yes, No}\}\}$.

Then the model parameters are marginal probabilities of X and Y,

$$\theta_{XY} = \{p_{i+k+} \mid i \in \{\text{Yes, No}\}, k \in \{\text{Yes, No}\}\},$$

where $p_{i+k+} = \sum_{j,l} p_{ijk}$. We utilize the subsample in Table 3 as hyper-parameters $\alpha_{XX'YY'}$ and obtain estimates of θ_{XY} by the DA algorithm. Table 4 shows the estimates and the SDs of θ_{XY} obtained by the DA algorithm, the exact Bayesian calculation, the Fisher scoring algorithm and the EM algorithm.

Table 3. Data of highway safety research (Hochberg, 1977)

		Main Sample		Subsample			
		X' = Yes	X' = No	X' = Yes		X' = No	
Y'	Y			X = Yes	X = No	X = Yes	X = No
Yes	Yes	1196	13562	17	3	10	258
	No			3	4	4	25
No	Yes	7151	58175	16	3	25	194
	No			100	13	107	1014

Table 4. Estimates and their SDs of θ_{XY}

X	Y	Exact Bayes	DA	Fisher scoring	EM
		Posterior means ± SD	Posterior means ± SD	Estimates ± SD	
Yes	Yes	0.0397 ± 0.0043	0.0389 ± 0.0041	0.0394 ± 0.0045	0.0394
	No	0.1293 ± 0.0065	0.1311 ± 0.0073	0.1190 ± 0.0076	0.1294
No	Yes	0.2558 ± 0.0079	0.2577 ± 0.0078	0.2563 ± 0.0103	0.2559
	No	0.5752 ± 0.0093	0.5722 ± 0.0092	0.5870 ± 0.0116	0.5752

The estimates using the DA algorithm can be found from simulated samples 100,000 after a burn-in samples 10,000 in two chains. The exact values of estimates of θ_{XY} using the Bayesian calculation are given by Geng and Asano (1989) who assumed the Jeffreys noninformative prior. The estimation

using the Fisher scoring algorithm were carried out with ℓ_{EM} developed by Vermunt (1997).

From these numerical results, it can be seen that the DA algorithm has the equivalent performance of the EM and the Fisher scoring algorithm in comparison with these estimates and SDs.

The EM and the Fisher scoring algorithms have disadvantages such that it is impossible to find the posterior distribution of model parameters, these algorithms may not be applied owing to unidentifiability of the model, and it may be difficult to calculate the Fisher information matrix needed in the Fisher scoring algorithm.

5 Concluding remarks

In this paper, we discussed the DA algorithm to estimate model parameters for misclassified categorical data. We gave the posterior distribution by exact Bayesian computation. To avoid complicated calculation, we used the DA algorithm and find the posterior distribution. It is easily seen that the DA algorithm is the iterative simulation version of the EM algorithm in which the imputation-step corresponds to the E-step and the posterior-step corresponds to the M-step.

In order to explore the possibility of parameter estimation by the DA algorithm, we provided two numerical experiments. In the first experiment, we evaluated accuracy of estimates in comparison with exact values. In the second experiment, we examined the performance of the DA algorithm. The results of both numerical experiments showed the advantage of applying the DA algorithm in terms of accuracy of estimates and in terms of algorithm simplicity to find the posterior distribution.

For the inference of multidimensional contingency tables, the Bayesian inference by the DA algorithm can be easily extended and also widely utilized. Then we may need to take account of conditional independence between variables in the models. For parameters assuming the conditional independence model, a prior Dirichlet distribution that has hyper Markov laws by Dawid and Lauritzen (1993) is very suitable. A future problem is how to incorporate prior information with the hyper-parameters of a hyper Dirichlet prior distribution without consistency.

Acknowledgements

The authors would like to thank an editor and a referee for helpful suggestions.

REFERENCES

Chen, T.T. (1989). A review of methods for misclassified categorical data in epidemiology. *Statistics in Medicine* 8, 1095-1106.

Dawid, A.P. and Lauritzen, S.L. (1993). Hyper Markov laws in the statistical analysis of decomposable graphical models. *Ann. Statist.* 21, 1272-1317.

Diamond, E.L. and Lilienfeld, A.M. (1962). Effects of errors in classification and diagnosis in various types of epidemiological studies. *American Journal of Public Health* 52, 1137-1144.

Espeland, M.A. and Odoroff, C.L. (1985). Log-linear models for doubly sampled categorical data fitted by the EM algorithm. *Journal of the American Statistical Association* 80, 663-670.

Espeland, M.A. and Hui, S.L. (1987). A general approach to analyzing epidemiologic data that contain misclassification errors. *Biometrics* 43, 1001-1012.

Evans, M., Guttman, I., Hatiovsky, Y. and Swartz, T. (1996). Bayesian analysis of binary data subject to misclassification. In *Bayesian Analysis in Statistics and Economics*, D. Berry, J. Geweke and K.M. Chaloner (editors). New York: John Wiley and Sons, 67-77.

Geng, Z. and Asano, Ch. (1989). Bayesian estimation methods for categorical data with misclassification. *Communications in Statistics* 8, 2935-2954.

Hochberg, Y. (1977). On the use of double sampling schemes in analyzing categorical data with misclassification errors. *Journal of the American Statistical Association* 72, 914-921.

Tanner, M.A. and Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* 82, 528-540.

Vermunt, J.K. (1997). ℓ_{EM}: *A general program for the analysis of categorical data.* Tilburg University.

Viana, M.A.G. (1994). Bayesian small-sample estimation of misclassified multinomial data. *Biometrics* 50, 237-243.

Walter, S.D. and Irwig, L.M. (1987). Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. *Journal of Clinical Epidemiology* 41, 923-937.

THE GLOBAL SOLUTION OF AN INITIAL VALUE PROBLEM FOR THE DAMPED BOUSSINESQ EQUATION

S. LAI
Department of Applied Mathematics
Southwest Jiaotong University
Chengdu 610031, P.R. China
E-mail: Laishay@main.sc.cninfo.net

Y. H. WU
Department of Mathematics and Statistics
Curtin University of Technology
GPO Box U 1987, Perth WA 6845, Australia
E-mail: yhwu@maths.curtin.edu.au

The aim of this paper is to establish the well-posedness of the global solution of an initial value problem for the damped Boussinesq equation

$$u_{tt} - 2bu_{txx} = -\alpha u_{xxxx} + u_{xx} + \beta(f(u))_{xx},$$

where $x \in (-\infty, \infty), t > 0$, α and b are positive constants, β is a real constant and $f \in C^{\infty}$. For the case $\alpha > b^2$, our result reduces to the well-posedness theorem obtained by Varlamov (1996).

Keywords: Boussinesq equation, initial value problem, global solution, existence.

1 Introduction

In the 1870's, Boussinesq derived various equations for modelling propagation of long waves with small amplitude on the surface of water. These equations were found to possess solutions of traveling waves known as solitary waves. Boussinesq also gave a scientific explanation on the existence of the waves, and proposed to use the so-called Lyapunov function to describe the stability of the waves.

The classical Boussinesq equation can be written as

$$u_{tt} = -\alpha u_{xxxx} + u_{xx} + \beta(u^2)_{xx}, \tag{1}$$

where $u(x,t)$ is the elevation of the free surface of fluid, the subscripts denote partial derivatives, and the constant coefficients α and β depend on the depth of fluid and the characteristic speed of the long waves. Some interesting comments on the history of the derivatives of (1) can be found in the review

articles Miles (1980) and Milewski and Keller (1996). Various versions of the Boussinesq equation discussed in the literature can be obtained in the way similar to that proposed in Milewski and Keller (1996). They all possess one common characteristics, namely, they are perturbations of the linear wave equation that take into account the effects of small nonlinearity and dispersion.

Various generations of the classical Boussinesq equation have been studied in various aspects, such as the well-posedness of the Cauchy problem for the following equation

$$u_{tt} = -u_{xxxx} + u_{xx} + (f(u))_{xx}. \tag{2}$$

In Tsutsumi and Matahashi (1991), the local and global well-posedness has been proved by firstly transforming the problem into a system of nonlinear Schrödinger equations. In Bona and Sachs (1988), it has been established that the special solitary wave solutions of (2) are nonlinearly stable for a range of wave speeds. The authors also drew a conclusion that the initial data lying relatively close to a stable solitary wave evolves into a global solution of the equation. This contrasts with the blow-up results obtained in Kalantarov and Ladyzhenskaya (1978) and Levine and Sleeman (1985), which show that certain solutions of the Cauchy problem cannot exist for all time in general, and casts some additional light on the results obtained for (1) by Deift, Tomei and Trubowitz (1982) via the inverse scattering theory.

In Varlamov (1996), an initial value problem for the following damped Boussinesq equation was considered

$$u_{tt} - 2bu_{txx} = -\alpha u_{xxxx} + u_{xx} + \beta(u^2)_{xx}, \tag{3}$$

where the second term on the left-hand side is responsible for dissipation, α and b are positive constants, β = constant $\in R$. The global solution of a small initial data problem for equation (3) was obtained in Varlamov (1996) under various assumptions regarding the smoothness of initial data.

In this paper, we study the existence of global solution of the initial value problem for the following damped Boussinesq equation

$$u_{tt} - 2bu_{txx} = -\alpha u_{xxxx} + u_{xx} + \beta(f(u))_{xx}, \tag{4}$$

where the constants are the same as those in (3).

Let $C([0,T], H^s(R))$ denote the Sobolev space with norm defined by

$$\|u\|_s = \sup_{t\in[0,T]} \|u\|_{H^s(R)},$$

where

$$\|u\|_{H^s(R)} = \int_{-\infty}^{\infty} <\xi>^s |\hat{u}(\xi,t)|^2 d\xi,$$

$$F[u] = \hat{u}(\xi, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ix\xi} u(x,t) dx,$$

$$< \xi > = \left(1 + |\xi|^2\right)^{\frac{1}{2}}.$$

We also require that the nonlinear term $f(u)$ satisfies

$$\|f(u)\|_{H^j(R)} \leq C \|u\|^p_{H^j(R)}, \qquad (*)$$

where $p \geq 1$, $j = s$ or $j = s+2$, $s > \frac{1}{2}$.

The main result obtained in our study is presented in the following section. The work can be regarded as the extension of the existence theorems discussed in Varlamov (1996) and Tsutsumi and Matahashi (1991). It should also be addressed that the methods used in this study are different from those used in Varlamov (1996) and Tsutsumi and Matahashi (1991).

2 Existence theorem

We pose the Cauchy problem for equation (4) adding the following initial conditions

$$u(x,0) = \varepsilon\varphi(x), \quad u_t(x,o) = \varepsilon\psi(x), \quad x \in (-\infty, \infty), \qquad (5)$$

where ε is a small positive parameter. As our main tool for solving (4) - (5) is the Fourier transform in x, we assume that $\varphi(x) \in H^{s+2}(R)$ and $\psi(x) \in H^s(R)$. In addition, we let

$$\chi_s(t) = C([0,T], H^{s+2}(R)) \cap C^1([0,T], H^s(R)),$$

with

$$\chi_s(\infty) = C([0,\infty], H^{s+2}(R)) \cap C^1([0,\infty], H^s(R)).$$

Theorem 1 *Suppose that $\varphi(x) \in H^{s+2}(R)$ and $\psi(x) \in H^s(R)$, $s > \frac{1}{2}$, $\alpha > b^2 > 0$ and $f(u)$ satisfies condition $(*)$. For a small $\varepsilon_0 > 0$, if $0 < |\varepsilon| \leq \varepsilon_0 \ll 1$, there must exist an $u(x,t) \in \chi_s(\infty)$ which solves the initial-value problem (4)-(5).*

Proof: Taking the Fourier transform with respect to x, we can obtain the following equations from (4)-(5)

$$\hat{u}_{tt}(\xi,t) + 2b\xi^2 \hat{u}_t(\xi,t) + \left(\alpha\xi^4 + \xi^2\right)\hat{u}(\xi) = \beta\xi^2 \widehat{f(u)},$$
$$\hat{u}(\xi,0) = \varepsilon\hat{\varphi}(\xi),$$
$$\hat{u}_t(\xi,0) = \varepsilon\hat{\psi}(\xi).$$

Solving above equations, we get

$$\hat{u}(\xi,t) = \varepsilon e^{-b\xi^2 t}\left\{\left[\cos(\sigma t) + b\xi^2\frac{\sin(\sigma t)}{\sigma}\right]\hat{\varphi}(\xi) + \frac{\sin(\sigma t)}{\sigma}\hat{\psi}(\xi)\right\}$$
$$-\frac{\beta\xi^2}{\sigma}\int_0^t \exp\left[-b\xi^2(t-\tau)\right]\sin(\sigma(t-\tau))\widehat{f(u)}d\tau, \tag{6}$$

where $\sigma = \xi\sqrt{k\xi^2+1}$, $k = \alpha - b^2 > 0$.

Let

$$u_0 = \varepsilon F^{-1}\left[e^{-b\xi^2 t}\left\{\left[\cos(\sigma t) + b\xi^2\frac{\sin(\sigma t)}{\sigma}\right]\hat{\varphi}(\xi) + \frac{\sin(\sigma t)}{\sigma}\hat{\psi}(\xi)\right\}\right], \tag{7}$$

where F^{-1} is the inverse transform of F. We immediately have

$$\|u_0\|_{s+2} \le \varepsilon\left(\|\psi\|_{s+2} + \|\psi\|_s\right) = \varepsilon A.$$

Now, we introduce the sequence $\{u_n\}$ as follows:

$$\hat{u}_n(\xi,t) = \hat{u}_0 - \frac{\beta\xi^2}{\sigma}\int_0^t \exp\left[-b\xi^2(t-\tau)\right]\sin\left[\sigma(t-\tau)\right]\hat{f}(u_{n-1})d\tau$$
$$(n = 1,2,3,\ldots) \tag{8}$$

It then follows, from (8), that

$$\|u_1\|_{s+2} \le \varepsilon A + \frac{\beta C_0}{kb}\|u_0\|_{s+2}^p$$
$$\le \varepsilon A + \frac{\beta C_0}{kb}(\varepsilon A)^p,$$

where C_0 is independent of ε. Choosing ε to be sufficiently small such that

$$\frac{\beta C_0}{kb}(\varepsilon A)^{p-1} < 1,$$

we have

$$\|u_1\|_{s+2} \le 2\varepsilon A,$$
$$\|u_2\|_{s+2} \le \varepsilon A + \frac{\beta C_0}{kb}(2\varepsilon A)^p.$$

Choosing ε to be suitably small such that

$$\frac{\beta C_0}{kb}(2\varepsilon A)^{p-1} < \frac{1}{2},$$

we have

$$\|u_2\|_{s+2} \le 2\varepsilon A.$$

By induction, we can easily derive that

$$\|u_n\|_{s+2} \leq 2\varepsilon A. \tag{9}$$

In the same proof as that of (9), we deduce that

$$\|\frac{\partial}{\partial t}u_n\|_s \leq 2\varepsilon A. \tag{10}$$

Therefore, we can conclude that there exists a sub-sequence $\{u_{n_k}\}$ of $\{u_k\}$, and that $\{u_{n_k}\}$ is a Cauchy sequence in the space $\chi_s(\infty)$. Since $\chi_s(\infty)$ is complete, there exists an $u \in \chi_s(\infty)$ such that $u_{n_k}(x,t)$ uniformly converges to u in the space $\chi_s(\infty)$. Clearly $u(x,t)$ satisfies (4) and (5). This completes the proof of the theorem.

REFERENCES

Bona, J. and Sachs, R. (1988). Global existence of smooth solutions and stability of solitary waves for a generalized Boussinesq equation. *Commun. Math. Phys.* 118, 12-29.

Deift, P., Tonei, C. and Trubowitz, E. (1982). Inverse scattering and the Boussinesq equation. *Commun. Pure. Appl. Maths.* 35, 567-628.

Kalantarov, V.K. and Ladyzhenskaya, O.A. (1978). The occurrence of collapse for quasilinear equations of parabolic and hyperbolic types. *J. Sov. Maths.* 10, 53-70.

Levine, H.A. and Sleeman, B.D. (1985). A note on the nonexistence of global solutions of initial-boundary value problems for the Boussinesq equation $u_{tt} = 3u_{xxxx} + u_{xx} - 12(u^2)_{xx}$. *J. Math. Anal. Appl.* 107, 206-210.

Miles, J.W. (1980). Solitary waves. *Ann. Rev. Fluid Mech.* 12, 11-43.

Milewski, P.A. and Keller, J.B. (1996). Three-dimensional water waves. *Studies in Applied Math.* 37, 149-166.

Tsutsumi, M. and Matahashi, T. (1991). On the Cauchy problem for the Boussinesq-type equation. *Math. Japonica* 36, 317-379.

Varlamov, V.V. (1996). On the Cauchy problem for the damped Boussinesq equation. *Differential and Integral Equations* 9, 619-634.

TESTING EQUALITY OF CORRESPONDING VARIANCES FROM MULTIPLE COVARIANCE MATRICES

KATRINA J. MATHER AND J.C.W. RAYNER
School of Mathematics and Applied Statistics
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: kjm02@uow.edu.au, john_rayner@uow.edu.au

The Wald statistic for testing equality of corresponding variances from multiple covariance matrices is introduced. The test is asymptotically optimal and derived under the assumption of independent and Normally distributed parent populations. The empirical size of the test in small samples is investigated via a Monte Carlo study. Use of the statistic is illustrated with a motivating example.

Keywords: asymptotically optimal test, empirical size study, variance homogeneity, Wald statistic.

1 Introduction

The problem we consider is motivated by a specific example, relating to data collected by Hermon Bumpus (1898). Eight morphological measurements were taken of sparrows injured in a storm. It is of interest to compare the corresponding variances of the survivors' and the nonsurvivors' covariance matrices, with the covariances considered as nuisance parameters; that is, not specified by the null and alternative hypotheses. An overall reduction in variance for the survivors would support the notion of stabilizing selection. The lower triangular part of the survivors' covariance matrix based on a sample of size 21 and given here to three decimal places is

$$\begin{pmatrix} 11.040 & & & & & & & \\ 9.100 & 17.500 & & & & & & \\ 1.557 & 1.910 & 0.531 & & & & & \\ 0.860 & 1.306 & 0.189 & 0.173 & & & & \\ 0.816 & 1.313 & 0.257 & 0.155 & 0.263 & & & \\ 0.045 & 0.078 & 0.016 & 0.007 & 0.013 & 0.001 & & \\ 0.532 & 0.460 & 0.145 & 0.022 & 0.017 & 0.003 & 0.110 & \\ 1.279 & 0.895 & 0.244 & 0.134 & 0.043 & -0.001 & 0.089 & 0.576 \end{pmatrix}.$$

The lower triangular part of the non-survivors' covariance matrix based on a sample of size 28 and given here to three decimal places is

$$\begin{pmatrix} 15.070 & & & & & & & \\ 17.190 & 32.550 & & & & & & \\ 2.243 & 3.398 & 0.728 & & & & & \\ 1.750 & 2.938 & 0.468 & 0.434 & & & & \\ 1.442 & 2.425 & 0.484 & 0.400 & 0.520 & & & \\ 0.101 & 0.195 & 0.029 & 0.028 & 0.028 & 0.002 & & \\ 0.914 & 1.546 & 0.291 & 0.227 & 0.284 & 0.015 & 0.213 & \\ 2.931 & 4.042 & 0.554 & 0.505 & 0.516 & 0.034 & 0.341 & 1.323 \end{pmatrix}.$$

The variables, in order, are total length, alar extent, length of beak and head, length of humerus, length of femur, length of tibio-tarus, width of skull, length of keel of sternum. All measurements are in millimetres. It should be noted that some previous analyses of the data have contained certain discrepancies. For example, the treatment found in Manly and Rayner (1987) analyses the data with some of the variables measured in millimetres and others in inches. The copy of the data found in Manly (1985) provides some measurements in inches and others in millimetres. We have made an effort to convert all measurements to millimetres before conducting our analysis; hence the covariance matrices given in this paper may appear somewhat different from those in other publications. The scale invariance property of our proposed Wald statistic makes our test indifferent to any such discrepancy. However we prefer to convert all variables to the one unit of measurement for clarification.

If the population covariance matrices are Σ_S and Σ_N, then we test the null hypothesis H: $(\Sigma_S)_{uu} = (\Sigma_N)_{uu}$, for $u = 1, \ldots, 8$, against K: not H. Testing hierarchically, Manly and Rayner (1987) found that the corresponding covariances (correlations) are collectively not equivalent when testing at the 0.1% level. Their hierarchic test would have proceeded to assess corresponding variances had corresponding covariances been judged collectively equal.

Section 2 of the paper defines the general problem that we wish to investigate and introduces the proposed Wald statistic; Section 3 provides the results from a preliminary simulation study and Section 4 illustrates use of the test statistic with the data collected by Bumpus (1898).

2 Formal testing

In the scenario we envisage we have k random samples from p-variate Normal populations $N_p(\mu_i, \Sigma_i)$ for $i = 1, \ldots, k$. The ith random sample, $\mathbf{X}_1, \ldots, \mathbf{X}_{n_i}$, is of size n_i. We wish to test H: $(\Sigma_1)_{uu} = \ldots = (\Sigma_k)_{uu}$, for $u = 1, \ldots, p$,

against K: not H.

In the univariate case we have k populations and we wish to test for homogeneity of variance. If $S_i^2 = \sum_{j=1}^{n_i} (X_{ij} - \overline{X}_i)^2 / (n_i - 1)$ in which $\overline{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$, put

$$\mathbf{V} = \begin{pmatrix} S_1^2 \\ S_2^2 \\ \vdots \\ S_k^2 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & -1 \end{pmatrix} \text{ and } \hat{\mathbf{\Phi}} = \mathbf{CV}.$$

The Wald statistic for testing H against K is $\hat{\mathbf{\Phi}}' \left(\widehat{\operatorname{Cov}\, \hat{\mathbf{\Phi}}}\right)^{-1} \hat{\mathbf{\Phi}}$. The usable form has substituted the maximum likelihood estimators of the parameters in the inverse of the covariance matrix of $\hat{\mathbf{\Phi}}$, or estimators that are as efficient. Finding this statistic requires the nontrivial inversion of the tri-diagonal matrix

$$\operatorname{Cov} \hat{\mathbf{\Phi}} = \begin{pmatrix} v_1 + v_2 & -v_2 & 0 & \cdots & \cdots & 0 \\ -v_2 & v_2 + v_3 & -v_3 & 0 & \cdots & \vdots \\ 0 & -v_3 & v_3 + v_4 & -v_4 & \cdots & \vdots \\ \vdots & \cdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \cdots & 0 & -v_{k-2} & v_{k-2} + v_{k-1} & -v_{k-1} \\ 0 & \cdots & \cdots & 0 & -v_{k-1} & v_{k-1} + v_k \end{pmatrix},$$

where

$$v_i = \operatorname{Var}(S_i^2) = \frac{2\sigma_i^4}{n_i - 1}, \quad i = 1, \ldots, k.$$

The ultimate form of the Wald statistic for k univariate samples is

$$\widehat{W}_1 = \frac{\sum_{r=1}^{k-1} \sum_{s>r}^{k} (n_r + 1)(n_s + 1)\left(S_r^2 - S_s^2\right)^2 / \left(S_r^4\, S_s^4\right)}{2 \sum_{r=1}^{k} (n_r + 1) / S_r^4}$$

With $k = 2$ this reduces to

$$\widehat{W}_1 = \frac{\left(S_1^2 - S_2^2\right)^2}{2\left(S_1^4/(n_1+1) + S_2^4/(n_2+1)\right)},$$

as derived in Rayner (1997).

These usable forms of the Wald statistic use optimal estimators of the σ_r^4, $r = 1, \ldots, k$. Using the well-known Rao-Blackwell theorem, it follows that unique estimators having minimum variance in the class of unbiased estimators are

$$\hat{\sigma}_r^4 = (n_i - 1)\, S_i^4 / (n_i + 1), \quad r = 1, \ldots, k.$$

It is not necessarily the case that such estimators will result in tests with achieved size close to the nominal size, or with greater power than tests that use other estimators. However, in large samples these tests should not be inferior to those that use other estimators.

When the data are p-variate rather than univariate and when there are k populations to be compared, let S_{ur}^2 be the unbiased sample variance of the uth variable in the rth population. Define

$$\widehat{W}_{u1} = \frac{\sum_{r=1}^{k-1} \sum_{s>r}^{k} (n_r + 1)(n_s + 1)\left(S_{ur}^2 - S_{us}^2\right)^2 / \left(S_{ur}^4\, S_{us}^4\right)}{2\sum_{r=1}^{k} (n_r + 1)/S_{ur}^4}$$

to be the univariate Wald statistic in the uth of p variables when sampling from k populations. The Wald statistic for k p-variate samples can be shown to be

$$\widehat{W}_{.p} = \widehat{W}_{11} + \ldots + \widehat{W}_{p1}.$$

Under the null hypothesis, $\widehat{W}_{.p}$ follows the $\chi^2_{p(k-1)}$ distribution.

3 Simulation study

Small samples versus asymptotics

Wald statistics asymptotically follow the χ^2 distribution; different Wald statistics will converge to their asymptotic distribution at varying rates. For particular Wald statistics, the question arises of how large a sample size is required for the χ^2 distribution to be satisfactory. To investigate this, we

conducted a Monte Carlo assessment of the empirical size of the statistic for nominal 5% level tests.

For various combinations of p, k, and n, we generated k random samples of equal size n; each sample is generated from a standard p-variate Normal parent population with zero mean vector and identity covariance matrix. The simulations discussed here are purely preliminary; more complex consideration as to the effects of nuisance parameters, etcetera will be given in later simulations. For each combination of p and k considered, the sample sizes were taken to be $n = 5, 10\ (10)\ 100$. The Wald statistic for the generated data is calculated and compared with the $\chi^2_{p(k-1)}$ 5% critical value. For each $p/k/n$ combination, this process was repeated 10,000 times and the empirical size calculated as the proportion of statistics that exceeded the $\chi^2_{p(k-1)}$ critical value. The empirical size of the Wald statistic was then plotted as a function of n. The graphs in Figure 1 below show a cross-section of the results obtained, for various values of p and k.

As expected, the minimum sample size required for the empirical size to converge to the nominal size (5%) depends strongly on the number of variables (p) and the number of samples (k). It is evident that for fixed p, there is more variation in the empirical size for different values of k than there is for different values of p given fixed k. Although both p and k affect the rate of convergence of the empirical size, it seems k has a greater affect of the two parameters. For a smaller number of samples, a sample size of around 50 seems large enough for the χ^2 distribution to accurately approximate the Wald statistic's distribution; however, for larger k, sample sizes of up to 150-200 may be required. As a rough rule of thumb it seems that if $pk < 20$, a sample size of at least 30 gives reasonable agreement between actual and nominal size, based on the results given.

Competitors

It is also of interest to compare the above performance of the Wald statistic with other competitor tests available in the literature. Here we only consider the comparison in the univariate ($p = 1$) case when the data are Normal. Two competing tests were chosen: the Likelihood Ratio (LR) test (more commonly known as Bartlett's test) and Levene's test. The LR test, like the Wald, is an asymptotically optimal test derived under Normality. For two samples, the LR test reduces to the well-known quotient of the sample variances S_1^2/S_2^2 and is Uniformly Most Powerful Unbiased for given significance level. Bartlett's test appears to be the most popular choice of test when Normality may be safely assumed. Levene's test has been proposed as a robust alternative to the LR test and is one of the more common choices when

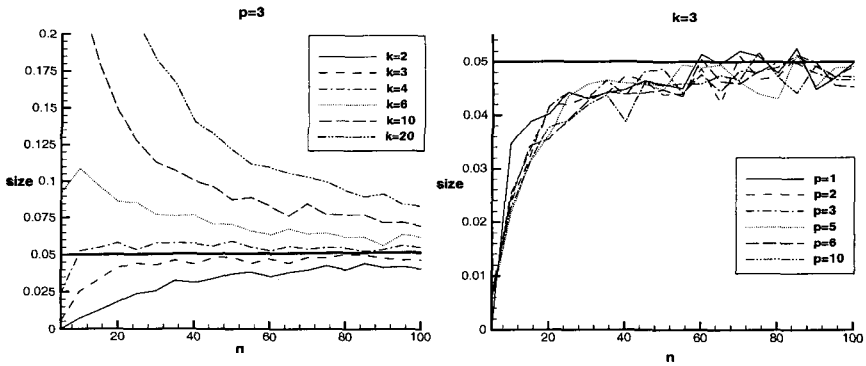


Figure 1. Empirical size of the Wald test for various p and k

Normality is in doubt. Levene's test performs a one-way ANOVA on residuals. The version employed here uses the sample mean in calculation of the residuals.

Figure 2 shows the empirical size of all three tests for two and six samples. It is obvious from the graphs that, unlike the Wald test, the LR test does not require any minimum sample size for its empirical size to match the nominal size of 5%. This is largely because the LR test employs a "Bartlett correction", a scaling constant derived solely to improve the approximation of the statistic's distribution in small samples. Our Wald test, as it stands, does not employ such a correction. However, we expect that once derived, the corresponding "Bartlett correction" for the Wald statistic will greatly improve the χ^2 approximation in small samples. We would then expect the size performance of our corrected test to be comparable with that of the LR. Note that for small samples, the empirical size of Levene's test is inferior to that of the LR test: Levene's test is an alternative to the LR test and is best recommended for situations involving non-Normality, see Levene (1960).

Bootstrapping

It has become evident that for small sample sizes, the χ^2 approximation to the distribution of the Wald statistic is inaccurate. In such samples, what difference, if any, does bootstrapping the test statistic make? For many of the p/k combinations considered above, and small to moderate sample sizes: $n = 10\ (20)\ 70$, random standard Normal samples were generated and used to bootstrap the Wald statistic. For each $p/k/n$ combination investigated,

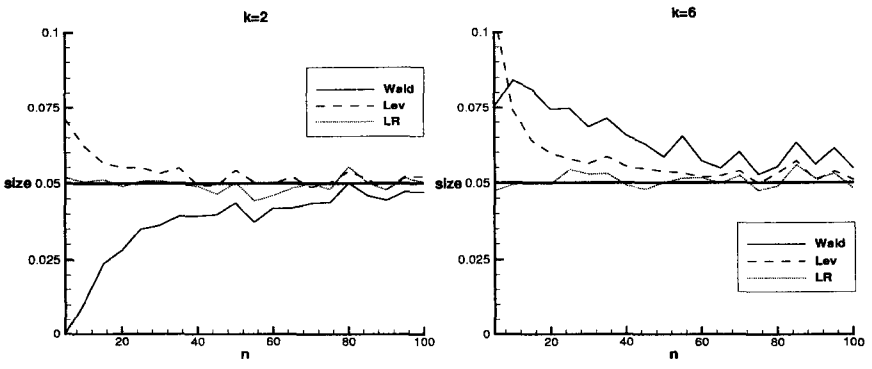


Figure 2. Empirical size of all three tests

100 random samples were generated and used to produce 100 pairs of (x, y) p-values. First, the Wald statistic was calculated on the random data and the p-value found using the $\chi^2_{p(k-1)}$ distribution. This is our x abscissa in Figure 3. Second, the data were resampled 99,999 times to obtain 99,999 values of the bootstrapped Wald statistic using method **B1** described in Section 4 below, resampling from the k samples separately. The associated bootstrapped p-value is then the proportion of times the statistics exceeded the value of the Wald statistic from the original data. This is our y ordinate in Figure 3. For each p/k combination, a graph of the y-values versus the x-values for each of the four sample sizes considered was produced; it is of interest to see how the bootstrapped p-values compare with those obtained from the χ^2 distribution.

The two graphs for $p = 5$, $k = 6$ and sample sizes of 10 and 30 are shown below in Figure 3. If there is perfect agreement between the two approaches, all points will lie on the line $y = x$. To aid in determining how large a sample size is required before the bootstrapped p-values closely match the χ^2 p-values, this line is displayed on each graph. For a sample size of ten, there is a noticeable difference between the p-values; those obtained using the χ^2 distribution tend to be smaller than the p-values found from bootstrapping. In particular, the majority of significant χ^2 p-values possess non-significant bootstrapped counterparts. For a sample size of 30, there exists closer agreement between the two sets of p-values. There was found to be little improvement obtained from increasing the sample size to 50 or 70. The most noticeable difference in agreement is that obtained by an increase in n from ten to 30.

The larger the sample size, the closer we expect the p-values to agree; the

aim here is to find the minimum n such that the agreement is satisfactory and such that increasing n further does not result in a significant improvement. Relating back to our rough rule of thumb from Section 3.1, here we have $p\,k = 30 > 20$ and hence we would expect a slightly larger sample size is needed in order to use the χ^2 distribution. For this particular p/k combination, we recommend bootstrapping the Wald test statistic for sample sizes that are smaller than around 40. Of course, this recommendation does depend on the values of the p and k parameters: for an increased number of samples we recommend a larger sample size before bootstrapping fails to result in a significant improvement over using the χ^2 distribution. We conclude that using the χ^2 distribution to obtain a p-value with small sample sizes inflates the chance of committing a Type I error, due to the inaccurate approximation of the χ^2 distribution and recommend bootstrapping the statistic in this case.

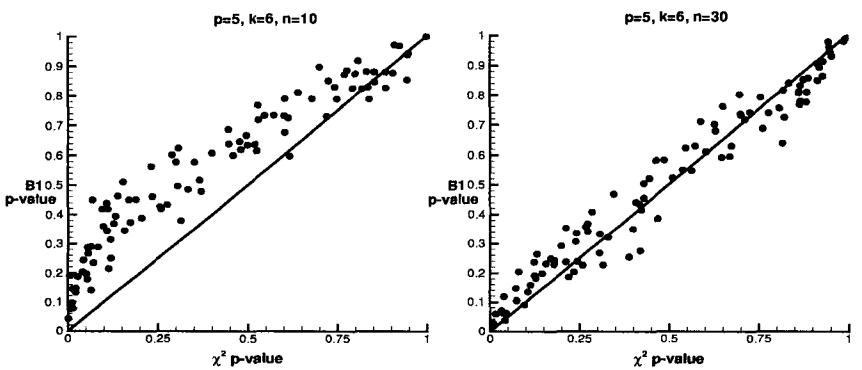


Figure 3. Bootstrapped p-values versus χ^2 p-values

4 Example

We now revisit Bumpus' motivating example. Simply by eye-balling each of the diagonal elements in the two covariance matrices, we see that each variable has a larger variance in the sample of non-survivors. If we apply the Wald test to formally test for a difference in the variances between the two samples, we find the statistic has a value of 21.047. The associated p-value, assuming the statistic is distributed as χ^2_8, is 0.0070. This is obviously significant at the 1% level. However, bearing in mind the poor performance of the χ^2 approximation for small sample sizes, we used simulation to calculate

the empirical size of the test as in Section 3.1. The empirical size at the nominal 5% level was calculated under Normality for the parameters present in the Bumpus problem: $p = 8$, $k = 2$, $n_1 = 21$, $n_2 = 28$. It was found to be 0.0227; roughly half the nominal size. This suggests use of the χ^2 distribution with the present parameters may be inadequate and bootstrapping the test a preferable alternative.

Table 1 shows a summary of p-values for the Wald test obtained from five different methods: using the χ^2 distribution, two methods of bootstrapping, a randomisation test and empirically from Monte-Carlo simulations. Method **B1** bootstraps the test statistic by resampling the data from the two samples of survivors and non-survivors separately; **B2** bootstraps by resampling from an aggregated data pool formed by combining the survivors and non-survivors together. **R** is a standard randomisation test: the two samples are combined into one and 21 observations randomly selected to constitute the survivors; the remaining 28 form the sample of non-survivors. The three data-based methods used 99,999 resampling repetitions each to obtain their associated p-values. To calculate the empirical p-value, 1,000,000 standard Normal random data sets were generated using the Bumpus values of the p, k, n_1 and n_2 parameters given above. For each data set, the Wald statistic was calculated. The empirical p-value is then the proportion of statistics which exceed the value of the Wald statistic from the Bumpus data: 21.047. Examining the results, the two bootstrap methods resulted in similar p-values to one another and agree closely with the p-value obtained using the χ^2 distribution. While the empirical p-value is slightly smaller, all four of these methods result in p-values that are significant at the 1% level. The p-value obtained using the randomisation test, while inflated compared to the other p-values, is still significant at the 5% level.

We conclude there is evidence of a difference in the variances of the eight morphological measurements between the survivors and non-survivors. Comparing the two covariance matrices, it appears that the variances are larger in the population of non-survivors. Our conclusion is consistent with Bumpus' hypothesised stabilizing selection phenomenon. The findings presented are in agreement with other analyses of the data contained in the literature; see Manly (1985) and Manly and Rayner (1987).

Table 1. p-values for the Bumpus data

Method	χ^2	B1	B2	R	empirical
p-value	0.0070	0.0054	0.0079	0.0263	0.0011

5 Conclusion

Monte Carlo investigation of the Wald statistic's empirical size in small samples found the χ^2 approximation to be unsatisfactory. Bootstrapping the test statistic is instead recommended for small samples. The minimum sample size required for the χ^2 distribution to serve as an accurate approximation is dependent upon the number of variables, p, and the number of samples, k. For most cases however, a sample size of 40-50 is recommended in order to use the χ^2 distribution. We stress that the simulation results presented in this paper are strictly preliminary; we expect future work involving extensive simulations to reveal further properties of our test statistic. For example, power and aspects such as the effect of nuisance parameters are yet to be considered. As previously mentioned, we expect to derive a "Bartlett correction" for the Wald statistic and anticipate this will greatly improve the χ^2 approximation to the small-sample distribution of the statistic, just as the correction did for the Likelihood Ratio test. We further hope to break down our Wald statistic into individual "components" and conduct least significant difference comparisons, which will allow us to ascertain which variables in what samples contribute to an overall rejection of the null hypothesis.

REFERENCES

Bumpus, H.C. (1898). The elimination of the unfit as illustrated by the introduced sparrow, passer domesticus. In *Biology Lectures.* Woods Hole: Marine Biology Laboratory, 209-226.

Levene, H. (1960). Robust tests for equality of variances. In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, Olkin, I., Ghurye, S.G., Hoeffding, W., Madow, W.G. and Mann, H.B. (editors). Palo Alto, Calif.: Stanford University Press, 278-292.

Manly, B.F.J. (1985). *The Statistics of Natural Selection on Animal Populations.* London: Chapman and Hall.

Manly, B.F.J. and Rayner, J.C.W. (1987). The comparison of sample covariance matrices using likelihood ratio tests. *Biometrika* 74, 841-847.

Rayner, J.C.W. (1997). The asymptotically optimal tests. *J.R.S.S., Series D (The Statistician)* 46, 337-346.

CONTEXTUAL IMAGE FUSION BASED ON MARKOV RANDOM FIELDS AND ITS APPLICATIONS TO GEO-SPATIAL IMAGE ENHANCEMENT

YOJI MORISAKI
Graduate School of Engineering
Hiroshima University, Kagamiyama 1-7-1
Higashi-Hiroshima 739-8521, Japan.
E-mail: morisaki@mis.hiroshima-u.ac.jp

RYUEI NISHII
Faculty of Integrated Arts and Sciences
Hiroshima University, Kagamiyama 1-7-1
Higashi-Hiroshima 739-8521, Japan.
E-mail: nishii@mis.hiroshima-u.ac.jp

Consider contextual data fusion of multispectral spatial imagery with different spatial resolutions. We introduce a new method for image enhancement based on Markov random fields so as to take contextual information into account. The low-resolution data are corrected by the high-resolution data and by the clustering result of the pixels. Our method is applied to geo-spatial data sensed by an artificial satellite, and shows a good performance.

Keywords: data fusion, ICM algorithm, image segmentation, MAP estimate.

1 Introduction

Data fusion with multispectral images with different spatial resolutions is an important and basic technology in satellite image analysis. The technology is widely required because many satellites are equipped with sensors of different resolutions. In this article, we consider a problem to improve multispectral low-resolution images by multispectral high-resolution images. The numbers of spectral bands with low resolution and with high resolution are set to ℓ and h, respectively.

As an example, consider the case such that multispectral low-resolution data and panchromatic high-resolution data are given, i.e., $\ell \geq 3$ and $h = 1$. For instance, the spatial resolution of visible lights of the satellite IKONOS is 4m, whereas that of the panchromatic sensor is 1m. In this case, $\ell = 3$ and $h = 1$. Also, the resolutions of SPOT are respectively 20m and 10m ($\ell = 3$, $h = 1$), and those of LANDSAT 7 are 30m and 15m ($\ell = 6$, $h = 1$).

The panchromatic images aim to supply supplementary information for the low-resolution images.

The commonly-used method for the enhancement of colored images is based on the Hue-Saturation-Value (HSV) transform Smith (1978). The three-dimensional data corresponding to the visible lights are converted to HSV, after which the values are replaced by the panchromatic values of the higher resolution. Then, the result is transformed back, and an enhanced colored image is obtained. However, the HSV method reduces spectral information, and is valid only when the number of low-resolution images is three. We note that the HSV is a non-linear method.

For the enhancement of images, linear methods are proposed, see e.g., Núñez et al. (1999). Many references can be found therein. These methods may keep local information and preserve the spectral information better than the HSV method. Another method for image correction is based on non-contextual clustering Zhukov et al. (1995). This method may be powerful, but needs much computation and overfits to the data. A Bayes method is found in a review paper Zhukov et al. (1999). A fully-statistical approach is proposed by Nishii et al. (1996). They derived a predictor based on the conditional expectation given the high-resolution data. However, this method also overfits to the data. A common drawback of these methods is that spatial information, which is important for geo-spatial data, is ignored. To take spatial information of images into account, another statistical approach so-called cokriging is proposed by Morisaki and Nishii (2000).

In this paper, we propose a new enhancement method taking contextual information into account through Markov random fields (MRFs), which provide a theoretically robust and mathematically tractable way of characterizing contextual information. MRFs are widely used in image segmentation and restoration, see, e.g., Chapter 13 of McLachlan (1992). The application of MRF for the clustering of multispectral magnetic resonance images is illustrated in McLachlan et al. (1996).

Our enhancement method is as follows. First, we classify all pixels into one of clusters by using Gaussian MRFs (GMRFs). Then, the low-resolution data are corrected by predictors based on parameters of the clusters.

In Section 2, we introduce notations and distributional assumptions on spectral data. In Section 3, predictors for low-resolution geo-spatial data based on the conditional distribution are derived. In Section 4, MRFs for contextual clustering are reviewed. In Section 5, a contextual prediction procedure based on the ICM (iterated conditional modes) algorithm due to Besag (1986) is presented. In Section 6, the proposed procedure is examined through the actual LANDSAT data, and compared with the standard methods. In

Section 7, conclusions are given.

2 Distributions of spectral data

Our aim is to correct low-resolution data at a pixel by using high-resolution data in its neighborhood. So, we focus our attention on a local window, and move it in the whole image. Figure 1 illustrates two neighborhoods consisting of four or eight pixels. They will be referred as W4 and W8 in the sequel.

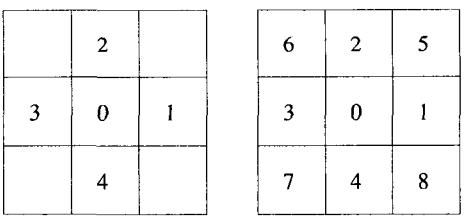


Figure 1. Neighborhoods of the pixel 0
left: Four-adjacent-pixels window (W4),
right: All-square-pixels window (W8)

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Figure 2. Local regions R_{16}

Suppose that images with two spatial-resolutions are given. We will derive the method to correct low-resolution data at a fixed pixel by using high-resolution data observed in the neighborhood W4 or W8. Then, the method is applied to all pixels by shifting the window.

Let

$$\boldsymbol{Z}_i = \begin{pmatrix} \boldsymbol{H}_i \\ \boldsymbol{L}_i \end{pmatrix} : (h+\ell) \times 1, \quad i = 0, 1, \ldots, 8, \tag{1}$$

be a random vector corresponding to spectral data at the ith pixel, where $\boldsymbol{H}_i : h \times 1$ and $\boldsymbol{L}_i : \ell \times 1$ represent random vectors in a local window of the images. Suppose that observations on $\boldsymbol{H}_i$ are available, say $\boldsymbol{h}_i$, $i = 0, 1, \ldots, 8$, whereas individual random vector $\boldsymbol{L}_i$ are not exactly observed. Only smoothed vectors, say $\boldsymbol{\ell}_i$, are available. Henceforth, the observed vectors $\boldsymbol{h}_i$ and $\boldsymbol{\ell}_i$ are respectively called by high-resolution data and low-resolution data.

Here, we regard the low-resolution vector $\boldsymbol{\ell}_0$ is an observation of average vector $\sum_{j=0}^{4} \boldsymbol{L}_j/5 \equiv \overline{\boldsymbol{L}}_5$ or $\sum_{j=0}^{8} \boldsymbol{L}_j/9 \equiv \overline{\boldsymbol{L}}_9$ corresponding to the window W4

or W8. We also assume that each pixel belongs to one of g clusters $C_1, \ldots, C_g$. A label of the cluster at the pixel i is denoted by $y_i (i = 1, \ldots, g)$. The label y_i is supposed to be a realization of a discrete random variable Y_i, and the distribution is discussed in Section 4. In Sections 2 and 3, we assume that all labels $Y_i = y_i$ are given.

Suppose that the conditional distribution of $(\boldsymbol{H}_i', \boldsymbol{L}_i')'$ given $Y_i = y_i$ is an $(h + \ell)$-variate normal distribution with mean vector $(\boldsymbol{\mu}(y_i)', \boldsymbol{\nu}(y_i)')'$ and common variance-covariance matrix Σ denoted by

$$\begin{pmatrix} \boldsymbol{H}_i \\ \boldsymbol{L}_i \end{pmatrix} \Big| \; y_i \sim \mathrm{N}_{h+\ell}\left(\begin{pmatrix} \boldsymbol{\mu}(y_i) \\ \boldsymbol{\nu}(y_i) \end{pmatrix}, \Sigma \right), \quad \Sigma \equiv \begin{pmatrix} \Sigma_{hh} & \Sigma_{h\ell} \\ \Sigma_{\ell h} & \Sigma_{\ell\ell} \end{pmatrix}. \tag{2}$$

Further, we assume that they are independent.

3 Contextual prediction for low-resolution data

We will propose a predictor for low-resolution data by clustering all pixels.

Prediction at the center of a local window

Consider the window W4 of Figure 1 consisting of four pixels. Under the normality assumption (2) and the conditional independence of $(\boldsymbol{H}_i', \boldsymbol{L}_i')'|y_i$ for $i = 0, 1, \ldots, 4$, the conditional distribution of the vector $\boldsymbol{Z} \equiv (\boldsymbol{H}_0', \boldsymbol{L}_0', \ldots, \boldsymbol{H}_4', \boldsymbol{L}_4')' : (5h + 5\ell) \times 1$ given label vector $\boldsymbol{y}_5 = (y_0, y_1, \ldots, y_4)'$ is again a normal distribution with mean vector $\boldsymbol{\xi} \equiv (\boldsymbol{\mu}(y_0)', \boldsymbol{\nu}(y_0)', \ldots, \boldsymbol{\mu}(y_4)', \boldsymbol{\nu}(y_4)')'$ and variance-covariance matrix $\Lambda \equiv \Sigma \otimes I_5$, where $\otimes$ and I denote the Kronecker product and the identity matrix respectively.

We shall derive the joint distribution of $\boldsymbol{L}_0$ and all random vectors whose observations are available. The random vector $(\boldsymbol{H}_0', \boldsymbol{H}_1', \ldots, \boldsymbol{H}_4', 5\overline{\boldsymbol{L}}_5', \boldsymbol{L}_0')' : (5h + 2\ell) \times 1$ is expressed by a linear mapping $A\boldsymbol{Z}$ with a matrix A defined by

$$A = \begin{pmatrix} I_h & O & O & O & \cdots & O & O \\ O & O & I_h & O & \cdots & O & O \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ O & O & O & O & \cdots & I_h & O \\ O & I_\ell & O & I_\ell & \cdots & O & I_\ell \\ O & I_\ell & O & O & \cdots & O & O \end{pmatrix} : (5h + 2\ell) \times (5h + 5\ell). \tag{3}$$

Hence the random vector $A\boldsymbol{Z}$ given label vector $\boldsymbol{y}_5$ follows the multivariate normal distribution with mean vector $A\boldsymbol{\xi}$ and variance-covariance matrix

$A\Lambda A'$, i.e.,

$$\begin{pmatrix} \boldsymbol{H}_0 \\ \vdots \\ \boldsymbol{H}_4 \\ 5\overline{\boldsymbol{L}}_5 \\ \boldsymbol{L}_0 \end{pmatrix} \Bigg| \; \boldsymbol{y}_5 \sim \mathrm{N}_{5h+2\ell}\left(\begin{pmatrix} \boldsymbol{\mu}(y_0) \\ \vdots \\ \boldsymbol{\mu}(y_4) \\ 5\overline{\boldsymbol{\nu}}_5 \\ \boldsymbol{\nu}(y_0) \end{pmatrix}, \begin{pmatrix} \Sigma_{hh} & \cdots & O & \Sigma_{h\ell} & \Sigma_{h\ell} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ O & \cdots & \Sigma_{hh} & \Sigma_{h\ell} & O \\ \Sigma_{\ell h} & \cdots & \Sigma_{\ell h} & 5\Sigma_{\ell\ell} & \Sigma_{\ell\ell} \\ \Sigma_{\ell h} & \cdots & O & \Sigma_{\ell\ell} & \Sigma_{\ell\ell} \end{pmatrix} \right), \tag{4}$$

where $\overline{\boldsymbol{\nu}}_5 = \sum_{j=0}^{4} \boldsymbol{\nu}(y_j)/5$. Using the joint distribution (4), we can derive the conditional distribution of $\boldsymbol{L}_0$ as follows. The covariance matrix $A\Lambda A'$ in (4) is partitioned as $A\Lambda A' = \begin{pmatrix} \Xi_{11} & \Xi_{12} \\ \Xi_{21} & \Sigma_{\ell\ell} \end{pmatrix}$ with size $\Xi_{11} : (5h+\ell) \times (5h+\ell)$ and $\Xi_{12} : (5h+\ell) \times \ell$. After some algebra we have

$$\Xi_{21}\Xi_{11}^{-1} = \frac{1}{5}(4\Sigma_{\ell h}\Sigma_{hh}^{-1}, -\Sigma_{\ell h}\Sigma_{hh}^{-1}, \ldots, -\Sigma_{\ell h}\Sigma_{hh}^{-1}, I_\ell), \tag{5}$$

$$\Sigma_{\ell\ell} - \Xi_{21}\Xi_{11}^{-1}\Xi_{12} = 4\Sigma_{\ell\ell \cdot h}/5, \quad \Sigma_{\ell\ell \cdot h} \equiv \Sigma_{\ell\ell} - \Sigma_{\ell h}\Sigma_{hh}^{-1}\Sigma_{h\ell}, \tag{6}$$

where $\Sigma_{\ell\ell \cdot h}$ denotes a conditional variance-covariance matrix. From the distribution (4) and the relations (5) and (6), the conditional distribution of $\boldsymbol{L}_0$ is derived as

$$\begin{aligned} &\boldsymbol{L}_0 \,|\{\boldsymbol{h}_0, \boldsymbol{h}_1, \ldots, \boldsymbol{h}_4, \boldsymbol{\ell}_0, \boldsymbol{y}_5\} \\ &\sim \mathrm{N}_\ell\left(\boldsymbol{\ell}_0 + \Sigma_{\ell h}\Sigma_{hh}^{-1}(\boldsymbol{h}_0 - \overline{\boldsymbol{h}}_5) + \boldsymbol{\Delta}(y_0) - \overline{\boldsymbol{\Delta}}_5, \ \frac{4}{5}\Sigma_{\ell\ell \cdot h} \right) \end{aligned} \tag{7}$$

where $\overline{\boldsymbol{h}}_5 = \sum_{j=0}^{4} \boldsymbol{h}_j/5$, $\boldsymbol{\Delta}(k) = \boldsymbol{\nu}(k) - \Sigma_{\ell h}\Sigma_{hh}^{-1}\boldsymbol{\mu}(k)$ $(k = 1, \ldots, g)$, $\overline{\boldsymbol{\Delta}}_5 = \sum_{j=0}^{4} \boldsymbol{\Delta}(y_j)/5$, and $\overline{\boldsymbol{\mu}}_5 = \sum_{j=0}^{4} \boldsymbol{\mu}(y_j)/5$. The vector $\boldsymbol{\Delta}(k) : \ell \times 1$ means the residual vector such that the mean vector $\boldsymbol{\nu}(k)$ of low-resolution data is approximated by the regression due to the mean vector $\boldsymbol{\mu}(k)$ of high-resolution. It is obvious that the conditional expectation of (7) minimizes the mean squared prediction errors. Thus, a predictor for $\boldsymbol{L}_0$ at the center pixel is proposed by

$$\widehat{\boldsymbol{L}}_{W4} = \boldsymbol{\ell}_0 + \Sigma_{\ell h}\Sigma_{hh}^{-1}(\boldsymbol{h}_0 - \overline{\boldsymbol{h}}_5) + \boldsymbol{\Delta}(y_0) - \overline{\boldsymbol{\Delta}}_5. \tag{8}$$

We note that the predictor due to Nishii, Kusanobu and Tanaka (1996) was proposed by the same conditional distribution (7) as

$$\widehat{\boldsymbol{L}}_0 = \boldsymbol{\ell}_0 + \Sigma_{\ell h}\Sigma_{hh}^{-1}(\boldsymbol{h}_0 - \overline{\boldsymbol{h}}_5). \tag{9}$$

They did not consider the categories, and assumed that the term $\boldsymbol{\Delta}(y_0) - \overline{\boldsymbol{\Delta}}_5$ is negligible. This reduction may cause the predictor (9) to overfit the data when the dimension h is large.

Next, we consider spectral data observed in the window W8, see the right hand side of Figure 1. By the similar discussion leading the formula (8), the predictor at the center pixel of W8 is given by

$$\widehat{\boldsymbol{L}}_{W8} = \boldsymbol{\ell}_0 + \Sigma_{\ell h}\Sigma_{hh}^{-1}(\boldsymbol{h}_0 - \overline{\boldsymbol{h}}_9) + \boldsymbol{\Delta}(y_0) - \overline{\boldsymbol{\Delta}}_9, \tag{10}$$

where $\overline{\boldsymbol{h}}_9$ and $\overline{\boldsymbol{\Delta}}_9$ are similarly defined as the window W4.

The predictors (8) and (10) are proposed for improving the predictor (9). Hence they were derived under the same assumption of homogeneous covariance matrices, as assumed in Nishii, Kusanobu and Tanaka (1996).

Prediction in non-overlapping square regions

In the previous subsection, we take a pixel in the image at first, then, we consider its neighborhood. So, neighborhoods of adjacent pixels overlap each other, which means that two predictors use common high-resolution data. We consider here another degradation setup of low-resolution images. We partition the whole region into non-overlapping squares consisting of $m(= 2^2$ or $4^2)$ pixels. Then, we suppose that the low-resolution images are obtained by the averaging process in each region.

Let R_m be a square region with m pixels. Figure 2 shows the local region R_{16}. It is assumed that each random vector $\boldsymbol{H}_i$ is observed for $i = 1, 2, \ldots, m$. Concerning $\boldsymbol{L}_i$, however, only averaged vector $\sum_{j=1}^{m} \boldsymbol{L}_j/m \equiv \overline{\boldsymbol{L}}_m$ is observed in the region R_m. Observations of $\boldsymbol{H}_i$ and $\overline{\boldsymbol{L}}_m$ are respectively denoted by $\boldsymbol{h}_i$ and $\overline{\boldsymbol{\ell}}_m$. Under these assumptions, we derive the predictor for $\boldsymbol{L}_i$ by incorporating mean vectors on respective clusters as follows:

$$\widehat{\boldsymbol{L}}_i = \overline{\boldsymbol{\ell}}_m + \Sigma_{\ell h}\Sigma_{hh}^{-1}(\boldsymbol{h}_i - \overline{\boldsymbol{h}}_m) + \boldsymbol{\Delta}(y_i) - \overline{\boldsymbol{\Delta}}_m \quad \text{for} \quad i = 1, \ldots, m, \tag{11}$$

where $\overline{\boldsymbol{h}}_m$ and $\overline{\boldsymbol{\Delta}}_m$ are similarly defined as in the previous subsection.

Note that all predictors proposed in this section can be derived without normality assumption. Our predictors give the minimum mean prediction error as long as moments of the feature vectors up to second degree are same as the normal distribution assumed here.

4 Spatial distributions due to MRFs

In this section, we consider a spatial model due to MRF for contextual clustering. We suppose that there is a rectangular scene S consisting of n pixels. All pixels in S are numbered from $i = 1$ to n. We define the following vectors

with all high-resolution data or labels by

$$\boldsymbol{H} \equiv \begin{pmatrix} \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_n \end{pmatrix} : hn \times 1, \quad \boldsymbol{Y} \equiv \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix} \text{ and } \boldsymbol{Y}_{-i} \equiv \boldsymbol{Y} \text{ with deleted } Y_i. \quad (12)$$

By the distributional assumption (2), it follows that $\boldsymbol{H}_i|y_i \sim \mathrm{N}_h(\boldsymbol{\mu}(y_i), \Sigma_{hh})$ and they are independent for $i = 1, \ldots, n$. Hence, the conditional density of all high-resolution vectors is given by the product of the cluster-conditional densities as

$$\prod_{i=1}^{n} f(\boldsymbol{h}_i|y_i) \equiv \prod_{i=1}^{n} |2\pi\Sigma_{hh}|^{-1/2} \exp\Big[-\{\boldsymbol{h}_i - \boldsymbol{\mu}(y_i)\}' \Sigma_{hh}^{-1} \{\boldsymbol{h}_i - \boldsymbol{\mu}(y_i)\}/2\Big]. \quad (13)$$

Further, we assume that the cluster-label vector $\boldsymbol{Y}$ follows a locally-dependent MRF, i.e., the conditional probability of Y_i given all labels except y_i is given by

$$\begin{aligned} &\Pr\{Y_i = k \mid \boldsymbol{y}_{-i}\} \\ &= \exp\left\{-\beta \sum_{j \in U_r(i)} D(k, y_j)\right\} \Big/ \sum_{k'=1}^{g} \exp\left\{-\beta \sum_{j \in U_r(i)} D(k', y_j)\right\} \end{aligned} \quad (14)$$

for $\beta \geq 0$ and $r = 1, \sqrt{2}, \ldots$, where $D(k, k') = \{\boldsymbol{\mu}(k) - \boldsymbol{\mu}(k')\}' \Sigma_{hh}^{-1} \{\boldsymbol{\mu}(k) - \boldsymbol{\mu}(k')\}$ denote the squared Mahalanobis distance between the normal populations C_k and $C_{k'}$ for $k, k' = 1, \ldots, g$, $U_r(i)$ indicates a set of pixels whose distance between the pixel i does not exceed a constant r (radius), and $\boldsymbol{y}_{-i} : (n-1) \times 1$ denotes a realization of $\boldsymbol{Y}_{-i}$, see (12). Figure 3 illustrates the neighborhood $U_1(i)$ with radius $r = 1$. The parameter β is a measure of spatial dependency of the MRF.

5 Prediction by the adaptive ICM algorithm

In the framework of the previous section, the problem is to estimate the label vector $\boldsymbol{Y}$ of the scene S by using the observed feature data $\boldsymbol{h}$ of high resolution. Then, it is possible to correct the low-resolution images by the predictor (8) or (10) or (11) based on the estimated categories. The estimation of the label vector can be performed by maximizing the posterior probability $\Pr\{\boldsymbol{Y} = \boldsymbol{y} \mid \boldsymbol{h}\}$ (MAP estimate). The maximization problem is clearly formidable because there are g^n possible vectors of $\boldsymbol{y}$. Simulated annealing due to Geman and Geman (1984) can be used to approximate the MAP estimate. However, this algorithm needs much computation and it is known that the precise estimate

of the spatial parameters does not necessarily result in good estimates, see Aykroyd and Green (1991) and Besag et al. (1991). In this paper, we use the iterative conditional mode (ICM) algorithm due to Besag (1986) instead. MAP estimates are obtained by the following adaptive ICM algorithm.

First, fix the parameters β and r. (We will find the optimal parameters β and r by the grid search.) Let $\boldsymbol{y}^{(0)}$ be the initial estimate of cluster label $\boldsymbol{y}$ obtained by a non-contextual rule. In our case, the K-means clustering algorithm Jain and Dubes (1988) is used for the initial segmentation of the given scene. Let $\boldsymbol{y}^{(t)}$ be the estimate after completion of the t-th cycle of the ICM algorithm. The $(t+1)$-st cycle for the iterative estimate of $\boldsymbol{y}$ and the predictions for low-resolution data are performed in the following EM algorithm with prediction step (EMP algorithm).

(i) E-step: The mean vectors $\boldsymbol{\mu}^{(t)}(k)$ and $\boldsymbol{\nu}^{(t)}(k)$ for $k = 1, \ldots, g$ and the common covariance matrix $\begin{pmatrix} \Sigma_{hh}^{(t)} & \Sigma_{h\ell}^{(t)} \\ \Sigma_{\ell h}^{(t)} & \Sigma_{\ell\ell}^{(t)} \end{pmatrix}$ are estimated by using the current estimate $\boldsymbol{y}^{(t)}$, the high-resolution data and the low-resolution data.

(ii) **(a)** M-step: Using the ordinary ICM algorithm, we update the current estimate $\boldsymbol{y}^{(t)}$ to $\boldsymbol{y}^{(t+1)}$ based on $\boldsymbol{\mu}^{(t)}(\cdot)$ and $\Sigma_{hh}^{(t)}$.

(b) Prediction-step: The low-resolution data at respective pixels are corrected consecutively by the predictor (8) or (10) or (11) with the mean vectors $\boldsymbol{\mu}^{(t)}(\cdot)$ and $\boldsymbol{\nu}^{(t)}(\cdot)$, and the regression matrix $\Sigma_{\ell h}^{(t)} \Sigma_{hh}^{(t)-1}$.

The algorithm is repeated until the image obtained by the prediction step converges. Note that the converged image depends on β and r because they are used in the M-step. Put the estimated label vector by $\widehat{\boldsymbol{y}}(\beta, r) = (\widehat{y}_1(\beta, r), ..., \widehat{y}_n(\beta, r))'$.

Next we consider selection of optimal parameters β and radius r. (Consequently we get the enhanced image based on the clustering result due to the selected parameters.) Besag (1986) took the pseudo-likelihood approach for this purpose, see Chapter 13.6 of McLachlan (1992). In this article, we maximizes the product of the following posterior probabilities.

In general, the posterior probability of the label $Y_i = k$ given the high-resolution vector $\boldsymbol{h}_i$ and the label vector $\boldsymbol{y}_{-i}$ except ith label is given by

$$\Pr\{Y_i = k | \boldsymbol{h}_i, \boldsymbol{y}_{-i}\} = \frac{f(\boldsymbol{h}_i \,|\, k) \Pr\{Y_i = k \,|\, \boldsymbol{y}_{-i}\}}{\sum_{k'=1}^{g} f(\boldsymbol{h}_i | k') \Pr\{Y_i = k' \,|\, \boldsymbol{y}_{-i}\}}, \tag{15}$$

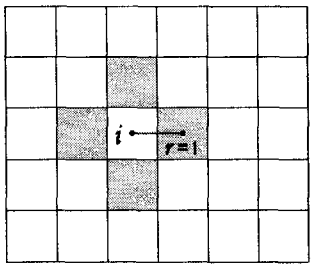


Figure 3. Neighborhood $U_1(i)$ of a pixel i with radius $r = 1$

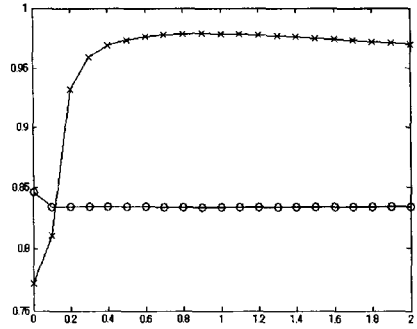


Figure 4. RSSE (-o-) and posterior $\overline{p}(\beta,\ \sqrt{8})$ (-×-) versus the parameter β for spatial dependency

where $f(\cdot|\cdot)$ is the cluster-conditional density defined by (13) and the conditional probability $\Pr\{Y_i = k' \mid \boldsymbol{y}_{-i}\}$ is given by the formula (14). Thus we assess the accuracy of the classification $\widehat{y}_i(\beta, r)$ given $\boldsymbol{h}_i$ and $\widehat{\boldsymbol{y}}_{-i}(\beta, r)$ through the marginal posterior (15) with $\boldsymbol{y} = \boldsymbol{y}(\beta, r)$, say $p_i(\beta,\ r)$. Then, the parameters β and r are determined by the product of the marginal posteriors $p_i(\beta,\ r)$ as

$$(\widehat{\beta},\ \widehat{r}) \equiv \arg\max_{\beta,\ r} \overline{p}(\beta,\ r), \quad \text{where } \overline{p}(\beta,\ r) = \prod_{i=1}^{n} p_i(\beta,\ r)^{1/n}. \tag{16}$$

Thus, we can derive the MAP estimate $(\widehat{\beta},\ \widehat{r})$. Finally, the enhanced image based on the optimal label vector $\widehat{\boldsymbol{y}}(\widehat{\beta},\ \widehat{r})$ is supposed to be the best enhanced image by the high-resolution data.

6 Contextual enhancement of geo-spatial images

We apply our methods to LANDSAT 7 data of Hasselt, Belgium taken on Oct. 18, 1999 sensed by ETM+ sensor. Our aim is to improve the color image of 30m resolution by the panchromatic image with 15m resolution. The proposed predictors given by the formulas (8), (10) and (11) are compared with the HSV method, the conditional expectation (9) due to Nishii et al. (1996) and the cokriging method Morisaki and Nishii (2000).

In order to compare the results numerically, we simulate images as follows. The original panchromatic image with resolution 15m and the visible bands 1 to 3 with resolution 30m are degraded to 30m and 60m resolution respectively.

Using the high-resolution panchromatic image, we correct the low-resolution multispectral images of size 150 × 150. The results are evaluated through the ratio of sums of squared errors (RSSE):

$$\text{RSSE} = \sum_{b=1}^{3} \sum_{i=1}^{150^2} \left(\widehat{L}_i^{(b)} - L_i^{(b)}\right)^2 \Big/ \sum_{b=1}^{3} \sum_{i=1}^{150^2} \left(\overline{L}_i^{(b)} - L_i^{(b)}\right)^2, \tag{17}$$

where $L_i^{(b)}$, $\overline{L}_i^{(b)}$ and $\widehat{L}_i^{(b)}$ denote respectively original values to be predicted, averaged values of the original values by 2×2 pixels, and the predicted values at pixels i in the whole square image. Here, $b(= 1, 2, 3)$ is an index for spectral bands corresponding to visible lights.

Table 1 tabulates optimal parameters β and maximum averaged posterior probabilities $\overline{p}(\beta,\ r)$ defined at (16) under the given radius r. RSSEs due to the proposed and the standard methods are also shown there. We note that the number of clusters selected in this application is five ($g = 5$). This table shows that our method derived by the region R_4 with $r = \sqrt{8}$ and $\beta = 0.8$ is best. We also see that the best result gives the maximum averaged posterior probability in the third column of the table.

Figure 4 shows RSSEs and averaged posterior probabilities $\overline{p}(\beta,\ \sqrt{8})$ for varying $\beta = 0.0(0.1)2.0$. We see from this figure that averaged posterior probability increases as RSSE decreases. These results imply that our strategy for choosing the optimal parameter succeeded in finding the optimal clustering result.

Figure 5 shows a panchromatic image of size 120 × 120 portion of the town of Hasselt, which includes urban area, arable land and river.

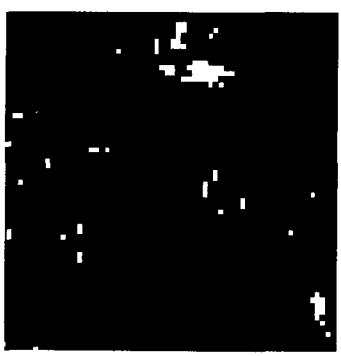


Figure 5. Panchromatic LANDSAT 7 image of size 120×120 courtesy ESA 1999 – distribution Eurimage, Hasselt, Belgium.

Five further color figures of the same region, namely:

Table 1. The ratios of sums of squared errors due to correction methods through the images of Hasselt of size 150×150.

r	$\widehat{\beta}(r)$	$\overline{p}(\widehat{\beta}(r),\ r)$	RSSE (%)		
			W4	W8	R_4
1	1.0	50.42	87.83	95.87	84.32
$\sqrt{2}$	0.4	89.49	87.17	95.09	83.51
2	0.6	95.88	87.18	95.12	83.80
$\sqrt{5}$	0.9	97.38	87.01*	94.96**	83.48*
$\sqrt{8}$	0.8	97.87**	87.01	95.00*	83.39**
3	0.6	97.73*	87.00**	95.00*	83.65
$\sqrt{10}$	0.2	95.71	87.04	95.09	83.67
$\sqrt{13}$	1.0	97.51	87.11	95.11	83.70
Conditional expectation					85.73
Cokriging					88.86
HSV					261.84

** and * denote the best and the second best values.

(b) the true colored image of size 120×120,

(c) the 2×2 averaged image of (b),

(d) the corrected image by our method with $g = 5$, $\widehat{r} = \sqrt{8}$, $\widehat{\beta} = 0.8$,

(e) the corrected image by Nishii et al. (1996),

(f) the corrected image by the HSV method

can be obtained at "http://www.mis.hiroshima-u.ac.jp/ ~nishii". Using the panchromatic image (a), we correct the low-resolution colored image (c) by the three methods and obtain the figures (d), (e) and (f). The optimal parameters $\widehat{r} = \sqrt{8}$ and $\widehat{\beta} = 0.8$ of the proposed method are chosen by maximizing the posterior probabilities (16). The number of clusters $g = 5$ is pre-assigned.

We see that the figure (f) due to HSV is quite different to the true image (b). The RSSE 261.85 % of (f) listed in Table 1 exceeds 100 %, which means that (f) is worse than the averaged image (c). On the other hand, both of the figures (d) due to our method and (e) due to Nishii et al. 1996) seem to be very close to the true image (b). By Table 1, it is shown that RSSE's are respectively given by 83.39 % and 85.73 %. Thus, our method shows a good performance.

7 Conclusion

In this paper, we proposed a new method for data fusion based on GMRF. The features of our approach are as follows.

- We considered two degradation cases of imagery, and proposed the predictors (8), (10) and (11) corresponding to the respective cases.
- The linear predictors for low-resolution data are derived in terms of the high-resolution data as well as the clusters of the pixels.
- The clustering method based on the adaptive ICM algorithm is employed and the spatial dependency parameter and radius of the locally-dependent MRF are successfully chosen by maximizing the averaged posteriors.
- Our methods were applied to LANDSAT data, and showed better performance than the standard methods.

We have proposed several predictors for image enhancement. Hence, the problem for selecting an appropriate predictor for given imagery still remains. The most important but difficult problem is to estimate the regression matrix $\Sigma_{\ell h}\Sigma_{hh}^{-1}$ used in our predictors because there is no exact observation on $\boldsymbol{L}_i$ itself. The predictor based on the better estimate would improve our approach. Further, we need to consider predictors under the case with heterogeneous covariance matrices.

Acknowledgments

The authors wish to thank the anonymous referee for careful reading, valuable comments and new references. The authors are also grateful for IEEE GRSS Data Fusion Committee for their permission to use the data set "grss_dfc_0002" of LANDSAT 7.

REFERENCES

Aykroyd, R.G. and Green, P.J. (1991). Global and local priors, and the location of lesions using gamma-camera imagery. *Philosophical Transactions on the Royal Society of London*, Series A 337, 323-342.

Besag, J. (1986). On the statistical analysis of dirty pictures. *J. Roy. Stat. Soc. B* 48, 259-302.

Besag, J., York, J. and Mollie, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43, 1-59.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images. *IEEE Trans. Pattern Anal. Machine Intell.* PAMI-6, 721-741.

Jain, A.K. and Dubes, R.C. (1988). *Algorithms for Clustering Data.* Englewood Cliffs, NJ: Prentice-Hall.

McLachlan, G.J. (1992). *Discriminant Analysis and Statistical Pattern Recognition.* New York: John Wiley.

McLachlan, G.J., Ng, S.K., Galloway, G. and Wang, D. (1996). Clustering of magnetic resonance images. *Proceedings of the American Statistical Association (Statistical Computing Section)*, Chicago, August 1996. Alexandria, Virginia: American Statistical Association, 12-17.

Morisaki, Y. and Nishii, R. (2000). Spatial enhancement of thermal infrared image based on cokriging. *Proc. Symp. 22nd Remote Sensing for Environmental Sciences*, 23-28.

Nishii, R., Kusanobu, S. and Tanaka, S. (1996). Enhancement of low spatial resolution image based on high resolution bands. *IEEE Trans. Geosci. Remote Sensing* 34, 1151-1158.

Núñez, J., Otazu, X., Fors, O., Prades, A., Palà, V., and Arbiol, R. (1999). Multiresolution-based image fusion with additive wavelet decomposition. *IEEE Trans. Geosci. Remote Sensing* 37, 1204-1211.

Smith, A.R. (1978). Color gamut transformation pairs. *Computers Graphics* 12, 12-19.

Zhukov, B., Oertel, D. and Lanzl, F. (1995). A multiresolution multisensor technique for satellite remote sensing. *Proc. 1995 Int. Geosci. Remote Sensing Symp.* I, 51-53.

Zhukov, B., Oertel, D., Lanzl, F. and Reinhäckel, G. (1999). Unmixing-based multisensor multiresolution image fusion, *IEEE Trans. Geosci. Remote Sensing* 37, 1212-1226.

AUTOREGRESSION, ESTIMATING FUNCTIONS, AND OPTIMALITY CRITERIA

URSULA U. MÜLLER
Fachbereich 3 Mathematik
Universität Bremen
Postfach 330 440, 28334 Bremen, Germany.
E-mail: uschi@math.uni-bremen.de

WOLFGANG WEFELMEYER
Fachbereich 6 Mathematik
Universität Siegen
Walter-Flex-Str. 3, 57068 Siegen, Germany.
E-mail: wefelmeyer@mathematik.uni-siegen.de

We consider Markov chains of order d that satisfy a conditional constraint of the form $E(a_\vartheta(\mathbf{X}_{i-1}, X_i) \mid \mathbf{X}_{i-1}) = 0$, where $\mathbf{X}_{i-1} = (X_{i-1}, \ldots, X_{i-d})$. These comprise quasi-likelihood models and nonlinear and conditionally heteroscedastic autoregressive models with martingale innovations. Estimators for ϑ can be obtained from estimating equations $\sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) = 0$. We review different criteria for choosing good weights $W_\vartheta(\mathbf{X}_{i-1})$. They usually lead to weights that depend on unknown features of the transition distribution and must be estimated. We compare the approach via estimating functions with other ways of constructing estimators for ϑ, and discuss efficiency of the estimators in the sense of Hájek and LeCam. Analogous comparisons may be made for regression models.

Keywords: generalized quasi-likelihood, extended quasi-likelihood, ARCH model, generalized method of moments, conditional least squares, influence function, gradient, variance bound.

1 Introduction

Let $X_{1-d}, \ldots, X_n$ be observations from a homogeneous and geometrically ergodic d-order Markov chain on some arbitrary state space. Write $\mathbf{X}_{i-1} = (X_{i-1}, \ldots, X_{i-d})$, and assume that the chain meets the conditional constraint

$$E(a_\vartheta(\mathbf{X}_{i-1}, X_i) \mid \mathbf{X}_{i-1}) = 0, \tag{1}$$

where $a_\vartheta(\mathbf{x}, y)$ with $\mathbf{x} = (x_1, \ldots, x_d)$ is a known k-dimensional vector of functions involving an unknown p-dimensional parameter ϑ. We are interested in optimal estimators of ϑ.

In Section 2 we derive an asymptotic lower bound for estimators of ϑ in the sense of Hájek and Le Cam, and give a characterization of efficient estimators.

In Section 3 we consider estimating equations for ϑ of the form

$$\sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) = 0,$$

with $W_\vartheta(\mathbf{x})$ a $k \times p$ matrix of weights. The weights minimizing the asymptotic covariance matrix depend, through conditional expectations of certain functions, on the unknown transition distribution of the chain. Hence the optimal estimating function cannot be used as it stands for estimating ϑ. We indicate that replacing the optimal weights by appropriate estimators does not change the asymptotic covariance matrix, and show that the resulting estimating function with estimated optimal weights is efficient. We also introduce generalized quasi-likelihood estimating functions, replacing the optimal weights by parametric models for the conditional expectations. These estimating functions are easier to calculate, but inefficient both for correctly specified and for misspecified conditional expectations.

We discuss these findings in more specific situations. A particular class of examples of constraints (1) are *quasi-likelihood models*, with real state space and parametric models for the conditional means and variances,

$$E(X_i \mid \mathbf{X}_{i-1}) = r_\vartheta(\mathbf{X}_{i-1}), \tag{2}$$

$$E((X_i - r_\vartheta(X_i))^2 \mid \mathbf{X}_{i-1}) = v_\vartheta(\mathbf{X}_{i-1}). \tag{3}$$

Then $a_\vartheta(\mathbf{x}, y) = (y - r_\vartheta(\mathbf{x}), (y - r_\vartheta(\mathbf{x}))^2 - v_\vartheta(\mathbf{x}))^\top$.

Quasi-likelihood models can be written as

$$X_i = r_\vartheta(\mathbf{X}_{i-1}) + v_\vartheta(\mathbf{X}_{i-1})^{1/2}\varepsilon_i, \tag{4}$$

with innovations ε_i that are martingale increments, $E(\varepsilon_i \mid \mathbf{X}_{i-1}) = 0$, and that satisfy $E(\varepsilon_i^2 \mid \mathbf{X}_{i-1}) = 1$ for identifiability. The submodel with *independent* innovations ε_i is called *nonlinear and heteroscedastic p-order autoregressive model.* We indicate that the estimating function with (estimated) optimal weights is not efficient in this submodel because it does not use the information that the innovations are independent.

2 Efficiency

In this section we derive a characterization of efficient estimators of ϑ in the d-order Markov chain model constrained by (1).

Consider first the *nonparametric* d-order Markov chain model, without constraint (1). Write $Q(\mathbf{x}, dy)$ for the transition distribution of X_i given $\mathbf{X}_{i-1} = \mathbf{x}$, and assume that the chain is geometrically ergodic under Q. Let $\pi(d\mathbf{x})$ be the stationary law of $\mathbf{X}_{i-1}$. Write $(\pi \otimes Q)(d\mathbf{x}, dy) = \pi(d\mathbf{x})Q(\mathbf{x}, dy)$ for the joint law of $(\mathbf{X}_{i-1}, X_i)$, and $Q(\mathbf{x}, f) = \int Q(\mathbf{x}, dy)f(\mathbf{x}, y)$ for the conditional expectation of $f(\mathbf{X}_{i-1}, X_i)$ given $\mathbf{X}_{i-1} = \mathbf{x}$. Whenever the argument $\mathbf{x}$ is omitted, we find it convenient to use the shorter notation Qf for $Q(\cdot, f)$.

The nonparametric model is *locally asymptotically normal* in the following sense. Introduce (Hellinger differentiable) perturbations

$$Q_{nh}(\mathbf{x}, dy) \doteq Q(\mathbf{x}, dy)(1 + n^{-1/2}h(\mathbf{x}, y)),$$

with h in the *tangent space*

$$H = \{h \in L_2(\pi \otimes Q) : Q(\mathbf{x}, h) = 0 \text{ for all } \mathbf{x}\}.$$

Since h may take large negative values, we cannot simply define Q_{nh} replacing $\doteq$ by an equality sign. There are three ways to take care of this problem: truncation of h, transformation of the density, or, simplest, restriction to bounded h (which are dense in H). The condition $Q(\mathbf{x}, h) = 0$ is required for Q_{nh} to be a transition distribution. Write P_{nh} and P_n for the joint law of $X_{1-p}, \ldots, X_n$ under Q_{nh} and Q, respectively. The log-likelihood ratio has the stochastic expansion

$$\log \frac{dP_{nh}}{dP_n} = n^{-1/2} \sum_{i=1}^{n} h(\mathbf{X}_{i-1}, X_i) - \frac{1}{2}(\pi \otimes Q)(h^2) + o_{P_n}(1).$$

For bounded h see Penev (1991). For general Hellinger differentiable perturbations, the stochastic expansion may be obtained by modifying Höpfner (1993a). See also Höpfner, Jacod and Ladelli (1990) and Höpfner (1993b). By a martingale central limit theorem, $n^{-1/2} \sum_{i=1}^{n} h(\mathbf{X}_{i-1}, X_i)$ is asymptotically normal with variance $(\pi \otimes Q)(h^2)$.

Now suppose that the model is constrained by (1). Relation (1) may be written $Q(\mathbf{x}, a_\vartheta) = 0$. The perturbed transition distribution Q_{nh} must also fulfill the constraint, possibly with perturbed parameter, say $\vartheta_{nu} \doteq \vartheta + n^{-1/2}u$:

$$0 = Q_{nh}(\mathbf{x}, a_{\vartheta_{nu}}) \doteq Q(\mathbf{x}, a_\vartheta) + n^{-1/2}(Q(\mathbf{x}, a_\vartheta h) + Q(\mathbf{x}, \dot{a}_\vartheta)u). \tag{5}$$

Hence the tangent space of the constrained model is the union, call it H_*, of the affine spaces

$$H_u = \{h \in H : Q(\mathbf{x}, a_\vartheta h) = -Q(\mathbf{x}, \dot{a}_\vartheta)u \text{ for all } \mathbf{x}\}.$$

We recall the following definitions and results from Le Cam's and Hájek's theory of efficient estimation. The standard reference for the i.i.d. case is Bickel, Klaassen, Ritov and Wellner (1998); for Markov chains see also Wefelmeyer (1999). A p-dimensional functional $t(Q)$ is called *differentiable* at Q with *gradient* g if $g \in H^p$ and

$$n^{1/2}(t(Q_{nh}) - t(Q)) \to (\pi \otimes Q)(gh) \quad \text{for } h \in H_*. \tag{6}$$

The *canonical* gradient g_* is the componentwise projection of g onto the tangent space H_*. An estimator $\hat{t}$ for $t(Q)$ is called *regular* at Q with *limit* L if

$$n^{1/2}(\hat{t} - t(Q_{nh})) \Rightarrow L \quad \text{under } P_{nh} \quad \text{for } h \in H_*.$$

The Convolution Theorem says that if $\hat{t}$ is regular for $t(Q)$ with limit L, then

$$L = (\pi \otimes Q)(g_* g_*^\top)^{1/2} N + M \quad \text{in distribution,}$$

where N a p-dimensional standard normal random vector, and M a random vector independent of N. This justifies calling a regular estimator *efficient* for $t(Q)$ if its limit is

$$L = (\pi \otimes Q)(g_* g_*^\top)^{1/2} N \quad \text{in distribution.}$$

An estimator $\hat{t}$ for $t(Q)$ is called *asymptotically linear* at P with *influence function* f if $f \in H^p$ and

$$n^{1/2}(\hat{t} - t(Q)) = n^{-1/2} \sum_{i=1}^{n} f(\mathbf{X}_{i-1}, X_i) + o_{P_n}(1). \tag{7}$$

Such an estimator is asymptotically normal with covariance matrix $(\pi \otimes Q)(ff^\top)$. We have the following two characterizations.

1. An asymptotically linear estimator for $t(Q)$ is regular if and only if its influence function is a gradient for $t(Q)$.

2. An estimator for $t(Q)$ is (regular and) efficient if and only if it is asymptotically linear with influence function equal to the canonical gradient of $t(Q)$.

Now we apply these results to estimation of ϑ. Consider the parameter ϑ as a functional of the transition distribution by setting $t(Q) = \vartheta$ if $Q(\mathbf{x}, a_\vartheta) = 0$. We have

$$n^{1/2}(t(Q_{nh}) - t(Q)) \doteq n^{1/2}(\vartheta_{nu} - \vartheta) \doteq u \quad \text{for } h \in H_u.$$

Hence, by (6), the canonical gradient is characterized as the vector $g_* \in H_*^p$ such that

$$(\pi \otimes Q)(g_* h) = u \quad \text{for } h \in H_u.$$

We show that the canonical gradient is $g_* = J^{-1}\ell$ with

$$\ell(\mathbf{x}, y) = -Q(\mathbf{x}, \dot{a}_\vartheta^\top) Q(\mathbf{x}, a_\vartheta a_\vartheta^\top)^{-1} a_\vartheta(\mathbf{x}, y),$$
$$J = (\pi \otimes Q)(\ell\ell^\top) = \pi(Q\dot{a}_\vartheta^\top Q(a_\vartheta a_\vartheta^\top)^{-1} Q\dot{a}_\vartheta).$$

We have

$$Q(\mathbf{x}, a_\vartheta \ell^\top) = -Q(\mathbf{x}, \dot{a}_\vartheta).$$

Hence the j-th component ℓ_j of ℓ is in H_{e_j}, where e_j denotes the j-th p-dimensional unit vector. It follows that ℓ and hence $J^{-1}\ell$ is in H_*^p. Furthermore, for $h \in H_u$,

$$(\pi \otimes Q)(J^{-1}\ell \cdot h) = -\pi(Q\dot{a}_\vartheta^\top Q(a_\vartheta a_\vartheta^\top)^{-1} Q\dot{a}_\vartheta)^{-1}\, \pi(Q\dot{a}_\vartheta^\top Q(a_\vartheta a_\vartheta^\top)^{-1} Q(a_\vartheta h))$$
$$= u.$$

This completes the proof that $J^{-1}\ell$ is the canonical gradient of ϑ. Using the above characterization of efficient estimators, we arrive at the following result.

Characterization. The canonical gradient of ϑ is $g_* = J^{-1}\ell$. Hence an estimator $\hat{\vartheta}$ for ϑ is regular and efficient if and only if

$$n^{1/2}(\hat{\vartheta} - \vartheta) = J^{-1} n^{-1/2} \sum_{i=1}^{n} \ell(\mathbf{X}_{i-1}, X_i) + o_{P_n}(1). \tag{8}$$

Its asymptotic covariance matrix is J^{-1}.

We see that ℓ and J play the roles of *score function* and *Fisher information* for ϑ.

The characterization sketched in this section is analogous to that obtained in Müller and Wefelmeyer (2001b) for the corresponding *regression* model, with i.i.d. observations (X_i, Y_i) meeting the conditional constraint $E(a_\vartheta(X_i, Y_i) \mid X_i) = 0$. A (different) derivation of the asymptotic variance bound J^{-1} is already sketched in Chamberlain (1987), with generalizations in (Chamberlain, 1992). Reviews are Newey (1990, 1993). Similar arguments as above are used in Müller and Wefelmeyer (2001a) for models with i.i.d. observations X_i satisfying an unconditional constraint $Ea_\vartheta(X_i) = 0$. Estimators of the *stationary law* π in our model (1) are constructed in Schick and Wefelmeyer (1999).

3 Estimating functions

The characterization (8) of efficient estimators for ϑ suggests to construct an efficient estimator as a one-step Newton–Raphson improvement of an initial,

inefficient, estimator $\bar{\vartheta}$,

$$\hat{\vartheta} = \bar{\vartheta} + \bar{J}^{-1}\frac{1}{n}\sum_{i=1}^{n}\bar{\ell}(\mathbf{X}_{i-1}, X_i),$$

with appropriate estimators $\bar{J}$ and $\bar{\ell}$ for J and ℓ. This construction does however not take advantage of the special feature of our model and is not recommended.

The conditional constraint (1) says that $a_\vartheta(\mathbf{X}_{i-1}, X_i)$ is a martingale increment. This suggests estimating ϑ by solutions $\hat{\vartheta}$ of martingale estimating equations

$$\sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) = 0, \tag{9}$$

with $W_\vartheta(\mathbf{x})$ a $k \times p$-matrix of weights. The asymptotic distribution of $\hat{\vartheta}$ is obtained from a Taylor expansion

$$0 \doteq \sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) + \sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top \dot{a}_\vartheta(\mathbf{X}_{i-1}, X_i)(\hat{\vartheta} - \vartheta),$$

with $\dot{a}_\vartheta(\mathbf{x}, y)$ the $k \times p$-matrix of partial derivatives of $a_\vartheta(\mathbf{x}, y)$ with respect to ϑ. Here we have used that $\dot{W}_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i)$ is a martingale increment and therefore negligible. If $(\pi \otimes Q)(W_\vartheta^\top \dot{a}_\vartheta)$ is invertible, we obtain the stochastic approximation

$$n^{1/2}(\hat{\vartheta} - \vartheta) = -\Big(\frac{1}{n}\sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top \dot{a}_\vartheta(\mathbf{X}_{i-1}, X_i)\Big)^{-1}$$
$$n^{-1/2}\sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) + o_{P_n}(1). \tag{10}$$

By ergodicity, we may replace the average in (10) by its mean $(\pi \otimes Q)(W_\vartheta^\top \dot{a}_\vartheta)$. Then $\hat{\vartheta}$ is seen to be asymptotically linear (7) with influence function

$$f(\mathbf{x}, y) = -(\pi \otimes Q)(W_\vartheta^\top \dot{a}_\vartheta)^{-1} W_\vartheta(\mathbf{x})^\top a_\vartheta(\mathbf{x}, y).$$

Hence $\hat{\vartheta}$ is asymptotically normal with covariance matrix

$$(\pi \otimes Q)(W_\vartheta^\top \dot{a}_\vartheta)^{-1}(\pi \otimes Q)(W_\vartheta^\top a_\vartheta a_\vartheta^\top W_\vartheta)(\pi \otimes Q)(\dot{a}_\vartheta W_\vartheta^\top)^{-1}$$
$$= \pi(W_\vartheta^\top Q\dot{a}_\vartheta)^{-1}\pi(W_\vartheta^\top Q(a_\vartheta a_\vartheta^\top)W_\vartheta)\pi(Q\dot{a}_\vartheta^\top W_\vartheta)^{-1}. \tag{11}$$

By the Cauchy–Schwarz inequality, the optimal weights are

$$W_\vartheta(\mathbf{x}) = W_\vartheta^*(\mathbf{x}) = Q(\mathbf{x}, a_\vartheta a_\vartheta^\top)^{-1} Q(\mathbf{x}, \dot{a}_\vartheta). \tag{12}$$

For these weights, the covariance matrix (11) is

$$\pi(Q\dot{a}_\vartheta^\top Q(a_\vartheta a_\vartheta^\top)^{-1} Q\dot{a}_\vartheta)^{-1}.$$

This is the asymptotic variance bound J^{-1} obtained in Section 2.

Minimizing the matrix (11) is also suggested by the *non-asymptotic optimality criterion* of Godambe (1985) and Godambe and Heyde (1987).

The average $\frac{1}{n}\sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top \dot{a}_\vartheta(\mathbf{X}_{i-1}, X_i)$ in (10) may also be replaced by $\frac{1}{n}\sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top Q(\mathbf{X}_{i-1}, \dot{a}_\vartheta)$. The *asymptotic optimality criterion* of Godambe and Heyde (1987) suggests minimizing the matrix

$$\begin{aligned}&\Big(\sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top Q(\mathbf{X}_{i-1}, \dot{a}_\vartheta)\Big)^{-1} \\ &\quad \sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top Q(\mathbf{X}_{i-1}, a_\vartheta a_\vartheta^\top) W_\vartheta(\mathbf{X}_{i-1}) \\ &\qquad \Big(\sum_{i=1}^n Q(\mathbf{X}_{i-1}, \dot{a}_\vartheta)^\top W_\vartheta(\mathbf{X}_{i-1})\Big)^{-1}.\end{aligned} \tag{13}$$

This leads to the same optimal weights. We refer to Heyde (1997) for uses of this criterion.

The optimal weights depend, through $Q(\mathbf{X}_{i-1}, a_\vartheta a_\vartheta^\top)$ and $Q(\mathbf{X}_{i-1}, \dot{a}_\vartheta)$, on the unknown transition distribution of the Markov chain. Hence the corresponding optimal estimating function cannot be used as it stands for estimating ϑ. We will call such an estimating function *undetermined*.

Generalized method of moments. The martingale estimating equation (9) results in an estimator that is asymptotically equivalent to the *GMM estimator* obtained from the *generalized method of moments*, the minimizer $\hat{\vartheta}$ of

$$\sum_{i=1}^n a_\vartheta(\mathbf{X}_{i-1}, X_i)^\top W_\vartheta(\mathbf{X}_{i-1}) M_n \sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i), \tag{14}$$

where M_n is a random symmetric $p \times p$ matrix converging to a deterministic matrix M, say. To prove the asymptotic equivalence, we write the GMM estimator as solution of an estimating equation. Taking partial derivatives with respect to ϑ, we see that $\hat{\vartheta}$ solves

$$\sum_{i=1}^n \dot{a}_{\hat{\vartheta}}(\mathbf{X}_{i-1}, X_i)^\top W_{\hat{\vartheta}}(\mathbf{X}_{i-1}) M_n \sum_{i=1}^n W_{\hat{\vartheta}}(\mathbf{X}_{i-1})^\top a_{\hat{\vartheta}}(\mathbf{X}_{i-1}, X_i)$$

$$+\sum_{i=1}^{n} a_{\hat{\vartheta}}(\mathbf{X}_{i-1}, X_i)^\top \dot{W}_{\hat{\vartheta}}(\mathbf{X}_{i-1}) M_n \sum_{i=1}^{n} W_{\hat{\vartheta}}(\mathbf{X}_{i-1})^\top a_{\hat{\vartheta}}(\mathbf{X}_{i-1}, X_i) = 0.$$

Again, the term involving $\dot{W}_{\hat{\vartheta}}$ is negligible because $\dot{W}_{\vartheta}(\mathbf{X}_{i-1})^\top a_{\vartheta}(\mathbf{X}_{i-1}, X_i)$ is a martingale increment. Using this argument repeatedly, we obtain by a Taylor expansion,

$$\begin{aligned} 0 \doteq & \sum_{i=1}^{n} \dot{a}_{\vartheta}(\mathbf{X}_{i-1}, X_i)^\top W_{\vartheta}(\mathbf{X}_{i-1}) M_n \sum_{i=1}^{n} W_{\vartheta}(\mathbf{X}_{i-1})^\top a_{\vartheta}(\mathbf{X}_{i-1}, X_i) \\ & + \sum_{i=1}^{n} \dot{a}_{\vartheta}(\mathbf{X}_{i-1}, X_i)^\top W_{\vartheta}(\mathbf{X}_{i-1}) M_n \sum_{i=1}^{n} W_{\vartheta}(\mathbf{X}_{i-1})^\top \dot{a}_{\vartheta}(\mathbf{X}_{i-1}, X_i)(\hat{\vartheta} - \vartheta). \end{aligned}$$

If M and $(\pi \otimes Q)(W_{\vartheta}^\top \dot{a}_{\vartheta})$ are invertible, we obtain

$$\begin{aligned} n^{1/2}(\hat{\vartheta} - \vartheta) = & -\Big((\pi \otimes Q)(\dot{a}_{\vartheta}^\top W_{\vartheta}) \cdot M \cdot (\pi \otimes Q)(W_{\vartheta}^\top \dot{a}_{\vartheta})\Big)^{-1} \\ & (\pi \otimes Q)(\dot{a}_{\vartheta}^\top W_{\vartheta}) \cdot M \cdot n^{-1/2} \sum_{i=1}^{n} W_{\vartheta}(\mathbf{X}_{i-1})^\top a_{\vartheta}(\mathbf{X}_{i-1}, X_i) \\ & + o_{P_n}(1) \\ = & -(\pi \otimes Q)(W_{\vartheta}^\top \dot{a}_{\vartheta})^{-1} n^{-1/2} \sum_{i=1}^{n} W_{\vartheta}(\mathbf{X}_{i-1})^\top a_{\vartheta}(\mathbf{X}_{i-1}, X_i) \\ & + o_{P_n}(1). \end{aligned}$$

Hence the GMM estimator has the same influence function as the estimator obtained from estimating equation (9). The optimal weights are therefore again given by (12). The generalized method of moments was developed by Hansen (1982, 1985). The optimal weights for this method were first obtained by Newey (1993). For reviews see Newey and McFadden (1994) and Wooldridge (1994). Note that the influence function of the GMM estimator does not involve the matrix M. Hence the random matrix M_n in (14) plays no role.

Generalized quasi-likelihood. One way of dealing with the problem of undetermined estimating functions is to specify parametric models for the conditional expectations involved in the optimal weights:

$$\Sigma_{\vartheta}(\mathbf{x}) = Q(\mathbf{x}, a_{\vartheta} a_{\vartheta}^\top) \quad \text{and} \quad A_{\vartheta}(\mathbf{x}) = Q(\mathbf{x}, \dot{a}_{\vartheta}).$$

This leads to the estimating equation

$$\sum_{i=1}^{n} A_{\vartheta}(\mathbf{X}_{i-1})^\top \Sigma_{\vartheta}(\mathbf{X}_{i-1})^{-1} a_{\vartheta}(\mathbf{X}_{i-1}, X_i) = 0. \tag{15}$$

We call the estimating function on the left (score function of the) *generalized quasi-likelihood.*

If Σ_ϑ and A_ϑ are correctly specified, this amounts to an additional restriction on the model. In this case, we can find new estimating functions besides (9) by using, in addition to $a_\vartheta(\mathbf{X}_{i-1}, X_i)$, further martingale increments

$$a_\vartheta(\mathbf{X}_{i-1}, X_i)a_\vartheta(\mathbf{X}_{i-1}, X_i)^\top - \Sigma_\vartheta(\mathbf{X}_{i-1}) \quad \text{and} \quad \dot{a}_\vartheta(\mathbf{X}_{i-1}, X_i) - A_\vartheta(\mathbf{X}_{i-1}).$$

Hence the generalized quasi-likelihood is inefficient, in general.

If Σ_ϑ and A_ϑ are misspecified, then the generalized quasi-likelihood still gives a consistent estimator, but is again inefficient, in general, now in model (1), since the weights will be different from the optimal ones.

We note that since $Q(\mathbf{x}, a_\vartheta a_\vartheta^\top)$ is $k \times k$ and symmetric, and $Q(\mathbf{x}, \dot{a}_\vartheta)$ is $k \times p$, the generalized quasi-likelihood requires modeling up to $\frac{1}{2}k(k+1) + kp$ functions in addition to the k components of a_ϑ.

We can summarize the above discussion in the following statement.

Dichotomy. The estimating equation (9) with optimal weights (12) is undetermined; the generalized quasi-likelihood (15) is inefficient.

Another, more satisfactory way of dealing with the problem of undetermined optimal weights is to replace them with estimators. It is not difficult to see that the stochastic approximation (10) remains valid if we replace the weights $W_\vartheta(\mathbf{X}_{i-1})$ by appropriate estimators $\hat{W}_\vartheta(\mathbf{X}_{i-1})$. The reason is that the weights are predictable. This argument is well known: $n^{-1/2}\sum_{i=1}^n(\hat{W}_\vartheta(\mathbf{X}_{i-1}) - W_\vartheta(\mathbf{X}_{i-1}))a_\vartheta(\mathbf{X}_{i-1}, X_i)$ is (approximately) conditionally centered with negligible terms, and therefore negligible. For heteroscedastic linear models $Y_{ij} = \vartheta^\top x_i + H(x_i)\varepsilon_{ij}$ and $Y_{ij} = \vartheta^\top x_i + H(\vartheta^\top x_i)\varepsilon_{ij}$ with unknown function H see Carroll (1982). For quasi-likelihood models (2) and (3) see Wefelmeyer (1996, 1997). For nonparametric regression models $Y_i = g(\vartheta^\top x_i) + v(g(\vartheta^\top x_i))^{1/2}\varepsilon_i$ with unknown function v and unknown or known function g see Chiou and Müller (1998, 1999). We arrive at the following result.

Estimated weights. If $\hat{W}_\vartheta^*(\mathbf{x})$ is an appropriate consistent estimator (possibly depending on ϑ) for

$$W_\vartheta^*(\mathbf{x}) = Q(\mathbf{x}, a_\vartheta a_\vartheta^\top)^{-1}Q(\mathbf{x}, \dot{a}_\vartheta),$$

then an efficient estimator for ϑ is obtained from the estimating equation with estimated optimal weights,

$$\sum_{i=1}^n \hat{W}_\vartheta^*(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) = 0.$$

Müller and Wefelmeyer (2001b) obtain an analogous result for the corresponding regression model, with i.i.d. observations (X_i, Y_i) satisfying $E(a_\vartheta(X_i, Y_i) \mid X_i) = 0$. Let us briefly sketch two specific methods of estimating the optimal weights $W^*_\vartheta(\mathbf{x})$.

Kernel estimators and penalized empirical variance. The optimal weights $W^*_\vartheta(\mathbf{x})$ involve conditional expectations. One way of estimating them is to use kernel estimators $\hat{\Sigma}_\vartheta(\mathbf{x})$ and $\hat{A}_\vartheta(\mathbf{x})$ for $Q(\mathbf{x}, a_\vartheta a_\vartheta^\top)$ and $Q(\mathbf{x}, \dot{a}_\vartheta)$. Such estimators require fairly large sample sizes. A different approach is developed by Li (2000, 20001), exploiting ideas of Lindsay (1985). Li considers i.i.d. observations (X_i, Y_i) with $E(Y_i \mid X_i) = \mu(\vartheta^\top X_i)$ and $E((Y_i - \mu(\vartheta^\top X_i))^2 \mid X_i) = \nu(\vartheta^\top X_i)$. For our constrained model (1), the approach consists in determining, for fixed ϑ, weights $\hat{W}^*_\vartheta(\mathbf{x})$ that minimize the appropriately penalized empirical version of the covariance matrix (11),

$$\Big(\frac{1}{n}\sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top \dot{a}_\vartheta(\mathbf{X}_{i-1}, X_i)\Big)^{-1}$$
$$\Big(\frac{1}{n}\sum_{i=1}^n W_\vartheta(\mathbf{X}_{i-1})^\top a_\vartheta(\mathbf{X}_{i-1}, X_i) a_\vartheta(\mathbf{X}_{i-1}, X_i)^\top W_\vartheta(\mathbf{X}_{i-1}) + \lambda I\Big)$$
$$\Big(\frac{1}{n}\sum_{i=1}^n \dot{a}_\vartheta(\mathbf{X}_{i-1}, X_i)^\top W_\vartheta(\mathbf{X}_{i-1})\Big)^{-1}.$$

In the following we illustrate the above remarks on optimal estimating functions with five examples.

Quasi-likelihood. Suppose the state space is real, and we have a parametric model for the conditional mean of the Markov chain,

$$E(X_i \mid \mathbf{X}_{i-1}) = r_\vartheta(\mathbf{X}_{i-1}). \tag{16}$$

This is a conditional constraint with $a_\vartheta(\mathbf{x}, y) = y - r_\vartheta(\mathbf{x})$.

A simple estimator for ϑ is the *conditional least squares estimator*, the minimizer $\hat{\vartheta}$ of

$$\sum_{i=1}^n (X_i - r_\vartheta(\mathbf{X}_{i-1}))^2.$$

See Klimko and Nelson (1978) and Tjøstheim (1986). Taking partial derivatives with respect to ϑ, we see that $\hat{\vartheta}$ solves

$$\sum_{i=1}^n \dot{r}_\vartheta(\mathbf{X}_{i-1})^\top (X_i - r_\vartheta(\mathbf{X}_{i-1})) = 0.$$

Here $\dot{r}_\vartheta(\mathbf{x})$ is the *row* vector of partial derivatives with respect to ϑ.

The martingale estimating equations (9) corresponding to model (16) are

$$\sum_{i=1}^{n} W_\vartheta(\mathbf{X}_{i-1})^\top (X_i - r_\vartheta(\mathbf{X}_{i-1})) = 0,$$

with W_ϑ a $p \times 1$ vector of weights. Here $Q(\mathbf{x}, \dot{a}_\vartheta) = -\dot{r}_\vartheta(\mathbf{x})$ does not involve the (unknown) transition distribution Q. The optimal weights (12) are

$$W_\vartheta^*(\mathbf{x}) = -\Big(\int Q(\mathbf{x}, dy)(y - r_\vartheta(\mathbf{x}))^2 \Big)^{-1} \dot{r}_\vartheta(\mathbf{x}).$$

An efficient estimator for ϑ is obtained from the estimating function

$$\sum_{i=1}^{n} \dot{r}_\vartheta(\mathbf{X}_{i-1})^\top \hat{v}_\vartheta(\mathbf{X}_{i-1})^{-1} (X_i - r_\vartheta(\mathbf{X}_{i-1})) = 0, \tag{17}$$

with $\hat{v}_\vartheta(\mathbf{x})$ an appropriate estimator of the conditional variance $\int Q(\mathbf{x}, dy)(y - r_\vartheta(\mathbf{x}))^2$; see Wefelmeyer (1997). The *quasi-likelihood estimator* replaces $\hat{v}_\vartheta(\mathbf{x})$ by a parametric model

$$v_\vartheta(\mathbf{x}) = \int Q(\mathbf{x}, dy)(y - r_\vartheta(\mathbf{x}))^2. \tag{18}$$

The discussion of estimating equation (15) has shown that the quasi-likelihood estimator does not use the information about ϑ in the additional specification (18).

Extended quasi-likelihood. Suppose the state space is real, and we have parametric models (16) and (18) for the conditional mean and variance of the Markov chain. Then $a_\vartheta(\mathbf{x}, y) = (y - r_\vartheta(\mathbf{x}), (y - r_\vartheta(\mathbf{x}))^2 - v_\vartheta(\mathbf{x}))^\top$. Hence

$$Q(\mathbf{x}, \dot{a}_\vartheta) = -\begin{pmatrix} \dot{r}_\vartheta(\mathbf{x}) \\ \dot{v}_\vartheta(\mathbf{x}) \end{pmatrix},$$

$$Q(\mathbf{x}, a_\vartheta a_\vartheta^\top) = \begin{pmatrix} v_\vartheta(\mathbf{x}) & \mu_3(\mathbf{x}) \\ \mu_3(\mathbf{x}) & \mu_4(\mathbf{x}) - v_\vartheta(\mathbf{x})^2 \end{pmatrix},$$

where $\mu_j(\mathbf{x}) = \int Q(\mathbf{x}, dy)(y - r_\vartheta(\mathbf{x}))^j$, $j = 3, 4$, are the third and fourth centered conditional moments of the chain. An efficient estimator for ϑ is obtained from the corresponding estimating equation with estimated optimal weights; see Wefelmeyer (1996). It requires estimators for $\mu_3(\mathbf{x})$ and $\mu_4(\mathbf{x})$. The *extended quasi-likelihood estimator* replaces these moments by parametric models; again it does not use the information about ϑ in the additional specifications. For the extended quasi-likelihood estimator in the case when

$\mu_3(\mathbf{x}) = 0$, see Crowder (1986, 1987), Godambe (1987), and Godambe and Thompson (1989); for the general case see Heyde (1987, 1997).

Nonlinear autoregression. A submodel of the Markov chain model with parametric specification (16) of the conditional mean is the *nonlinear d-order autoregressive model*

$$X_i = r_\vartheta(\mathbf{X}_{i-1}) + \varepsilon_i,$$

where the innovations are i.i.d. with density f having mean 0 and variance σ^2, say. Then $Q(\mathbf{x}, dy) = f(y - r_\vartheta(\mathbf{x}))dy$. The conditional variance $\int Q(\mathbf{x}, dy)(y - r_\vartheta(\mathbf{x}))^2$ reduces to σ^2, and the optimal estimating equation (17) simplifies to the equation defining the conditional least squares estimator,

$$\sum_{i=1}^{n} \dot{r}_\vartheta(\mathbf{X}_{i-1})^\top (X_i - r_\vartheta(\mathbf{X}_{i-1})) = 0.$$

This estimating equation does not require estimators for the weights. It is not efficient because it does not use the information that the innovations are i.i.d. Efficient estimators for ϑ are constructed in Hwang and Basawa (1993), Jeganathan (1995), Drost, Klaassen and Werker (1997), and Koul and Schick (1997).

Nonlinear and heteroscedastic autoregression. A submodel of the quasi-likelihood model (16) and (18) is the *nonlinear and heteroscedastic d-order autoregressive model*

$$X_i = r_\vartheta(\mathbf{X}_{i-1}) + v_\vartheta(\mathbf{X}_{i-1})^{1/2}\varepsilon_i,$$

where the innovations are i.i.d. with density f having mean 0 and variance 1. Then

$$Q(\mathbf{x}, dy) = \frac{1}{v_\vartheta(\mathbf{x})^{1/2}} f\Big(\frac{y - r_\vartheta(\mathbf{x})}{v_\vartheta(\mathbf{x})^{1/2}}\Big) dy,$$

$$Q(\mathbf{x}, a_\vartheta a_\vartheta^\top) = \begin{pmatrix} v_\vartheta(\mathbf{x}) & v_\vartheta(\mathbf{x})^{3/2}\mu_3 \\ v_\vartheta(\mathbf{x})^{3/2}\mu_3 & v_\vartheta(\mathbf{x})^2(\mu_4 - 1) \end{pmatrix},$$

where μ_3 and μ_4 are the third and fourth (centered) moments of the innovation distribution. The optimal weights are therefore easy to estimate: simply replace μ_j by the empirical estimator

$$\hat{\mu}_{j\vartheta} = \frac{1}{n}\sum_{i=1}^{n} (X_i - r_\vartheta(\mathbf{X}_{i-1}))^j, \quad j = 3, 4.$$

Then the estimating equation with estimated optimal weights is

$$\sum_{i=1}^{n}(\dot{r}_\vartheta(\mathbf{X}_{i-1})^\top, \dot{v}_\vartheta(\mathbf{X}_{i-1})^\top)\begin{pmatrix} v_\vartheta(\mathbf{X}_{i-1}) & v_\vartheta(\mathbf{X}_{i-1})^{3/2}\hat{\mu}_{3\vartheta} \\ v_\vartheta(\mathbf{X}_{i-1})^{3/2}\hat{\mu}_{3\vartheta} & v_\vartheta(\mathbf{X}_{i-1})^2(\hat{\mu}_{4\vartheta}-1)\end{pmatrix}^{-1}$$
$$\begin{pmatrix} X_i - r_\vartheta(\mathbf{X}_{i-1}) \\ (X_i - r_\vartheta(\mathbf{X}_{i-1}))^2 - v_\vartheta(\mathbf{X}_{i-1})\end{pmatrix} = 0. \tag{19}$$

Again this estimator is not efficient. See Drost, Klaassen and Werker (1997) for efficient estimators of ϑ.

ARCH. A special case of the heteroscedastic d-order autoregressive model is the *ARCH(d) model*

$$X_i = v_\vartheta(\mathbf{X}_{i-1})^{1/2}\varepsilon_i \quad \text{with} \quad v_\vartheta(\mathbf{x}) = \vartheta_0 + \sum_{j=1}^{d}\vartheta_j x_j^2,$$

with $(d+1)$-dimensional parameter $\vartheta = (\vartheta_0, \ldots, \vartheta_d)$. The innovations are again assumed i.i.d. with mean 0 and variance 1. It is convenient to introduce $\mathbf{Y}_{i-1} = (1, X_{i-1}^2, \ldots, X_{i-d}^2)$. Then $v_\vartheta(\mathbf{X}_{i-1}) = \vartheta^\top\mathbf{Y}_{i-1}$. The optimal estimating equation (19) reduces to

$$\sum_{i=1}^{n}(\vartheta^\top\mathbf{Y}_{i-1})^{-2}\mathbf{Y}_{i-1}(X_i^2 - \vartheta^\top\mathbf{Y}_{i-1}) = 0.$$

Since the weights $(\vartheta^\top\mathbf{Y}_{i-1})^{-2}$ depend on ϑ, we cannot solve the equation explicitly. However, as seen above, we may replace the weights by estimators without changing the influence function of the solution of the estimating equation. A simple estimator for ϑ is the conditional least squares estimator

$$\overline{\vartheta} = \Big(\sum_{i=1}^{n}\mathbf{Y}_{i-1}\mathbf{Y}_{i-1}^\top\Big)^{-1}\sum_{i=1}^{n}X_{i-1}^2\mathbf{Y}_{i-1}.$$

The solution of the estimating equation with estimated optimal weights is

$$\hat{\vartheta} = \Big(\sum_{i=1}^{n}(\overline{\vartheta}^\top\mathbf{Y}_{i-1})^{-2}\mathbf{Y}_{i-1}\mathbf{Y}_{i-1}^\top\Big)^{-1}\sum_{i=1}^{n}(\overline{\vartheta}^\top\mathbf{Y}_{i-1})^{-2}X_{i-1}^2\mathbf{Y}_{i-1}.$$

For a direct derivation see Chandra and Taniguchi (2001). The estimator is not efficient. For efficient estimators see Engle and Gonzáles-Rivera (1991), Linton (1993), and Drost and Klaassen (1997).

Acknowledgment. We are most grateful to the referee for several useful suggestions that have made the paper more readable.

REFERENCES

Bickel, P.J., Klaassen, C.A.J., Ritov, Y. and Wellner, J.A. (1998). *Efficient and Adaptive Estimation for Semiparametric Models.* New York: Springer.

Carroll, R.J. (1982). Adapting for heteroscedasticity in linear models. *Ann. Statist.* 10, 1224-1233.

Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *J. Econometrics* 34, 305-334.

Chamberlain, G. (1992). Efficiency bounds for semiparametric regression. *Econometrica* 60, 567-596.

Chandra, S.A. and Taniguchi, M. (2001). Estimating functions for nonlinear time series models. *Ann. Inst. Statist. Math.* 53, 125-141.

Chiou, J.M. and Müller, H.G. (1998). Quasi-likelihood regression with unknown link and variance functions. *J. Amer. Statist. Assoc.* 93, 1376-1387.

Chiou, J.M. and Müller, H.G. (1999). Nonparametric quasi-likelihood. *Ann. Statist.* 27, 36-64.

Crowder, M. (1986). On consistency and inconsistency of estimating equations. *Econometric Theory* 2, 305-330.

Crowder, M. (1987). On linear and quadratic estimating functions. *Biometrika* 74, 591-597.

Drost, F.C. and Klaassen, C.A. (1997). Efficient estimation in semiparametric GARCH models. *J. Econometrics* 81, 193-221.

Drost, F.C., Klaassen, C.A. and Werker, B.J.M. (1997). Adaptive estimation in time-series models. *Ann. Statist.* 25, 786-817.

Engle, R.F. and González-Rivera, G. (1991). Semiparametric ARCH models. *J. Business Economic Statist.* 9, 345-359.

Godambe, V.P. (1985). The foundations of finite sample estimation in stochastic processes. *Biometrika* 72, 419-428.

Godambe, V.P. (1987). The foundations of finite sample estimation in stochastic processes II. In *Proceedings of the 1st World Congress of the Bernoulli Society*, Prohorov, Y. and Sazonov, V. V. (editors), 49-54. Utrecht: VNU Science Press.

Godambe, V.P. and Heyde, C.C. (1987). Quasi-likelihood and optimal estimation. *Internat. Statist. Rev.* 55, 231-244.

Godambe, V.P. and Thompson, M.E. (1989). An extension of quasi-likelihood estimation. *J. Statist. Plann. Inference* 22, 137-152.

Hansen, L.P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029-1054.

Hansen, L.P. (1985). A method for calculating bounds on the asymptotic covariance matrices of generalized method of moments estimators. *J. Econometrics* 30, 203-238.

Heyde, C.C. (1987). On combining quasi-likelihood estimating functions. *Stochastic Process. Appl.* 25, 281-287.

Heyde, C.C. (1997). *Quasi-Likelihood And Its Application, A General Approach to Optimal Parameter Estimation.* New York: Springer.

Höpfner, R. (1993a). On statistics of Markov step processes: Representation of log-likelihood ratio processes in filtered local models. *Probab. Theory Related Fields* 94, 375-398

Höpfner, R. (1993b). Asymptotic inference for Markov step processes: Observation up to a random time. *Stochastic Process. Appl.* 48, 295-310

Höpfner, R., Jacod, J. and Ladelli, L. (1990). Local asymptotic normality and mixed normality for Markov statistical models. *Probab. Theory Related Fields* 86, 105-129

Hwang, S.Y. and Basawa, I.V. (1993). Asymptotic optimal inference for a class of nonlinear time series models. *Stochastic Process. Appl.* 46, 91-113.

Jeganathan, P. (1995). Some aspects of asymptotic theory with applications to time series models. *Econometric Theory* 11, 818-887.

Klimko, L.A. and Nelson, P.I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.* 6, 629-642.

Koul, H.L. and Schick, A. (1997). Efficient estimation in nonlinear autoregressive time-series models. *Bernoulli* 3, 247-277.

Li, B. (2000). Nonparametric estimating equations based on a penalized information criterion. *Canad. J. Statist.* 28, 621-639.

Li, B. (2001). On quasilikelihood equations with nonparametric weights. *Scand. J. Statist.* 28, 577-602.

Lindsay, B.G. (1985). Using empirical partially Bayes inference for increased efficiency. *Ann. Statist.* 13, 914-931.

Linton, O. (1993). Adaptive estimation in ARCH models. *Econometric Theory* 9, 539-569.

Müller, U.U. and Wefelmeyer, W. (2001a). Estimators for models with constraints involving unknown parameters, to appear in: *Math. Methods Statist.* http://www.math.uni-siegen.de/statistik/wefelmeyer.html.

Müller, U.U. and Wefelmeyer, W. (2001b). Regression type models and optimal estimators, in preparation.

Newey, W.K. (1990). Semiparametric efficiency bounds. J. Appl. Econometrics 5, 99-135.

Newey, W.K. (1993). Efficient estimation of models with conditional moment

restrictions. In *Handbook of Statistics 11: Econometrics*, Maddala, G. S., Rao, C. R. and Vinod, H. D. (editors), 419-454. Amsterdam: Elsevier.

Newey, W.K. and McFadden, D.L. (1994). Large sample estimation and hypothesis testing. In *Handbook of Econometrics* 4, Engle, R. F. and McFadden, D. L. (editors), 2111–2245. Amsterdam: Elsevier.

Penev, S. (1991). Efficient estimation of the stationary distribution for exponentially ergodic Markov chains. *J. Statist. Plann. Inference* 27, 105-123.

Schick, A. and Wefelmeyer, W. (1999). Efficient estimation of invariant distributions of some semiparametric Markov chain models. *Math. Methods Statist.* 8, 119-134.

Tjøstheim, D. (1986). Estimation in nonlinear time series models. *Stochastic Process. Appl.* 21, 251–273.

Wefelmeyer, W. (1996). Quasi-likelihood models and optimal inference. Ann. Statist. 24, 405-422.

Wefelmeyer, W. (1997). Adaptive estimators for parameters of the autoregression function of a Markov chain. *J. Statist. Plann. Inference* 58, 389-398.

Wefelmeyer, W. (1999). Efficient estimation in Markov chain models: an introduction. In *Asymptotics, Nonparametrics, and Time Series*, Ghosh, S. (editor), 427-459. Statistics: Textbooks and Monographs 158. New York: Dekker.

Wooldridge, J.M. (1994). Estimation and inference for dependent processes. In *Handbook of Econometrics* 4, Engle, R.F. and McFadden, D.L. (editors), 2639-2738. Amsterdam: Elsevier.

SEMIPARAMETRIC DENSITY ESTIMATION WITH ADDITIVE ADJUSTMENT

KANTA NAITO
Department of Mathematics
Shimane University
Matsue 690-8504, Japan
E-mail: naito@math.shimane-u.ac.jp

This paper introduces a class of density estimators having both of parametric part and nonparametric factor. A plug-in parametric estimator is seen as an initial guess of the true density, and the proposed estimator is built up by adding a nonparametric adjustment factor and the initial estimator. Asymptotic theory is developed, and comparisons with the traditional kernel estimator and a multiplicative estimator are also reported.

Keywords: adjustment, density estimation, kernel, local fitting.

1 Introduction

Let $X_1, ..., X_n$ be independently and identically distributed with density f. There are two main approaches for estimating f from the data. One is called parametric approach. Under the assumption that the underlying f is exactly in a prepared parametric model $\{g(x,\theta) : \theta \in \Theta\}$ ruled by a finite dimensional parameter θ in $\Theta \subset R^p$, or at least in the neighborhood of it, estimating the density function f is replaced by estimation of θ in the model. And we finally obtain a parametric density estimator $\hat{f}(x) = g(x,\hat{\theta})$, by plug-in, where $\hat{\theta}$ is an estimator of θ. Another approach is called nonparametric. The kernel method is well known and representative nonparametric approach which can be utilized without the structural assumption such that the underlying structure is controlled or captured by finite dimensional parameter (cf. Wand and Jones, 1995). Thus nonparametric methods have attractive flexibility, however it is hard to throw away the parametric model since well-estimated structure by a parametric approach provides a clear understanding of the underlying structure.

In the present approach, a plug-in parametric density estimator $g(x,\hat{\theta})$ is utilized, but it is seen as a crude guess of f. And we aim to adjust this initial parametric approximation by $g(x,\hat{\theta}) + \xi$, where ξ is called the adjustment factor. The adjustment factor ξ is determined by the minimization of the

empirical version of a function

$$Q(x,\xi|\alpha) = \int K_h(t-x)\frac{[f(t)-\{g(t,\hat{\theta})+\xi\}]^2}{g(t,\hat{\theta})^\alpha}dt,$$

with respect to ξ, where α is a real number called index. This method is called *the local L_2-fitting criteria with index α*. Here, $K(z)$ is a symmetric density and $K_h(z) = K(z/h)/h$. The symmetric density K makes the adjustment locally around the target point x. This local approach is based on a simple intuition that an observed data which is far from the target point x does not have information about the adjustment.

In the next section, a class of density estimators is constructed and behavior of a typical estimator in the class is investigated. It will be seen that the elegant feature of the bias function enables us to find the best estimator in the class. Theoretical AMISE comparisons of new estimators with kernel estimator $\hat{f}_K(x) = n^{-1}\sum_{i=1}^n K_h(X_i - x)$ and multiplicative estimator $\hat{f}_{HG}(x) = n^{-1}\sum_{i=1}^n g(x,\hat{\theta})\{K_h(X_i-x)/g(X_i,\hat{\theta})\}$ proposed by Hjort and Glad (1995) are implemented for normal mixture densities in Section 3. Monte Carlo simulation study is reported in Section 4. In Section 5, some remarks and discussions are given in conjunction with our future works. Outline proof of a theoretical result is put on Section 6.

2 Estimators and their behavior

For a fixed target point x, the adjustment factor is obtained by minimization of the empirical version of $Q(x,\xi|\alpha)$ (omitting an irrelevant term),

$$Q_n(x,\xi|\alpha) = \xi^2\int\frac{K_h(t-x)}{g(t,\hat{\theta})^\alpha}dt - 2\xi\int\frac{K_h(t-x)}{g(t,\hat{\theta})^\alpha}\Big(dF_n(t) - g(t,\hat{\theta})dt\Big)$$

with respect to ξ, where F_n is the empirical distribution function. The minimizer is

$$\hat{\xi}_\alpha = \hat{\xi}_\alpha(x) = \frac{\int K_h(t-x)g(t,\hat{\theta})^{-\alpha}\{dF_n(t) - g(t,\hat{\theta})dt\}}{\int K_h(t-x)g(t,\hat{\theta})^{-\alpha}dt},$$

which gives the proposed estimator

$$\hat{f}_\alpha(x) = g(x,\hat{\theta}) + \frac{\int K_h(t-x)g(t,\hat{\theta})^{-\alpha}\{dF_n(t) - g(t,\hat{\theta})dt\}}{\int K_h(t-x)g(t,\hat{\theta})^{-\alpha}dt}. \qquad (1)$$

Different values of α produce different estimators, so that $\hat{f}_\alpha$ forms a class of estimators. When we deal with densities, the multiplicative adjustment

seems to be natural and in fact, $\hat{f}_{HG}$ is a typical example. But, there is a simple statistical result which provides a viewpoint of the additive adjustment. Under some regularity conditions, a statistic T_n, the standardized version of the sum of i.i.d data, has an asymptotic expansion of the density:

$$(d/dx)\Pr\{T_n \leq x\} = \phi(x)\left[1 + n^{-1/2}p(x)\right] + o(n^{-1/2}) \tag{2}$$

$$= \phi(x) + n^{-1/2}p(x)\phi(x) + o(n^{-1/2}), \tag{3}$$

for sufficient large n, where $p(x)$ is some polynomial and $\phi(x)$ is the standard normal density (see, Bhattacharya and Rao, 1976). Both (2) and (3) express adjustments of the asymptotic density $\phi(x)$: (2) is multiplicative, and (3) is additive. Further, as an earlier work, Olkin and Spiegelman (1987) discussed an estimator constructed by a convex combination of a plug-in parametric estimator and $\hat{f}_K$, which also can be seen as an additive adjustment. This paper, however, does not focus on their estimator, since it seems to be somewhat less attractive than $\hat{f}_{HG}$.

In the following, we investigate behavior of $\hat{f}_\alpha$, which is depending on that of $\hat{\theta}$ included in the initial estimator $g(x, \hat{\theta})$. For estimating θ, we utilize a functional estimator $\hat{\theta} = H(F_n)$ for a certain smooth functional H satisfying some regularity conditions, and having the influence function I (see, Shao, 1991). Using this functional expression, the least false value is defined as $\theta_0 = H(F)$ which yields the best parametric approximation $g_0(x) = g(x, \theta_0)$ to f among $g(x, \theta)$, $\theta \in \Theta$, where F is the cumulative of f. Notations $\mu_\ell = \int z^\ell K(z)dz$ and $R(K) = \int K(z)^2 dz$ are used throughout. Further we need the following assumptions: the densities f and g_0 are such that their first and second derivatives are continuous, square integrable. Then we have the following.

Proposition 1 *As* $n \to \infty$, $h \to 0$,

$$Bias\hat{f}_\alpha(x) = \frac{h^2}{2}\mu_2\left[g_0(x)^\alpha\left\{\left(\frac{f(x)-g_0(x)}{g_0(x)^\alpha}\right)'' - \{f(x)-g_0(x)\}(g_0(x)^{-\alpha})''\right\}\right]$$
$$+O\left(h^4 + \frac{h^2}{n} + \frac{1}{n^2}\right),$$
$$Var\hat{f}_\alpha(x) = \frac{R(K)}{nh}f(x) - \frac{f(x)^2}{n} + O\left(\frac{h}{n} + \frac{1}{n^2}\right).$$

Outline proof of Proposition 1 is included in Section 6. The essential difference of asymptotic behavior of $\hat{f}_K$, $\hat{f}_{HG}$ and $\hat{f}_\alpha$ appears in the bias, since the leading terms of their variances are the same. We see h^2-term of the bias is

vanished if f is in the model, that is, $f(x) = g_0(x)$. Further, an important point is that the bias can be rewritten as

$$g_0(x)^\alpha \left\{ \left(\frac{f(x) - g_0(x)}{g_0(x)^\alpha} \right)'' - \{f(x) - g_0(x)\}(g_0(x)^{-\alpha})'' \right\} = b_1(x) - \alpha b_2(x),$$

where

$$b_1(x) = f''(x) - g_0''(x) \quad \text{and} \quad b_2(x) = 2\{f'(x) - g_0'(x)\} \left\{ \frac{g_0'(x)}{g_0(x)} \right\}.$$

So the bias is linear in α and this fact yields the following theorem.

Theorem 1 *As $n \to \infty$, $h \to 0$, the leading term of* MISE *of $\hat{f}_\alpha$ is*

$$\text{AMISE}\left[\hat{f}_\alpha\right] = \frac{h^4}{4}\mu_2^2 \{C_1\alpha^2 - 2C_2\alpha + C_3\} + \frac{R(K)}{nh},$$

which is minimized on α at

$$\alpha_{opt} = \frac{C_2}{C_1}, \tag{4}$$

where

$$C_1 = \int b_2(x)^2 dx\ , \quad C_2 = \int b_2(x) b_1(x) dx \quad \textit{and} \quad C_3 = \int b_1(x)^2 dx.$$

3 AMISE comparison in normal mixture

In this section, we compare $\hat{f}_\alpha$ with $\hat{f}_K$ and $\hat{f}_{HG}$ by AMISE. The comparison by AMISE is acomplished by comparing $\mathcal{R}(\hat{f}_\alpha) = C_1\alpha^2 - 2C_2\alpha + C_3$ with the integrated squared biases of $\hat{f}_K$ and $\hat{f}_{HG}$ respectively given as $\mathcal{R}(\hat{f}_K) = \int\{f''(x)\}^2 dx$ and $\mathcal{R}(\hat{f}_{HG}) = \int\{g_0(x)(f(x)/g_0(x))''\}^2 dx$, since these variances are the same as mentioned before, where $C_i (i = 1, 2, 3)$ are those in Theorem 1. Here we adopt the normal mixture $f(x) = \sum_{i=1}^k p_i \phi_{\sigma_i}(x - \mu_i)$ as the true density, where $p_i \geq 0 (i = 1, 2, ..., k)$, $\sum_{i=1}^k p_i = 1$ and $\phi_\sigma(x-\mu)$ designates the normal (Gaussian) density with mean μ and variance σ^2. For $\hat{f}_\alpha$ and $\hat{f}_{HG}$, we use a normal density as an initial parametric model and MLE for estimation of θ_0, which is corresponding to the use of $\phi_{\sigma_0}(x - \mu_0)$ as $g_0(x)$, where μ_0 and σ_0^2 are given as $\mu_0 = \sum_{i=1}^k p_i \mu_i$ and $\sigma_0^2 = \sum_{i=1}^k p_i\{\sigma_i^2 + (\mu_i - \mu_0)^2\}$, respectively. Direct calculations mimicking those in Hjort and Glad (1995) give concrete expressions of $C_i (i = 1, 2, 3)$ for this normal mixture and normal start case. Detailed calculations are found in Naito (1998). The comparison is done for

Table 1. Values of $\sigma_0\mathcal{R}(\hat{f}_K)^{1/5}$, $\sigma_0\mathcal{R}(\hat{f}_{HG})^{1/5}$ and $\sigma_0\mathcal{R}(\hat{f}_\alpha)^{1/5}$ on $\alpha = 0, 1, 2, 3, \alpha_{opt}$ for normal mixture densities (#1 $\sim$ #15) utilized in Marron and Wand. Values of α_{opt} obtained by (4) are also tabulated in the last column.

f	$\hat{f}_K$	$\hat{f}_{HG}$	$\hat{f}_\alpha$					α_{opt}
			$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = \alpha_{opt}$	
#1	0.7330	0	0	0	0	0	0	——
#2	0.8921	0.6739	0.7519	0.7824	0.9102	1.0431	0.7433	0.3099
#3	5.6070	5.5985	5.6069	5.6110	5.6368	5.6833	5.6058	0.3135
#4	3.8664	3.8354	3.8657	3.8513	3.8383	3.8269	3.7935	9.1711
#5	2.3201	2.2088	2.3095	2.2748	2.2756	2.3119	2.2706	1.4766
#6	1.1183	1.0615	1.1403	1.1533	1.2631	1.4038	1.1304	0.3939
#7	2.0215	1.9579	2.0301	2.0595	2.2199	2.4318	2.0214	0.3198
#8	1.3753	1.3468	1.3787	1.4102	1.5325	1.6857	1.3751	0.2395
#9	1.5600	1.5335	1.5681	1.5813	1.6616	1.7789	1.5639	0.3299
#10	3.5571	3.5421	3.5570	3.5608	3.5937	3.6525	3.5550	0.3719
#11	12.4450	12.4447	12.4450	12.4451	12.4458	12.4472	12.4449	0.3591
#12	6.4350	6.4382	6.4350	6.4457	6.4800	6.5365	6.4350	0.0588
#13	11.1149	11.1147	11.1149	11.1151	11.1162	11.1181	11.1149	0.2833
#14	14.8802	14.8807	14.8802	14.8831	14.8930	14.9099	14.8802	0.0896
#15	9.6259	9.6261	9.6259	9.6333	9.6602	9.7058	9.6257	0.1186

representative normal mixture densities utilized in Marron and Wand (1992). See Hjort and Glad (1995) for the expressions of $\mathcal{R}(\hat{f}_K)$ and $\mathcal{R}(\hat{f}_{HG})$.

Table 1 tabulates the values for comparison. #1 is the case where f is exactly in the parametric model utilized, so that $\hat{f}_\alpha$ with any values of α are superior to $\hat{f}_K$. $\hat{f}_{\alpha_{opt}}$ surpasses $\hat{f}_K$ in all cases except #6 and #9. $\hat{f}_0$ is also superior to or at least competitive to $\hat{f}_K$. Multiplicative $\hat{f}_{HG}$ is good for almost all cases, but $\hat{f}_0$ is marginally better than $\hat{f}_{HG}$ in #12, #14 and #15. It can be concluded from these observations that $\hat{f}_\alpha$ is indeed an efficient class of density estimators.

4 Finite sample performances

The results of simulations are reported in this section. Normal mixture densities listed in Table 1 were again adopted here as true densities. In each case 1000 samples of size $n = 500$ were generated. The estimators compared in this study are those in Table 1. For a given bandwidth h, each density estimate was calculated on a grid of 301 points on $[-3, 3]$, and the numerical integral was done to obtain an approximation of ISE(h). The MISE(h) value is estimated by the average of these 1000 realizations of ISE(h). In order to obtain precise approximation to the minimum MISE, a grid search of the bandwidth was implemented. This is done after an initial screening has provided a suitable h interval containing the minimum. The Gaussian kernel was

Table 2. 10^5 times the values of estimated $\min_h$MISE based on 1000 simulations with sample size $n = 500$ for normal mixture densities (#1 $\sim$ #15).

f	$\hat{f}_K$	$\hat{f}_{HG}$	$\hat{f}_\alpha$				
			$\alpha = 0$	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = \alpha_{opt}$
#1	170	68	68	56	51	51	—
#2	266	207	216	219	256	294	210
#3	1401	1401	1401	1402	1411	1429	1400
#4	1312	1302	1312	1301	1291	1282	1248
#5	1700	1611	1691	1645	1643	1677	1638
#6	237	232	247	244	266	296	241
#7	341	332	344	347	377	416	342
#8	313	315	318	323	357	399	315
#9	301	301	307	309	332	363	304
#10	1163	1163	1163	1163	1177	1194	1162
#11	412	407	421	418	441	471	416
#12	969	975	970	973	996	1035	969
#13	669	670	672	668	676	691	669
#14	1523	1528	1523	1528	1540	1558	1524
#15	1441	1443	1441	1443	1450	1463	1441

used throughout. The results are tabulated in Table 2.

We can observe from Table 2 that $\hat{f}_{\alpha_{opt}}$ is the best in $\hat{f}_\alpha$ for almost all cases, which justifies Theorem 1. $\hat{f}_2$ and $\hat{f}_3$ are very good in #1. In ten densities from #2 to #11, $\hat{f}_{HG}$ is the best for six cases. But $\hat{f}_{\alpha_{opt}}$ is remarkably better than $\hat{f}_{HG}$ in #4. Among four cases from #12 to #15, both of $\hat{f}_0$ and $\hat{f}_1$ are marginally better than $\hat{f}_{HG}$ in #12, #14 and #15. A suggestion from this simulation is that $\hat{f}_\alpha$'s are good not only in the case where the initial parametric model is correct (#1), but also in the case where it is remarkably different in shape (#11 $\sim$ #15).

5 Discussion

It is easy to prove by its derivation that $\hat{\xi}_\alpha(x)$ is a consistent estimator of $f(x) - g_0(x)$ irrespective of the value of α, and hence $\hat{f}_\alpha$ also has consistency even when we utilize an *incorrect* model $g(x, \theta)$. It is noted that $\hat{f}_\alpha$ provides not only a density estimator but also a statistic $\hat{\xi}_\alpha$ which can be used to check the parametric model. For example ($\alpha = 0$), the plot of $\hat{\xi}_0(x)^2 = \{\hat{f}_K(x) - \int K_h(t - x) g(t, \hat{\theta}) dt\}^2$ against x is a graphical tool to assess the goodness of fit. Weak convergence of $\hat{\xi}_0(x)$ as a stochastic process under the hypothesis $H : f(x) \in \{g(x, \theta), \theta \in \Theta\}$ can be easily proved, from which we

can construct goodness of fit test. We defer such a theory to later work.

Data-based choice of the index α is practically important like the bandwidth selection problem (see, e.g., Wand and Jones, 1995); it is, however, also deferred to future work. The choices $\alpha = 0$ or $\alpha = 1$ are recommended for practical use, see the last column of Table 1.

6 Proof outline

Behavior of $\hat{f}_\alpha$ will be shown to be dominated by that of f_α^* in (5), and hence we need the following which can be obtained by Taylor expansion.

Lemma 1 *Let*

$$f_\alpha^*(x) = g_0(x) + \frac{\int K_h(t-x) g_0(t)^{-\alpha} \{dF_n(t) - g_0(t)dt\}}{\int K_h(t-x) g_0(t)^{-\alpha} dt}. \tag{5}$$

Then as $n \to \infty$, $h \to 0$,

$$Bias \hat{f}_\alpha^*(x) = \frac{h^2}{2}\mu_2 \left[g_0(x)^\alpha \left\{ \left(\frac{f(x) - g_0(x)}{g_0(x)^\alpha} \right)'' - \{f(x) - g_0(x)\}(g_0(x)^{-\alpha})'' \right\} \right] + O\left(h^4\right),$$

$$Var f_\alpha^*(x) = \frac{R(K)}{nh} f(x) - \frac{f(x)^2}{n} + O\left(\frac{h}{n}\right).$$

Outline Proof of PROPOSITION 1. Let us define

$$\zeta_0(x|\alpha) = \int K_h(t-x) g_0(t)^{-\alpha} dt, \quad \zeta_1(x|\alpha) = \int K_h(t-x) u_0(t) g_0(t)^{-\alpha} dt,$$

$$\zeta_2(x|\alpha) = \int K_h(t-x) \left\{ U_0(t) - \alpha u_0(t) u_0(t)' \right\} g_0(t)^{-\alpha} dt,$$

where $u_0(x) = (\partial/\partial\theta) \log g(x, \theta_0)$ and $U_0(x) = (\partial^2/\partial\theta\partial\theta') \log g(x, \theta_0)$. Taylor expansion of $\hat{\theta}$ around θ_0 gives that

$$\hat{f}_\alpha(x) = f_\alpha^*(x) + (\hat{\theta} - \theta_0)' \bar{\mathcal{B}}_n(x) + \frac{1}{2}(\hat{\theta} - \theta_0)' \bar{\mathcal{C}}_n(x)(\hat{\theta} - \theta_0) + o_p(n^{-1}),$$

where $f_\alpha^*(x)$ is in (5), $\bar{\mathcal{B}}_n(x) = (1/n)\sum_{i=1}^n \mathcal{B}_i(x)$, $\bar{\mathcal{C}}_n(x) = (1/n)\sum_{i=1}^n \mathcal{C}_i(x)$,

$$\mathcal{B}_i(x) = K_h(X_i - x) g_0(X_i)^{-\alpha} \left\{ \frac{\alpha}{\zeta_0(x|\alpha)^2} \zeta_1(x|\alpha) - \frac{\alpha}{\zeta_0(x|\alpha)} u_0(X_i) \right\}$$
$$- \frac{\alpha\zeta_0(x|\alpha - 1)}{\zeta_0(x|\alpha)^2} \zeta_1(x|\alpha) - \frac{1-\alpha}{\zeta_0(x|\alpha)} \zeta_1(x|\alpha - 1),$$

$$\mathcal{C}_i(x) = \frac{K_h(X_i - x)}{g_0(X_i)^\alpha}\left[\alpha h(x|\zeta,\alpha) - \frac{2\alpha^2}{\zeta_0(x|\alpha)^2}\zeta_1(x|\alpha)u_0(X_i)'\right.$$
$$\left. - \frac{\alpha}{\zeta_0(x|\alpha)}\{U_0(X_i) - \alpha u_0(X_i)u_0(X_i)'\}\right] - \alpha\zeta_0(x|\alpha-1)h(x|\zeta,\alpha)$$
$$-2\frac{\alpha(1-\alpha)}{\zeta_0(x|\alpha)^2}\zeta_1(x|\alpha)\zeta_1(x|\alpha-1)' - \frac{1-\alpha}{\zeta_0(x|\alpha)}\zeta_2(x|\alpha-1),$$

and $h(x|\zeta,\alpha) = \{2\alpha\zeta_1(x|\alpha)\zeta_1(x|\alpha)' + \zeta_0(x|\alpha)\zeta_2(x|\alpha)\}\zeta_0(x|\alpha)^{-3}$.

From the theory of M-estimator, we have the representation $\hat{\theta} - \theta_0 = H(F_n) - H(F) = \sum_{i=1}^n I_i/n + b/n + \varepsilon_n$, where $\varepsilon_n = O_p(1/n)$ with mean $O(1/n^2)$, b/n is the essential bias of $\hat{\theta}$ and $I_i = I(X_i)$ is the influence function of H with zero mean (see, Shao, 1991). Using this i.i.d. expression and somewhat tedious calculations yield that $E[\mathcal{B}_i(x)] = O(h^2)$, $E[\mathcal{B}_i(x)I_i] = O(h^2)$ and $E[\mathcal{C}_i(x)] = O(h^2)$. From these we have $E\left[(\hat{\theta}-\theta_0)'\bar{\mathcal{B}}_n(x)\right] = O(h^2/n + n^{-2})$ and $E\left[(\hat{\theta}-\theta_0)'\bar{\mathcal{C}}_n(x)(\hat{\theta}-\theta_0)\right] = O(h^2/n+n^{-2})$. Since the bias of f_α^* is given in Lemma 1, we have the bias term of $\hat{f}_\alpha$ in Proposition 1. The expression of variance can be obtained from similar calculations with referring to $Var\left[(\hat{\theta}-\theta_0)'\bar{\mathcal{B}}_n(x)\right] = O(h^4/n^2)$, $Var\left[(\hat{\theta}-\theta_0)'\bar{\mathcal{C}}_n(x)(\hat{\theta}-\theta_0)\right] = O(h^4/n^2)$ and $Cov\left[f_\alpha^*(x), (\hat{\theta}-\theta_0)'\bar{\mathcal{B}}_n(x)\right] = O(h^2/n + n^{-2})$. Detailed calculations are found in Naito (1998).

Acknowledgments

The author is grateful to the referee for comments. Thanks also go to M. Yoshizaki for his assistance in simulation. This research is supported in part by Grant-in-Aid for scientific research No.13780174 from Japan Society for the Promotion of Science.

REFERENCES

Bhattacharya, R. and Rao, R.R. (1976). *Normal Approximation and Asymptotic Expansions.* New York: Wiley.

Hjort, N.L. and Glad, I.K. (1995). Nonparametric density estimation with a parametric start. *Ann. Statist.* 23, 882-904.

Marron, J.S. and Wand, M.P. (1992). Exact mean integrated squared error. *Ann. Statist.* 20, 712-736.

Naito, K. (1998). Density estimation by local L_2-fitting. *Technical Report No. 98-7*, Statistical Research Group, Hiroshima University.

Olkin, I. and Spiegelman, C.H. (1987). A semiparametric approach to density estimation. *J. Amer. Statist. Assoc.* 82, 858-865.

Shao, J. (1991). Second-order differentiability and jackknife. *Statistica Sinica* 1, 185-202.

Wand, M.P. and Jones, M.C. (1995). *Kernel Smoothing.* London: Chapman and Hall.

A MODIFIED CYCLIC-COORDINATE EXCHANGE ALGORITHM AS ILLUSTRATED BY THE CONSTRUCTION OF MINIMUM-POINT SECOND-ORDER DESIGNS

NAM-KY NGUYEN
AV Biometrics, Department of
Natural Resources & Environment
Institute for Horticultural Development
Private Bag 15, Ferntree Gully
Delivery Centre, VIC 3156, Australia
E-mail: namky.nguyen@nre.vic.gov.au

Box and Draper (1974) reported five m-factor D-optimal minimum-point second order designs for $m = 2, \ldots, 5$. The m coordinates of each of the n design points lie in the interval $[-1, 1]$. This paper describes a method of constructing these designs and compares new designs with those of Box and Draper.

Keywords: computer-aided designs, D-optimality, second-order designs.

1 Introduction

A second-order response surface fitted for an experimental design with n runs and m explanatory variables may be written as:

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j x_{ij} + \sum_{j=1}^{m} \sum_{j'=j}^{m} \beta_{jj'} x_{ij} x_{ij'} + e_i \qquad (1)$$

where $i = 1, \ldots, n$; y is the response variable; and the m x's are scaled predictor variables. The minimum-point second order design problem can be formulated as being to find n design points $(x_{i1}, \ldots, x_{im})$, $i = 1, \ldots, n$ such that: (i) the coordinates of each design point lie in the interval $[-1, 1]$; (ii) n equals the number of parameters in the model, *i.e.* $(m+1)(m+2)/2$; and (iii) $|X'X|$ should take its maximum value where X is the $n \times n$ matrix whose ith row is $x_i' = (1, x_{i1}, \ldots, x_{im}, x_{i1}^2, x_{i1}x_{i2}, \ldots, x_{i,m-1}x_{im})$ (see Box and Draper, 1974).

The following example describes a possible use of a minimum-point second-order design. In chocolate manufacture, conching converts the refined powdery ingredient mix into a suspension with the characteristic flow and flavour properties of chocolate. Over time, shear forces and heat combine to coat the suspended particles in fat and to release and develop essential

and flavour properties of chocolate. Over time, shear forces and heat combine to coat the suspended particles in fat and to release and develop essential flavours. Problems faced by the chocolate manufacturer include minimizing the (costly) cocoa butter addition, reducing processing time and assessing new equipment while still maintaining product quality. A second-order design might be required to study the effects of four factors on the chocolate quality indices. These factors are (i) conching time (10-20 hours), (ii) mixing temperature (60^0-80^0C), (iii) mixing speed (100-500 RPM) and (iv) cocoa butter addition (25%-40%). While no restriction is put on the number of levels of each of the mentioned factors, the company aims to have the optimization results within one month during which a total of 15 runs could be completed.

2 Method of construction

A cyclic-coordinate exchange algorithm (CCEA) was used to construct the required designs. In a sense this algorithm is a hybrid of the adjustment algorithm (AA) of Donev and Atkinson (1988) and the CCEA of Meyer and Nachtsheim (1995). The new CCEA starts with a non-singular D-optimal starting design D with points in $\{-1, 0, 1\}^m$. This design can be constructed by any of the conventional D-optimal design exchange algorithms (see Nguyen and Miller (1992)). The new CCEA then follows the following steps:

- For each row i of D, pick a coordinate x_{ij}, $(j = 1, \ldots, m)$ at random and perturb this coordinate by an amount σ. Calculate $M = X'X + \epsilon I$, $|M|$ and M^{-1} where ϵ is a very small positive number say 1.0e-5. Note that the choice of n perturbations is empirically based and the the purpose of adding ϵI to $X'X$ is to avoid possible singularity caused by these perturbations.

- Pick a coordinate x_{ij}, $(i = 1, \ldots, n;\ j = 1, \ldots, m)$ at random among the nm coordinates of D.

- Remove the ith row x_i' from X. Update $|M|$ and M^{-1} by the formula:

$$|M - x_i x_i'| = |M|(1 - x_i' M^{-1} x_i),$$
$$(M - x_i x_i')^{-1} = M^{-1} + u_i u_i'/(1 - x_i' u_i)$$

  where $u_i = M^{-1} x_i$ (see Equations 3.3 and 3.4 of Nguyen and Miller (1992).

- Reinstate the deleted ith row x_i' of X with the coordinate x_{ij} replaced by a value which circles from -1 to $+1$ in with step length σ. Let x_{ij}^* be

a value which gives the biggest increase in $x_i' M^{-1} x_i$ where x_i' is now the reinstated row with x_{ij} replaced by x_{ij}^*. Update $|M|$ and M^{-1} by the formula:

$$\begin{aligned} |M + x_i x_i'| &= |M|(1 + x_i' M^{-1} x_i), (M + x_i x_i')^{-1} \\ &= M^{-1} - u_i u_i'/(1 + x_i' u_i) \end{aligned}$$

where $u_i = M^{-1} x_i$ (see Equations 3.1 and 3.2 of Nguyen and Miller (1992).

Steps 2-4 are repeated until the number of coordinates is exhausted and no further improvement in $|M|$ is found. In these steps, σ is first set at 1 and then at 0.1 and finally at 0.01. From our experience, there is no further gain in decreasing σ further.

- Remove the effect of ϵ in step 1 by recalculating $|X'X|$.

The five steps of the above CCEA correspond to a single try. Several tries are made for each (m, n) combination and the try which corresponds to the highest $|X'X|$ will be chosen.

Remark:

1. Both the AA and the CCEA of Meyer and Nachtsheim (1995) use a different starting design for each try. The new CCEA uses the same starting design for all the tries. This approach saves time as we only have to construct the starting design once. Note that there is a resemblance of this approach and the approach used in Nguyen and Williams (1993) where different row-column designs are constructed from a single incomple block design by shuffling the treaments in each block of this block design.

2. With the AA, the candidate set for each coordinate x_{ij} of D is $\{x_{ij} - \sigma, x_{ij} + \sigma\}$. With the CCEA, the candidate set for each coordinate x_{ij} of D is $\{-1, -1 + \sigma, \ldots, -\sigma, 0, \sigma, \ldots, 1 - \sigma, 1\}$.

3. While the new CCEA works with all the n points, the one of Meyer and Nachtsheim (1995) only works with k $(k < n)$ points which has the minimum $x_i' M^{-1} x_i$. While this practice saves some computer time, there is a substantial risk that their CCEA misses the best design.

4. While steps 3 and 4 in the new CCEA are separate steps, they are combined into single step in the one of Meyer and Nachtsheim (1995). Basically, they attempted to replace simultaneously row x_i' with another one with the coordinate x_{ij} replaced by x_{ij}^*. The update of $|M|$, unlike (2) and (4) requires more calculation (see Equations 3.5-3.9 of Nguyen and Miller (1992).

Table 1. Standardized $|X'X|$ for various designs, $2 \leq m \leq 5$

m	n	Box & Draper	Dubova & Federov	Donev & Atkinson	Nguyen
2	6	5.74e-3	5.72e-3	5.74e-3	5.74e-3
3	10	1.85e-4	1.85e-4	1.85e-4	1.85e-4
4	15	3.89e-7	3.37e-6	2.94e-6	3.45e-6
5	21	3.39e-11	-	8.46e-8	1.15e-7

3 Discussion

Table 1 presents the standardized $|X'X|$ $(= |X'X|/n^p)$ for minimum point second-order designs for $m = 2, \ldots, 5$ of Box and Draper (1974), Dubova and Federov (1972), Donev and Atkinson (1988) and new ones.

The new designs for m=2 and m=3 have a pattern similar to the corresponding ones of Box and Draper (1974). The new design for m=4 have a pattern similar to the corresponding one of Dubova and Federov (1972) (see Table 2 of Box and Draper, 1974)) except that the values of of the coordinates which are different form -1 or $+1$, i.e. $(\alpha, \beta, \gamma, \delta, \epsilon)$ are set at $(-0.22, -0.27, -0.63, 0.04, -0.04)$. Note that there was a misprint in Table 2 of Box and Draper (1974). The first coordinate of the first factor in this Table should be 1 instead of -1. If both δ and ϵ are set to 0, the maximum number of levels for each factor of this design (with the standardized $|X'X|$ equal to 3.44e-6) will be reduced from five to four. This is actually the design recommended for the experiment discussed in the Introduction. Table 2 gives the new design for m=5.

Each factor of the new designs in Table 1 has at most five levels. Note that the standardized $|X'X|$ for the corresponding D-optimal minimum-point second order designs in $\{-1, 0, 1\}^m$ for $m = 2, \ldots, 5$ are 5.49e-3, 1.33e-4 , 2.67e-6 and 7.89e-8. Hence, if there is no restriction to three level for each factor, the new designs in Table 1 are highly recommended.

As an additional note, the new CCEA improves $|X'X|$ of seven out of 17 designs (and matches $|X'X|$ of the remaining designs) in Table 1 of Donev and Atkinson (1988). The standardized $|X'X|$ of designs for (m, n)=(3,14),(3,16), (3,20), (4,15), (4,24), (5,21) and (5,26) in this table are 4.531e-4, 4.136e-4, 4.650e-4, 2.941e-6, 1.340e-5, 8.465e-8, 2.283e-7. The ones of the corresponding designs constructed by the new CCEA are 4.553e-4, 4.162e-4, 4.670e-4, 3.454e-6, 1.352e-5, 1.1474e-7 and 2.324e-7.

Table 2. Design for m=5†

x_1	x_2	x_3	x_4	x_5
-1	-1	-1	-1	-1
-1	1	1	1	1
1	-1	1	1	1
1	1	-1	1	1
1	1	1	-1	1
1	1	1	1	-1
1	1	-1	-1	-1
1	-1	1	-1	-1
1	-1	-1	-1	1
-1	1	-1	1	-1
-1	1	-1	-1	1
-1	-1	1	1	-1
-1	-1	1	-1	1
-1	-1	-1	1	1
α	$-\alpha$	-1	1	-1
1	-1	$-\alpha$	α	-1
$-\beta$	1	β	-1	-1
-1	β	1	$-\beta$	-1
γ	-1	1	γ	δ
-1	$-\gamma$	$-\gamma$	1	δ
-1	1	1	-1	$-\epsilon$

†$(\alpha, \beta, \gamma, \delta, \epsilon)$=(0.31,0.14,0.01,0.06,0.17)

In this paper, minimum-point second-order designs have been constructed in order to illustrate the performance of the new CCEA. However, this CCEA should work with any practical values of m and n. The use of this CCEA in constructing large designs of the cyclic family such as cyclic incomplete block designs and Alpha designs will be discussed elsewhere (see John and Williams (1987) Chapters 3-4 for the discussion of these types of designs). The byte code of the Java program implementing the CCEA discussed in this paper is available free of cost from the author.

Acknowledgments

The author would like to thank a referee for valuable comments and suggestions. This research was done while the author was a staff member of CSIRO

Mathematical and Information Sciences.

REFERENCES

Box, M.J. and Draper, N.R. (1974). On minimum-point second-order designs, *Technometrics* 16, 613-616.

Donev, A.N. and Atkinson, A.C. (1988). An adjustment algorithm for construction of the exact D-optimal experimental designs. *Technometrics* 30, 429-433.

Dubova, I.S. and Federov, V.V. (1972). Tables of optimum design II (saturated optimal designs on a cube). Preprint No. 40, in Russian. Issued by Interfaculty Laboratory of Statistical methods, Moscow University.

John, J.A. and Williams, E.R. (1987). *Cyclic designs and computer-generated designs.* New York: Chapman and Hall.

Meyer, R.K. and Nachtsheim, C.J. (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics* 37, 60-69.

Nguyen, N-K. and Miller, A.J. (1992). A review of exchange algorithms for constructing discrete D-optimal designs. *Comp. Stat. and Data Analysis* 14, 489-498.

Nguyen, N-K. and Williams, E.R. (1993). An algorithm for constructing optimal resolvable row-column designs. *Austral. J. of Stat.* 35, 363-370.

HAS STATISTICS A FUTURE? IF SO, IN WHAT FORM?

C. RADHAKRISHNA RAO
Emeritus Eberly Professor of Statistics
Pennsylvania State University
University Park, PA 16802
E-mail: crr1@email.psu.edu

The mathematical foundations of statistics as a separate discipline were laid by Fisher, Neyman and Wald during the second quarter of the last century. Subsequent research in statistics and the courses taught in the universities are mostly based on the guidelines set by these pioneers. Statistics is used in some form or other in all areas of human endeavor from scientific research to optimum use of resources for social welfare, prediction and decision-making. However, there are controversies in statistics, especially in the choice of a model for data, use of prior probabilities and subject-matter judgments by experts. The same data analyzed by different consulting statisticians may lead to different conclusions.
What is the future of statistics in the present millennium dominated by information technology encompassing the whole of communications, interaction with intelligent systems, massive data bases, and complex information processing networks? The current statistical methodology based on simple probabilistic models developed for the analysis of small data sets appears to be inadequate to meet the needs of customers for quick on line processing of data and making the information available for practical use. Some methods are being put forward in the name of data mining for such purposes. A broad review of the current state of the art in statistics, its merits and demerits, and possible future developments will be presented

Keywords: Bayesian analysis, cross validation, data mining, decision theory, estimation, hypothesis testing, large data sets, machine learning.

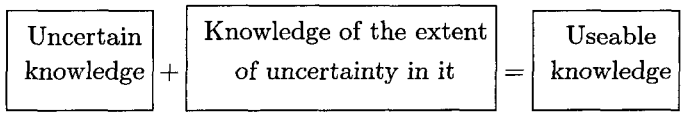
1 Statistics as a separate discipline

1.1 A paradigm for statistical theory and methods

The word statistics was coined by the German Scholar Gottfried Achenwall about the middle of the eighteenth century in the context of *collection, processing and use of data by government.*

During the nineteenth century, statistics acquired a new meaning as extraction of information from data for decision making. The need arose especially in testing hypotheses or making predictions or forecasts based on information in the observations made on natural phenomena or generated through well designed experiments. It was realized that the information contained in particular data, however well they are ascertained, is subject to some

uncertainty and consequently our conclusions based on observed data could be wrong. How then can we acquire new knowledge? We have to evolve a new methodology of data analysis with a view to estimate the amount of uncertainty in extracted information and to formulate rules for making decisions with minimal risk. The equation



is used as a *new paradigm* for statistical theory and methods. Thus, statistics acquired the status of a new discipline of study for

* acquiring data with maximum possible information for given cost,

* processing data to quantify the amount of uncertainty in answering particular questions, and

* making optimal decisions (subject to minimal risk) under uncertainty.

The first systematic efforts for the development of statistical methodology began only in the beginning of the 20-th century, and it is only in the first half of the century the basic concepts of statistical inference were introduced, (Table 1), which enabled rapid developments to take place for possible applications in all areas of human endeavor ranging from natural and social sciences, engineering and technology, management and economic affairs, to arts, literature, medicine and legal problems. Knowledge of statistics was considered to be essential in all fields of inquiry. Courses in statistics were introduced in the curriculum of social sciences. Specialized books dealing with the applications of statistics in particular areas were written as guidance to research workers. Referring to the ubiquity of statistics, Sir Ronald Fisher (1953), in his presidential address to the Royal Statistical Society in 1952, made the optimistic statement:

> *I venture to suggest that statistical science is the peculiar aspect of human progress which gives to the twentieth century its special character; and indeed members of my present audience will know from their own personal and professional experience that it is to the statistician that the present age turns for what is most essential in all its more important activities.*

Table 1. Mathematical foundation of theoretical statistics

	Pre-1900	
T. Bayes	1764	Bayes Theorem
	1900-1950	
K. Pearson	1900	Chi-square test
W.S. Gosset	1908	Students' t-test
F.A. Fisher	1915	Exact distributions of statistics
	1922	Estimation (maximum likelihood)
	1923	Analysis of Variance
	1926	Design of Experiments
J. Neyman E.S. Pearson	1928	*Likelihood ratio test
	1933	Testing of hypotheses
E.J.G. Pitman	1937	Nonparametric tests
H. Jeffreys	1939	Bayesian statistics
A. Wald	1943	*Asymptotic test
P.C. Mahalanobis M. Hansen	1944	Sample surveys
A. Wald	1947	Sequential tests
C.R. Rao	1948	*Score test
A. Wald	1950	Decision theory

* The three general asymptotic tests are referred to as the *Holy Trinity.*

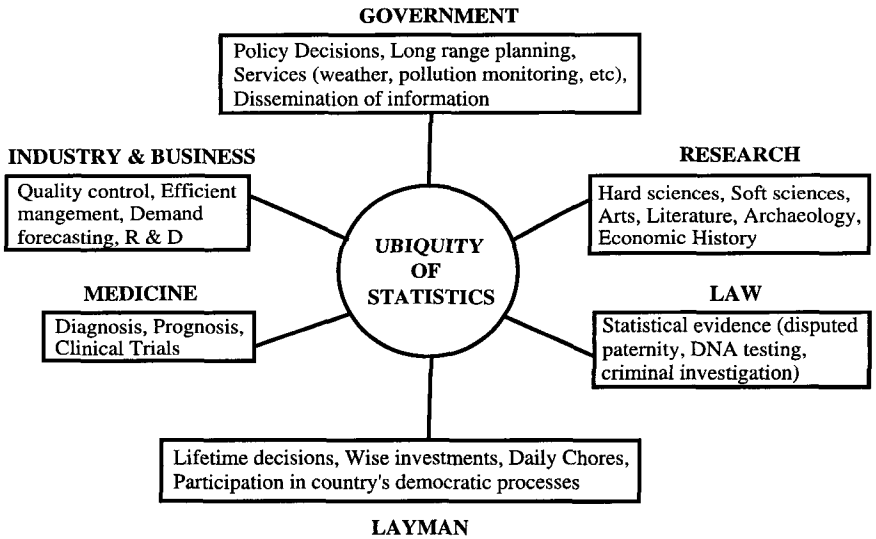


Figure 1. The Ubiquity of Statistics

The scope of statistics as it is understood, studied and practiced today extends to all areas of human activity as shown in Figure 1.

The *layman* uses statistics (information obtained through data of various kinds and their analyses published in newspapers and consumer reports) for taking decisions in daily life, or making future plans, deciding on wise investments in buying stocks and shares, etc.. Some amount of statistical knowledge may be necessary for a proper understanding and utilization of all the available information and to guard oneself against misleading advertisements. The need for statistical literacy in our modern age dominated by science and technology was foreseen by H.G. Wells:

> *Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.*

For the *government* of a country, statistics is the means by which it can make short and long range plans to achieve specified economic and social goals. Sophisticated statistical techniques are applied to make forecasts of population and the demand for consumer goods and services and to formulate economic plans using appropriate models to achieve a desired rate of progress

in social welfare.

In *scientific research*, statistics plays an important role in the collection of data through efficiently designed experiments, in testing hypotheses and estimation of unknown parameters, and in interpretation of results.

In *industry*, extremely simple statistical techniques are used to improve and maintain the quality of manufactured goods at desired levels. Experiments are conducted in R. & D. departments to determine the optimum mix (combination) of factors to increase the yield or give the best possible performance. It is a common experience all over the world that in plants where statistical methods are exploited, production has increased by 10% to 100% without further investment or expansion of plant. In this sense statistical knowledge is considered as a *national resource*. It is not surprising that a recent book on modern inventions lists statistical quality control as one of the great technological inventions of the last century.

In *business*, statistical methods are employed to forecast future demand for goods, to plan for production, and to evolve efficient management techniques to maximize profit.

In *medicine*, principles of design of experiments are used in screening of drugs and in clinical trials. The information supplied by a large number of biochemical and other tests is statistically assessed for diagnosis and prognosis of disease. The application of statistical techniques has made medical diagnosis more objective by combining the collective wisdom of the best possible experts with the knowledge on distinctions between diseases indicated by tests.

In *literature*, statistical methods are used in quantifying an author's style, which is useful in settling cases of disputed authorship.

In *archaeology*, quantitative assessment of similarity between objects has provided a method of placing ancient artifacts in a chronological order.

In *courts of law*, statistical evidence in the form of probability of occurrence of certain events, such as similarity of DNA, is used to supplement the traditional oral and circumstantial evidences in judging cases.

There seems to be no human activity whose value cannot be enhanced by injecting statistical ideas in planning and by using results for feedback and control. It is apodictic to claim: *If there is a problem to be solved, seek for statistical advise instead of appointing a committee of experts. Statistics and statistical analysis can throw more light than the collective wisdom of the articulate few.*

In the book on *Statistics and Truth* by Rao (1997b) numerous examples are given in Chapters 5 and 6 of applications of statistical techniques to a variety of problems ranging from disputed authorship, disputed paternity,

seriation of Plato's works, foliation of manuscripts, dating of publications and construction of language trees, to weather forecasting, public opinion polls and extra sensory perception.

1.2 What is statistics?

Is statistics a science, a technology, or an art? Statistics is not a subject like the basic disciplines of mathematics, physics, chemistry or biology. Each of these disciplines has a subject matter of its own and problems of its own which are solved by using the knowledge of the subject. There is nothing like a statistical problem which statistics purports to solve. Statistics is used to solve problems in other disciplines and appropriate methodology is developed for any given situation. The following Table 2 from a paper by Box (1980) shows how most of the important concepts in statistics were motivated by practical problems. In course of time, the subject matter of statistics grew from isolated methods applied to particular problems to the consolidation of different methods under a unified theory based on the concepts of probability. The basic problem of statistics is viewed as quantification of uncertainty, which may be considered as the subject matter of statistics for study and research. As it is practiced today, statistics appears to be a combination of science, technology and art.

It is *science* in the sense that it has an identity of its own with a large repertoire of techniques derived from some basic principles. These techniques cannot be used in a routine way; the user must acquire the necessary expertise to choose the right technique in a given situation and make modifications, if necessary. Further, there are philosophical issues connected with the foundations of statistics - the way uncertainty can be quantified and used - which can be discussed independently of any subject matter. Thus in a broader sense statistics is a separate discipline.

It is *technology* in the sense that statistical methodology can be built into any operating system to maintain a desired level and stability of performance, as in quality control programs in industrial production. Statistical quality control is described as one of the great technological inventions of the 20th century. Statistical methods can also be used to control, reduce and make allowance for uncertainty and thereby maximize the efficiency of individual and institutional efforts.

Statistics is also *art*, because its methodology which depends on inductive reasoning is not fully codified or free from controversies. Different statisticians may arrive at different conclusions working with the same data set. There are frequentists, Bayesians, neo-Bayesians and empirical Bayesians among statis-

Table 2. Practical problems motivating general statistical concepts (George Box (1980))

Practical problem	Investigator	Derived general concept
Analysis of Asteroid Data.How far is it from Berlin to Potsdam?	Gauss	Least squares
Are planetary orbits randomly distributed	Daniel Bernoulli	Hypothesis testing
What is the population of France?	Laplace	Ratio estimators
How to handle small samples of brewery data	Gosset	t-test
Improving agricultural practice by using field trials	Fisher	Design of experiments
Do potato varieties and fertilizers interact?	Fisher	Analysis of variance
Accounting for strange cycles in U.K. wheat prices.	Yule	Parametric time series models
Economic inspection (of ammunition).	Wald Barnard	Sequential tests
Need to perform large numbers of statistical tests in pharmaceutical industry before computers were available	Wilcoxon	Nonparametric tests
Advanced estimates of agricultural production	Mahalanobis	Sample surveys*

* not quoted by Box

ticians each one advocating a different approach to data analysis. (A familiar quote on statistics: *If there are 3 statisticians on a committee, there will be 4 minority reports.* See also Van den Berg (1992) who conducted a survey and found that statisticians with different backgrounds used different methods for the analysis of the same data.) There is usually more information in given

data than what can be extracted by available statistical tools. Making figures tell their own story depends on the skill and experience of a statistician, which makes statistics an art. Perhaps, *statistics is more a way of thinking or reasoning than a bunch of prescriptions for beating data to elicit answers.*

While mathematics is the logic of deducing consequences from given premises, statistics may be regarded as a rational approach to learning from experience and the logic of identifying the premises given the consequences, or inductive reasoning as it is called. Both mathematics and statistics are important in all human endeavors whether it is in the advancement of natural knowledge or in the efficient management of our daily chores.

1.3 Fisherian framework

In his fundamental paper on mathematical foundations of theoretical statistics, Fisher (1922) stated three methodological aspects of statistics:

- *Specification* (choice of a stochastic model for data)
- *Estimation* (of unknown parameters in the chosen model)
- *Testing of hypotheses* (seeking evidence from data for possible rejection of a specified hypothesis or theory)

Fisher's framework has been and still is the basis for the development of statistical methods. However, there are difficulties in using these concepts and methods based on them in statistical analysis of real data.

First is the specification, or the choice of a stochastic model for data: Fisher did not specify any statistical method for model selection. He acknowledged the usefulness of Karl Pearson's chi-square test for specification, only as a method for possible rejection of a given model and not for its acceptance. Reference may be made to Inman (1994) for a review of the controversy between Pearson and Fisher on the role of the chi-square criterion in accepting or rejecting a specified model for data.

In recent years, several model selection criteria have been suggested such as AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion), and GIC (General Information Criterion), an extensive review of which can be found in a paper by Rao and Wu (2001). These methods are not directly related to the performance of the estimated models for it is known that different models may have to be used in the analysis of the same data for different purposes as shown by Rao (1977, 1987). Further, model selection by using AIC, BIC and GIC depends on the sample size; a larger sample size may choose a more complex model.

Recent studies in chaos theory show that there are difficulties in distinguishing between sequences of observations produced by deterministic and random mechanisms. Attempts are beginning to be made for modeling a sequence of variables such as time series as a combination of deterministic and random components. (See Cox (1990), Lehmann (1990) and Rao (1997b, pp.26-28)).

In real situations scientists are looking for what may be called working hypotheses (which may not be strictly true) which enable prediction of events with reasonable accuracy. So the main question should be to ask how *good* a proposed model or theory is in explaining the observed data and in predicting future events, and not whether the proposed model is *true or false. A working hypothesis is rejected if a better working hypothesis is found. This is how science progresses creating useful knowledge from time to time.*

Fisherian framework provided the basis for the development of theoretical statistics during the first half of the 20-th century as shown in Table 1. Neyman and Pearson (1933) developed a theory of testing of hypotheses using the concept of the power function of a test (with respect to possible alternatives to the given hypothesis) for comparing different test criteria and choosing the one with some optimum properties for the power function. The theory provided a justification of some of the test criteria introduced by Fisher on an intuitive basis. Pitman (1937) developed nonparametric tests which do not depend on any stochastic model for data. Wald (1950) formulated estimation and testing of hypothesis in a decision theoretic set up considering a loss function as an input into the problem. Wald (1947) also developed the theory of sequential testing. General asymptotic test criteria, called the *Holy Trinity*, were introduced by Neyman and Pearson (1928), Wald (1941) and Rao (1949).

It is relevant to mention that Fisher introduced the basic concepts of statistical inference and developed the related statistical methodology when computers capable of performing complex computations were not available and there were serious limitations on acquisition of data. Under these restrictions, the statistical methodology developed was *mostly model oriented*, i.e., under the assumption that the observed data is a random sample from a population belonging to a specified family of distributions functions. Often a simple stochastic model was chosen, like the normal distribution, to provide exact results (closed form solutions to problems) involving minimum computations.

Tables of limited percentage points of test statistics were constructed (using desk computers) choosing the normal as the underlying distribution and rules were laid down for rejection of hypotheses at the tabulated levels

of significance, usually 5% and 1%. Limitation on the sample size made it difficult to verify model assumptions. [Commenting on the mistrust of British statistical methods by continental statisticians, Buchanan-Wollaston (1935) says, *"The fact that British methods "work" is due to prevalence in Nature of distributions similar to Gaussian rather than to any peculiar value in the methods themselves"*].

Second is the method of estimation: The method of maximum likelihood introduced by Fisher is valuable in the estimation of parameters when the model for data is known and the sample size is not small. However, it is not robust for slight departures from the specified model and for outliers in data. Robust methods of estimation known as M-estimation and associated tests of hypotheses have received much attention. A variety of procedures have been introduced (without any guidance on what to choose) to eliminate or minimize the influence of outliers or contamination in data. The theory is mostly asymptotic and the performance of M-estimates in small samples has not been adequately examined. The character of research in this area is described by Tukey (1993) as *asymptotitise*.

There are also controversies in expressing the precision of an estimator. Fisher suggested the conditional variance of an estimator given an ancillary statistic as a measure of precision. But, as pointed out by Basu (see Ghosh, 1988, pp.3-19), there are difficulties in such a procedure. First there is, in general, no maximal ancillary statistic and different choices of ancillary statistics lead to different expressions for precision, and there is no way of choosing one in preference to the other. Basu also gives examples where the conditional distribution given an ancillary statistic becomes degenerate (Ghosh, 1988, pp.161-167), and uninformative about the parameter.

Third is testing of hypotheses: Fisher considered a test of significance using an appropriate criterion as a method which can lead to possible rejection of a given hypothesis, but not for establishing a given hypothesis as certainly true. He used tests of significance in an ingenious way, an example of which is the discovery of the Rhesus factor described in Fisher (1948). It is a brilliant example of how hypothesis testing can be of help "in fitting one scrupulously ascertained fact into another, in building a coherent structure for knowledge and seeing how each gain can be used as a means for further research. Fisher also used tests of significance to detect irregularities in data such as lack of randomness, recording errors or bias in sampling. In this connection, reader is referred to Fisher (1936), where he shows that Mendel's data on his genetic studies are probably faked and to Fisher (1934) where he studies the effects of different methods of ascertaining data in genetic studies of inheritance of diseases.

There is, however, some debate among statisticians on the usefulness of tests of significance. The null hypothesis H_0 as formulated in many problems is known to be wrong and no test of significance is needed. What is of interest is to determine to what extent the true hypothesis can differ from H_0, which is a problem of estimation rather than of testing a hypothesis. Frank Yates, a long time associate of R.A. Fisher, mentioned in the obituary published in *Biographical Memoires of the Royal Society* that Fisher laid too much stress on hypothesis testing. He said that if we are comparing the yields of two varieties of corn, it is useful to ask what the *difference in yields is* rather than whether they *have the same yield*, which is seldom true.

In Fisher-Neyman framework of testing a null hypothesis, there is some controversy about the level of significance of a test. From the early use of tests of significance by Fisher and the axiomate set up by Neyman, by level of significance is meant "the frequency with which the hypothesis is rejected in repeated sampling of any fixed population allowed by the hypothesis". In his last book, Fisher (1956, p.91) disassociated himself from such a view by saying: "This intrusive axiom, which is foreign to the reasoning on which tests of significance were in fact based seems to be a real bar to progress". On p.77, Fisher says: "the population in question is hypothetical, that it could be defined in many ways..., or, that an understanding, of what the information is which the test is to supply, is needed before an appropriate population, if indeed we must express ourselves in this way, can be specified". Fisher has not made explicit how the level of significance can be ascertained given the data.

The well known mathematical statistician, J. Wolfowitz (1967) reviewing a popular book on testing of hypotheses made the following critical comment.

> *"... the history of testing of hypothesis is an example of collaboration between theoreticians and practical statisticians which has resulted in greater obfuscation of important statistical problems and side tracking of much statistical effort."*

Wolfowitz believed that a useful approach to statistical analysis of live data is *Decision Theory* as developed by Wald (1950), which needs inputs such as the class of alternative hypotheses, prior probabilities and losses associated with different possible decisions. Such a procedure of choosing a hypothesis to minimize the expected loss can be implemented in certain situations like acceptance sampling (such as accepting or rejecting batches of goods produced in a factory), but does not seem to be applicable in scientific research.

When a test is applied to test a hypothesis, there are two possible scenarios:

1. the hypothesis is rejected as not being true;
2. the hypothesis is not rejected, but this does not mean that it is accepted as true.

In either case, the scientist has to continue his search for an alternative hypothesis. Does statistics help in the search for an alternative hypothesis? There is no codified statistical methodology for this purpose. Text books on statistics do not discuss either in general terms or through examples *how to elicit clues from data to formulate an alternative hypothesis or theory when a given hypothesis is rejected.*

Imagine the following scenario of a possible dialogue between Einstein and some contemporary statisticians.

> **Einstein**: *I have a new theory for explaining some natural phenomena. Can statisticians help in testing it?*
>
> **Neyman and Pearson respond**: *Einstein, you have to do your own experiment, give us your data and also tell us what the possible alternatives are to your theory. We can then tell you the most powerful method of verifying your theory.*
>
> **Einstein**: *Alternative theories! There may be but I do not know.*
>
> **Fisher responds**: *I can give you the design of a perfect experiment to perform. The results can reject your theory if it is wrong and cannot confirm it if it is true.*
>
> **Einstein** : *I am disappointed, you cannot confirm it if it is true.*
>
> **Wald and Wolfowitz respond**: *We would like to review your problem in terms of decision theory. Apart from other inputs, we need to know the losses involved in accepting and rejecting your theory.*
>
> **Einstein**: "*If my theory is proven successful, Germany will claim me as a German and France will declare that I am a citizen of the world. Should my theory prove untrue France will say that I am German and Germany will declare that I am a Jew*". [*This is a true statement made by Einstein in an address at the Sorbonne.*]

Fisher (1935) also introduced nonparametric tests based on permutation distributions using the randomization principle, which was hailed as an important contribution by Neyman and others. Unfortunately, the randomization

principle is not without logical difficulties (see the paper by Basu on pp.290-312 in Ghosh (1988)).

A notable contribution to nonparametric testing is Efron's (1979) bootstrap, which has become popular with the enormous computing power we now have. However, its theoretical justification is again based on asymptotics and the consequences of bootstrapping in small samples have not been fully examined. A related method is Jackknife which Tukey (1977) recommends as an alternative. Efron (1979) gives a comparative study of bootstrap and Jackknife techniques.

1.4 Bayesian analysis

Some of the inconsistencies in the classical methods described in previous sections of this paper, which depend on properties based on repeated sampling from a population, led several statisticians to use Bayesian methods in data analysis. References to papers emphasizing the need for Bayesian analysis are Berger (2002) and Basu's contributions reproduced in Ghosh (1988). It is argued that in classical statistics of Fisher, Neyman and Wald, statistical methods for drawing inferences on unknown parameters are judged by their (average) performance in repeated sampling from a population. Such a procedure ignores the fact that *all samples are not equally informative* and inference on unknown parameters should be made conditional on the observed sample, which makes the use of Bayesian analysis inevitable.

Other arguments advanced by Bayesians against classical testing procedure refer to the interpretation of p-values and paradoxes associated with it. Lindley (1957) gave an example in which the p-value is fixed at .05, but as the sample size increases, the Bayesian posterior probability that the null hypothesis is true approaches unity. It is also shown that for a large class of priors

$$p = p(T > T_{obs}|H_0) \leq p(H_0|\text{Data})$$

where H_0 is a null hypothesis and T is a test statistic. The above inequality shows that the use of p-values *exaggerates* significance.

Bayesian analysis depends on Bayes theorem

$$p(s|x) = \frac{p(s)p(x|s)}{p(x)}$$

where $p(s)$ is the *prior probability* distribution on the space S of specified family of models, $p(x|s)$ is the probability density of the observation x for

specified $s \in S$, $p(x)$ is the overall probability density of x, and $p(s|x)$ is the *posterior probability* of s given x.

An attractive feature of Bayes theorem is that we can make probability statements about the probability models in the light of observed data x using the posterior density $p(s|x)$. But the question is: *Where does $p(s)$, the prior probability come from?* Berger (2002) lists five different approaches to the problem "each of which can be of great value in certain situations and for certain users":

- *Objective* (non-informative or default priors, maximum entropy and reference priors)
- *Subjective* and partly subjective and partly objective
- *Robust priors*
- *Frequentist*-Bayes
- *Quasi*-Bayes
- *Empirical* Bayes (strictly not a Bayesian analysis)

However, it appears that there is no unified approach to Bayesian analysis. How can Bayesian analysis be implemented in the context of a customer, who generates data to throw light on a particular problem, approaches a statistical consultant for analysis of data. Who supplies the input on prior? Surely, not the consultant, but should he (or she) accept the customer's prior? Reference may be made to Cox (2000) for further comments on Bayesian analysis.

1.5 Likelihood principle

It is generally agreed by statisticians belonging to different schools of thought that the likelihood function introduced by Fisher (1922) plays a pivotal role in statistical inference. The ideal situation is the combined use of the prior distribution and likelihood function to derive the posterior distribution. Attempts have been made in the case of a single parameter θ, to suggest plausible ranges of the parameter, without using priors, using only the ratio $r = L(\hat{\theta}|x)/L(\theta|x)$, where $L(\theta|x)$ is the likelihood function of θ given the observation x and $\hat{\theta}$ is the maximum likelihood estimate of θ. Fisher, Jeffreys and Royall suggested ranges of r as given in Table 3 for the classifications *very plausible*, *somewhat implausible* and *highly implausible.*

Table 3. Ranges to classify θ

θ Classification	Fisher	Jeffreys	Royall
Very plausible	$r \in (1,2)$	$r \in (1,3)$	$r \in (1,4)$
Somewhat implausible	$r \in (2,5)$	$r \in (3,10)$	$r \in (4,8)$
Highly implausible	$r \in (5,15)$	$r \in (10,100)$	$r \in (8,32)$

It is not clear how these ranges are obtained and how they could be used in practice. When the full likelihood cannot be used due to nuisance parameters, several modified versions have been suggested such as, *partial likelihood, pseudo likelihood, quasi-likelihood, empirical likelihood and a predictive likelihood.* For further details on the likelihood principle, reference may be made to Ghosh (1988, pp.313-320) and Reid (2002).

It is surprising that Fisher, who in his early research work emphasized tests of significance based on a test statistic and the p-value in the tails of its distribution, recommended in his last book (Fisher (1956)) the use of likelihood ratio without any reference to its probability distribution.

There is another problem about the use of likelihood which is often referred to as Rao's paradox in sample surveys (see Rao (1971), Cox (1997) and Smith (1997)). Consider a finite population defined by the set $\{(Y_i, X_i),\ i = 1, \ldots, N\}$, where Y_i is a label identifying the i-th member and X_i is, in a general set up, a random variable with probability density $f_i(X, \theta_i)$ depending on an unknown parameter θ_i. A sample is a selection of n pairs

$$(y_1, x_1), \ldots, (y_n, x_n)$$

drawn from $\{Y_i, X_i\}$, where y_i takes one of the values $(Y_1, \ldots, Y_n)$ and x_i is an observation on X_i. The problem is that of estimating a function of $\theta_1, \ldots, \theta_N$. The probability density at the observed values $\{(y_i, x_i)\}$, i.e., the likelihood of the parameters based on the sample is

$$\prod_{i=1}^{n} \frac{1}{N} f_{r_i}(x_i, \theta_{r_i})$$

where the i-th pair in the sample corresponds to the unit labeled r_i with parameter θ_i. The above likelihood function contains only n of the parameters $(\theta_1, \ldots, \theta_N)$ and has thus no information on the rest of the $N - n$ parameters. In such a case the likelihood approach based on the entire sample is not applicable.

However, if we disregard the labels and retain only $(x_1, \ldots, x_n)$, then the likelihood based on $(x_1, \ldots, x_n)$ under random sampling is

$$\prod_{i=1}^{n} \frac{1}{N} \sum_{r=1}^{N} f_r(x_i, \theta_r)$$

which contains all the unknown parameters. If one wants to use the likelihood principle, *it may be necessary to throw away part of the data*! This raises new problems on the choice of statistics (functions of the sample) for setting up the likelihood function without loss of information.

2 Statistics in the information age

2.1 Limitations of the current statistical methods

Mostly model oriented. Where does the stochastic model for data come from? This question has often be debated. From whoever it may come, either the customer who may have some knowledge of the data and the mechanism generating the data, or the statistical consultant from his previous experience of similar data, there is no reasonable statistical procedure for validating it for use on current data. The reader is referred to a recent paper by Breiman (2001) and the discussion by Cox, Efron, Parzen and others. The author mentions two cultures, data modeling (practiced by 98% of all statisticians) and algorithmic modeling (practiced by 2%) and makes a strong case for model free analysis using techniques such as *neural networks*, *genetic algorithms*, *machine learning* (support vector machines of Vipnik).

In the author's opinion, a combination of both cultures will be ideal. If one can succeed in getting a model, at least an approximate one, characterizing the source and the mechanism generating data, it may contribute to expansion of natural knowledge. One way of achieving this is to use a model, wherever it may come from, and validating it by algorithmic modeling. In this connection, it is interesting to note what Fisher (1956, 1960, Sec.21.1 at the end of Chapter 3) said about nonparametric tests which Fisher himself introduced round about 1935:

> *The utility of such nonparametric tests consists in their being able to supply confirmation whenever, rightly or, more often, wrongly it is suspected that the simpler tests have been appreciably injured by departures from normality.*

Lack of firm basis for measurement of uncertainty. There are various methods of expressing uncertainty such as the variance of estimators conditional on a certain configuration of the sample, confidence limits, fiducial limits, plausible limits based on the likelihood function or posterior probability and so on, which are all subject to debate.

Lack of methodology for distinguishing between random noise and chaos. The reader is referred to examples given in Rao (1997b, pp.26-28).

Methodology based on asymptotics. Current contributions to statistical theory are based on asymptotic behavior of estimators and tests (as the sample size tends to infinity) without examination of their usefulness in small samples.

2.2 Limitations of statisticians

In the early days of the development of statistics as a method of extracting information from data and taking decisions, research in statistics was motivated by practical problems in biological and natural sciences, as indicated in Table 2. Methods developed for use in one area found applications in other areas with minor modifications. Gradually, statistics came to be adopted as an inevitable instrument in all investigations scientific or otherwise as discussed in Section 1 of this paper. Then the need arose for training professionals in statistics to help the government and research organizations in the collection and analysis of data. Statistics was introduced as a compulsory subject in the curriculum of courses in some scientific and technological disciplines.

Gradually, universities started separate departments of statistics where statistical theory and methodology is taught without any serious focus on applications. Venues of interaction between faculty members in statistics and other departments have gradually closed, and the lack of contact with live problems has impeded the expansion of statistics in desired directions or sharpening of the existing tools.

Students graduating in statistics learn statistics as a set of rigid rules without acquiring any knowledge of their applications to practical problems. The students are not made aware that statistics is a dynamic and evolving discipline and fertile research in statistics can result only by collaborative work with researchers in other sciences.

Statistics Departments in the universities generally tend to produce statisticians as a separate breed of scientists, which is detrimental to their usefulness as professionals helping research workers in natural and social sciences in data collection and its analysis. They teach statistics as a deductive discipline of deriving consequences from given premises. *The need for examining the premises, which is important for practical applications of results of data*

analysis, is seldom emphasized.

It is also surprising that in many universities courses in design of experiments and sample surveys are not given, or listed as optional. Knowledge of these two methodological aspects of data collection is extremely important in all investigations.

Further, students specializing in statistics do not acquire in-depth knowledge of any basic discipline and are therefore unable to collaborate with scientists in research work. There has been some thinking on the education and training of statisticians, but no attempts have been made to change the present system (see Kettenring (1995), Parzen (1997), Rao (1997a) and other references in these papers.)

It is also relevant to add here what Fisher (1938) said on who should teach statistics.

> *I want to insist on the important moral that the responsibility for the teaching of statistical methods in our universities must be entrusted, certainly to highly trained mathematicians, but only to such mathematicians as have had sufficiently prolonged experience of practical research and of responsibility for drawing conclusions from actual data, upon which practical action is to be taken. Mathematical acuteness is not enough.*

This is generally disregarded in the recruitment of faculty members of statistics departments at the universities.

Regarding research work in statistics published in journals, S.C. Pearce says:

> *In many fields of statistics numerous techniques have been published with little to guide the practical man as to their spheres of influence.*

Current research in statistics should be directed to and made available for immediate use in problems waiting to be solved "rather than getting published in archival journals", as the editors of the newly started journal *Biostatistics* put it.

2.3 Needs of customers

Who needs statistics? The scientists use statistics in a marginal way. Current technology enables scientists to make measurements with a high degree of accuracy and generate large amounts of data under identical conditions. In such a situation, it *may* not be necessary to use sophisticated methods of data

analysis. There appears to be no substantial evidence in scientific literature of any major discovery being directly attributed to results or insight provided by statistical analysis. Let us look at the following quotations.

If your experiment needs statistics, you ought to have done a better experiment.

- Lord Rutherford (1871-1931)

A theory can be proved by an experiment but no path leads from experiment to theory.

- A. Einstein (1879-1955)

It is safe to say that no discovery of some importance would have been missed by lack of statistical knowledge.

- F.N. David (1909-1993)

All these statements do not imply that observational data cannot provide clues to scientific discovery. Perhaps, lack of interest in using statistical methods in scientific research may be due to the limited role of hypothesis testing as formulated by statisticians in knowledge discovery. The aim of statistical analysis should be not only to *answer specific questions* but also to *raise new questions* and indicate what further investigations are needed to answer them.

Perhaps, the greatest beneficiaries of statistics are the national governments (responsible for socio-economic development, optimum utilization of national resources, protecting the environment and providing essential public services), industry (in maintaining quality of manufactured goods, increasing productivity) and business (in efficient management and working out optimal strategies). Is the current statistical methodology adequate to meet the demands of customers in these areas?

With computerization of all activities in science, commerce and government, we will have access to unprecedented quantity and variety of data. We also have enormous computing power. These provide us an opportunity to meet the customer's demands for timely and useful information on a wide variety of issues.

There is need to develop new statistical methods for managing large data sets, on line automatic processing of data (OLAP) to judge the performance of existing practices (working hypotheses), extracting *new* information useful to customers rather than to answer specific questions, decision making and assessing the risks involved, and making automatic adjustments for missing or contaminated data. The limitations of the current statistical methods

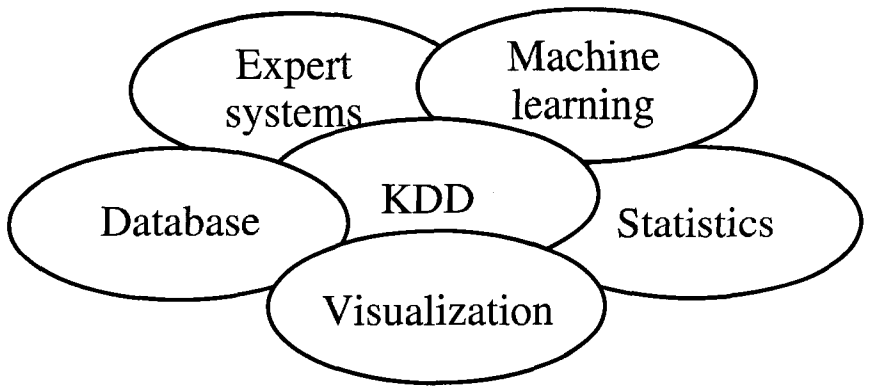


Figure 2. KDD: Knowledge Discovery in Database

in handling large data sets for extracting useful information have led computer scientists, engineers and operations research workers to suggest what is claimed to be a different approach to data analysis called *Data Mining* much to the surprise of statisticians.

3 Data mining

3.1 What is data mining

Is Data Mining (DM) a form of statistics or a revolutionary concept? Adriaans and Zantinge (1996, p.5) describe DM or a more general concept known as KDD (knowledge discovery in databases) as

> *the non-trivial extraction of implicit, previously unknown and potentially useful knowledge from data.*

It is conceived as a multidisciplinary field of research involving machine learning, database technology, statistics, expert systems and visualization (see Figure 2).

Some statisticians think that the concepts and methods of DM have their basis in statistics or already subsumed under current statistical methodology. We shall review the current literature on DM, examine to what extent they meet the needs of customers compared to the available statistical methodology, and comment on possible developments in the future.

3.2 Massive data sets

What motivated and made DM popular is the availability of large data sets which are automatically generated, stored and easily retrievable for analysis. They are high dimensional in terms of features, cases and classes. The stochastic model for the observations is generally not fully known. There may be some missing values and contaminated data. (See Table 4 for examples of such data sets). Generally, data relating to business transactions, services provided by the government and even scientific programs like the genome mapping and sky surveys in astronomy run into multi-gigabytes.

Conventional statistical methods of testing of hypotheses and building models for prediction may not be suitable. *Every conceivable hypothesis or model is bound to be rejected when a large data set is available.* Even the computation of test statistics and estimates of parameters such as the sample median may pose difficulties. What can we do when large data sets are available?

The characteristics of a large sample are, by asymptotic consistency theorems, close to that of the population on which observations are made. As such, inferences drawn from a sample will have a low degree of uncertainty. Further, the amount of uncertainty itself can be estimated with a high degree of precision by double cross validation (revalidation) as explained in Section 3.4 without any model assumptions, which cannot be achieved with small data sets.

3.3 Data mining versus traditional data base queries

Using traditional data analytic methods, we can estimate certain parameters of interest and examine the performance of certain decisions (or hypotheses) formulated on the basis of previous studies or some theoretical considerations. Such an analysis is often called on-line analytical processing (OLAP) or providing answers to certain queries.

In DM, through the use of specific algorithms or *search engines* as they are called, attempts are made to discover *previously unknown patterns and trends* of interest in the data and take decisions based on them. We shall examine some of the methods reported in the literature on DM which is described by Wegman (1998) as

> *"exploratory data analysis with little or no human intervention using computationally feasible techniques, i.e., the attempt to find interesting structures unknown a priori."*

Table 4. The number and type of features, classes and cases used for training, cross validation (Test 1) and revalidation (Test 2) in seven data sets.

	Cases			Features		
Dataset	Train	Test1	Test2	Num.	Type	Classes
Medical	2079	501	522	33	Num+Binary	2
Telecom	62414	34922	34592	23	Num+Binary	2
Media	7133	3512	3672	87	Num	2
Control	2061	685	685	22	Num	Real
Sales	10779	3591	6156	127	Num+Binary	3
Service	4826	2409	2412	215	Binary	2
Noise	20000	5000	5000	100	Num	2

3.4 Cross validation and revalidation

When a large data set, say with S cases, is available, we can divide it into subsets with S_1 and S_2 cases which are also sufficiently large. We can use the subset S_1 to formulate a certain decision rule R based on the discovery of patterns through a *search engine*. The second set S_2 can be used to evaluate the performance of R through some loss function. In view of the largeness of S_2, we expect to get a precise estimate of the average loss. This procedure known as cross validation is well known in statistical literature, but its application in small samples through methods such as LOO (leave one out) may not be effective.

There are other possibilities when a large sample is available, especially when the *search engine* suggests several possible rules $R_1, R_2, \ldots$ based on the subset S_1 of cases. We then divide S_2 into two subsets S_{21} and S_{22}, and use cross validation on rules $R_1, R_2, \ldots$ on S_{21} and choose the rule R^* with the minimum loss. Now, we can compute the loss in applying R^* on the second subset S_{22}. We thus have an unbiased estimate of loss in using the rule R^*. This method may be described as *revalidation*. (See Table 4 where different divisions of the available cases as Train (S_1), Test 1 (S_{21}) and Test 2 (S_{22}) are given in some real large data sets.).

As new data come in, we have a chance to evaluate the performance of rules in current practice and update if necessary.

3.5 Data mining techniques and algorithms

Visualization

The use of graphs in exploratory data analysis (for understanding the nature of observations and choosing an appropriate model), and in reporting the results of statistical analysis is well known in statistical literature. (See Fisher (1967, Chapter 2), Tukey (1977).) With increase in computing power and possibilities of viewing high dimensional data through parallel coordinates (Wegman (1990), Wegman and Luo (1997), Wilhelm, Symanzik and Wegman (1999)), projections in different directions (Friedman and Tukey (1974)), and data reduction by canonical coordinates (Rao (1948a)), principal components (Rao (1964)), correspondence analysis (Benzēcri (1992) and Rao (1995)), and multidimensional scaling (Kruskal and Wish (1978)), graphical analysis is becoming a valuable tool in discovering patterns in data.

3.6 Finding associations

A typical problem is that of finding association between items purchased by customers in a grocery shop (e.g., those who purchase bread also buy butter). In the abstract, the problem may be stated as follows. We have a set of vectors with zeros and ones such as (10010...), where 1 denotes the presence of a specific characteristic (such as purchase of an item) and 0 otherwise. The object is to find whether there is a high percentage of vectors with all 1's in certain positions. A fast algorithm for this purpose was developed by Agrawal, Imielinski and Swami (1993).

3.7 Clustering, pattern recognition and decision trees

These methods first introduced in statistical literature and developed by computer scientists and engineers for specific purposes are extensively used in data mining.

3.8 Machine learning, neural networks and genetic algorithms

Suppose the problem is that of predicting a target (or class) variable y using a concomitant vector variable x (called features). In statistics, we generally start with a probability model for the variables (x, y) and estimate the conditional distribution of y given x, on the basis of observed samples $(x_1, y_1), \ldots, (x_n, y_n)$. We can then use the conditional distribution of y given x to predict y. In machine learning, we do not explicitly use any probability model. We use an algorithm to find a function $f(\cdot)$ such that

$$\sum_{i=1}^{m} \phi[y_i - f(x_i)]$$

is minimized, where ϕ is a given loss function and $m(< n)$ is the number of samples set apart for learning. This is done by specifying a wide class of functions for f and using a search method like neural networks or genetic algorithms. The efficiency of an estimated function $\hat{f}$ is judged by cross validation, i.e., applying it on the remaining $(n - m)$ samples and computing the average loss

$$(n-m)^{-1} \sum_{m+1}^{n} \phi[y_i - \hat{f}(x_i)].$$

If the computed loss is large, we alter the class of functions f and search for an optimal solution. The final solution is obtained by a series of iterations.

4 Some final thoughts

We view this pile of data as an asset to be learned from. The bigger the pile, the better - if you have the tools to analyze it, to synthesize it and make yourself more and more creative.

- Britt Mayo
Director of Information Technology

Statistics is a broad based scientific discipline with theory and methods developed through the calculus of probability for taking optimal decisions under uncertainty. During the last century, research in statistics was directed to the concepts laid down by Fisher, Neyman and Wald. As pointed out in Section 1 of this paper, there are difficulties in formulating the problems to be solved and in applying these concepts to practical problems. [There has been an uncharitable criticism that statisticians are providing exact solutions to the wrong problems, where as in practice, what is needed is an approximate solution to the right problem]. Current statistical methodology has no satisfactory rules governing the choice of the inputs needed such as data modeling, prior probabilities, and expression of uncertainty in decision making. Data mining methods, applied on large datasets, seem to bypass stochastic considerations, and derive decision rules using "machine learning" methods and

evaluate their performance through cross validation. The techniques used in data mining problems such as pattern recognition, decision trees, clustering and cross validation have their roots in statistics, but perhaps not actively pursued by statisticians. We may agree with what Weiss and Indurkhya (1998) say:

> *Statistical models are competitive with those developed by computer scientists and may overlap in concept. Still, classical statistics may be saddled with a timidity that is not up to the speed of modern computers.*

In conclusion, I believe DM is a form of much needed statistics neglected by statisticians.

REFERENCES

Adriaans, P. and Zantinge, D. (1996). *Data Mining.* Addison Wesley.

Agrawal, R., Imielinski, T. and Swami, A. (1993). Mining association rules between sets of items in large databases. *Proc. Int. Conf.*, ASMSIGMOD. Washington, D.C., 207-216.

Barnard, G.A. (1996). Scientific practice and statistical inference. *Symposium on the Foundations of Statistical Inference, with Special Emphasis on Applications in honor of D.A. Sprott*, 1-9.

Benzēcri, J.P. (1992). *Correspondence Analysis Handbook.* New York: Marcel Dekker Inc.

Berger, J.O. (2002). Bayesian analysis: A look at today and thoughts of tomorrow. In *Statistics in 21st Century* (Eds: Raftery, Tanner and Wells). Chapman and Hall/CRC, 275-291.

Box, G.E.P. (1980). Comment (on A Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Industry). *The American Statistician* 34, 65-80.

Breiman, L. (2001). Statistical Modeling: Two cultures. *Statistical Science* 16, 199-231.

Buchanan-Wollaston, H.J. (1935). Statistical Tests. *Nature* 136, 182-183.

Cox, D.R. (1990). Role of models in statistical analysis. *Statistical Science* 5, 169-174.

Cox, D.R. (1997). The current position of statistics: A personal view. *The International Statistical Review* 65, 261-276.

Cox, D.R. (2000). The five faces of Bayesian statistics. *Calcutta Statistical Bulletin* 50, 127-136.

Efron, B. (1979). Bootstrap methods: Another look at jackknife. *Ann. Stat.* 7, 1-26.

Fisher, R.A. (1922). On the mathematical foundations of theoretical statistics. *Philos. Trans. Roy. Soc.* A222, 309-368.

Fisher, R.A. (1934). The effect of methods of ascertainment upon the estimation of frequencies. *Ann. Eugenics* 6, 13-25.

Fisher, R.A. (1935, 1960, 7th edition). *The Design of Experiment.* Edinburgh: Oliver and Boyd.

Fisher, R.A. (1936). Has Mendel's work been rediscovered? *Annals of Science* 1, 115-137.

Fisher, R.A. (1948). The Rhesus factor: A study in scientific research. *Am. Sc.* 35, 95-102, 113.

Fisher, R.A. (1953). The expansion of statistics. *J. Roy. Statist. Soc.* A 116, 1-6.

Fisher, R.A. (1956). *Statistical Methods and Scientific Inference.* Edinburgh: Oliver and Boyd.

Fisher, R.A. (1967). *Statistical Methods for Research Workers.* Hafner Publishing Company.

Friedman, J. and Tukey, J.W. (1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers* 23, 881-889.

Ghosh, J. (1988). *Statistical Information and Likelihood: A Collection of Critical Essays by D. Basu.* Lecture Notes in Statistics 45, Springer Verlag.

Inman, H.F. (1994). Karl Pearson and R.A. Fisher on statistical tests: A 1935 exchange from *Nature*. *The American Statistician* 48, 2-11.

Kettenring, J. (1995). What industry needs. *The American Statistician* 49, 2-4.

Kruskal, J.B. and Wish, M. (1978). *Multidimensional Scaling.* Sage Publications.

Lehmann, E.L. (1990). Model specification: The views of Fisher and Neyman and later developments. *Statistical Science* 5, 160-168.

Lindley, D.V. (1957). A statistical paradox. *Biometrika* 44, 187-192.

Neyman, J. and Pearson, E.S. (1928). On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika* 20A, 175-240 and 263-294.

Neyman, J. and Pearson, E.S. (1933). On the problem of most efficient tests of statistical hypotheses. *Philos. Trans. Roy. Soc.* A 231, 289-337.

Parzen, E. (1997). Data mining, statistical methods mining and history of statistics. Tech. Rept. Texas A& M University.

Pitman, E.J.G. (1937). Significance tests which may be applied to samples from any population. *J. Roy. Statist. Soc.* (supp), 119-130, 225-232 and *Biometrika* 29, 322-335.

Rao, C.R. (1948a). The utilization of multiple measurements in problems of biological classification. *J. Roy. Statist. Soc.* B 9, 128-140.

Rao, C.R. (1948b). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proc. Camb. Philos. Soc.* 44, 50-57.

Rao, C.R. (1964). The use and interpretation of principal components analysis in applied research. *Sankhyā* A 26, 329-358.

Rao, C.R. (1971). Some aspects of statistical inference in problems of sampling from finite populations. In *Foundations of Statistical Inference*, Godambe V.P. and Sprott D.A. (eds), 177-202. Toronto, Canada: Holt, Rinehart and Winston.

Rao, C.R. (1977). Prediction of future observations with special reference to linear models. *J. Multivariate Analysis* 4, 193-208.

Rao, C.R. (1987). Prediction of future observations in growth curve type models (with discussion). *Statistical Science* 2, 434-471.

Rao, C.R. (1995). A review of canonical coordinates and an alternative to correspondence analysis using Hellinger distance. *Qüestiiō* 9, 23-63.

Rao, C.R. (1997a). A cross disciplinary approach to teaching of statistics. *Proc. 51st session of the Int. Statist. Institute* Book 1, 441-444, Istanbul.

Rao, C.R. (1997b). *Statistics and Truth: Putting Chance to Work* (Second Edition). Singapore: World Scientific.

Rao, C.R. and Wu, Y. (2001). On model selection. *IMS Lecture Notes - Monograph Series* 38, 1-64.

Reid, N. (2002). Likelihood. In *Statistics in 21st Century* (Eds: Raftery, Tanner and Wells). Chapman and Hall/CRC, 419-431.

Smith, T.M.F. (1997). Social surveys and social science. *Canada J. Statist.* 25, 23-44.

Tukey, J.W. (1993). Major changes for multiple-response (and multiple adjustment) analysis. In *Multivariate Analysis: Future Directions* (Ed. C.R. Rao). North Holland, 401-422.

Tukey, J.W. (1977). *Exploratory Data Analysis.* Reading, Mass: Addison-Wesley.

Van den Berg, G. (1992). *Choosing an Analysis Method: An Empirical Study of Statistician's Ideas in View of the Design of Computerized Support.* Ph.D. Thesis, University of Leiden.

Wald, A. (1941). Asymptotically most powerful tests of statistical hypotheses. *Ann. Math. Statist.* 12, 1-19.

Wald, A. (1947). *Sequential Analysis.* New York: John Wiley.
Wald, A. (1950). *Statistical Decision Functions.* New York: John Wiley.
Weiss, S. M. and Indurkhya, N. (1998). *Predictive Data Mining: A Practical Guide.* Morgan Kaufmann Publishers, Inc.
Wegman, E. J. (1990). Hyperdimensional data analysis using parallel coordinates. *J. Am. Statist. Assoc.* 68, 664-675.
Wegman, E. J. (1998). Visions: The evolution of statistics. Keynote talk at the Conference, New Techniques and Technologies in Statistics, Sorrento, Italy.
Wegman, E. J. and Luo, Q. (1997). High dimensional clustering using parallel coordinates and grand tour. *Computing Science and Statistics* 28, 352-360.
Wilhelm, A., Symanzik, J. and Wegman, E.J. (1999). Visual clustering and classification: The oronsay particle size data revisited. *Computational Statistics* 14, 109-146.
Wolfowitz, J. (1967). Remark on the theory of testing of hypotheses. *The New York Statistician* 18, 1-3.

Comment

CHALLENGES FACING STATISTICS FOR THE 21ST CENTURY

C.C. Heyde
Australian National University and Columbia University
Australia
E-mail: chris@wintermute.anu.edu.au

1. Introduction

I am pleased to have the opportunity to join Professor Rao in commenting on what future there may be for statistics. He has an unrivalled perspective on the core content of the discipline while my vision is principally from the probabilistic side of the subject. But we are very much in agreement on the many challenges facing statistics and that the warning signs should not be ignored.

Prof. Rao has chosen a title for his paper which suggests that he has looked deeply into his crystal ball to see the possible future form for statistics, although in the text he hardly extrapolates beyond the present. He starts

from a deep perspective of what statistics is, and what are its principal tools. He then goes on to sketch current limitations, firstly of the methods, which are mostly model oriented, and secondly of the limitations of statisticians themselves, as a consequence of a training which all too frequently omits key topics and any serious focus on applications. He questions the extent to which the needs of statistical customers are being met. Finally, he notes that statisticians have been less than wholehearted in their acceptance of the opportunities offered by the power of modern computers, and he flags data mining as a "form of much needed statistics neglected by statisticians".

These are timely warnings to which I will add my own perspective. We must be aware of the possible future if the statistical profession maintains its current attitudes and practices. Much of this discussion is an abridged version of the lecture Heyde (2002) given at a NSF workshop in Washington D.C. in May 2002.

2. The current situation

At a superficial look, statistics is thriving as a profession. The American Statistical Association (ASA) has around 17,500 members and 21 special interest sections. The Royal Statistical Society (RSS) has 6300 members including a professionally accredited membership of some 1500. The societies, national and international, have an ever expanding list of programs reflecting the increasing importance of applications. And there is no shortage of jobs for statisticians. In some areas, for example biostatistical jobs in the pharmaceutical industry, the demand for new recruits is considerable.

However, if one looks at membership over time, the societies are pretty much static, save for special cases such as the International Statistical Institute (ISI). The RSS has been declining slowly in recent times, for example dropping by 100 in 2000. Its last big jump in membership was in the 1970s when it grew by 20%. Since then, if one corrects for the merger with the Institute of Statisticians (IOS), it has been static or in decline. The ASA has increased its membership by a little over 10% in the last 20 years. Statistical societies elsewhere have had roughly comparable experiences. For example, the membership of The Statistical Society of Australia peaked at 735 in 1973; 5 years later it had slipped to 650 and it has since stayed around this level. Whether these membership numbers are in line with a general downturn in participation in community type organizations (e.g Putnam (2000)) is an open question. Certainly the societies need to maintain their vitality and integrity of purpose, and to avoid excessive fragmentation. And unfortunately many new graduates are not joining.

Ageing of the societies is another issue. For example in the case of the

International Statistical Institute (ISI), the median age of members in June 1996 was 55 and by December 2000 it had grown to 59! Somewhat more than 10% of newly elected members are resigning after 3 or 4 years of membership.

Table 5. INTERNATIONAL STATISTICAL INSTITUTE (ISI)

Members by Age	June 1996	Dec. 2000
Over 70	14.6%	17.2%
Over 60	34.2%	45.2%
Over 50	71.9%	83.5%
Below 50	28.1%	16.5%

It would be instructive to examine age profiles in a wide variety of settings to aid in the assessment of general health.

A useful measure of activity in statistics is provided by the number of papers and books listed in the annual *Current Index to Statistics* published by ASA and IMS. Details are given in the following table. It should be noted that there was rapid growth in the 1970s but that numbers have now levelled out to around 8500 - 9000/year. This is coming from a list of core journals, currently numbering around 110 and expanding at the rate of around 2/annum, roughly 400 non-core journals and some 1500+ edited books and proceedings. Some details are given in Table 6 below and the numbers do suggest a hollowing out of core statistics as the diverse applications literature expands.

Table 6. Papers listed in *Current Index to Statistics*

1976	4300
1978	7500
1980-89	82591
1990-98	85658
1999	8585

It is instructive to look at the case of Data Mining, a major growth area whose importance for statistics has been reviewed by Professor Rao. *Current Index to Statistics* first recorded the title in 1996, and then flagged the subject with a heading in 1997, in which year 17 papers were listed. This came

about through the creation of a new (international) journal *Data Mining and Knowledge Discovery* published by Kluwer, with 4 issues/year. As of late April 2002 the *Current Index to Statistics* database produced 93 articles in response to a search for data mining articles. By contrast, a look at Ingenta Search (http://www.ingenta.com) on the same day produced 1256 articles spanning the period 1995-2002. It is clear enough that Statistics is not embracing Data Mining even though a few statistics departments seem to be acknowledging that it represents a new market niche.

Rather similar comments can be made about Bioinformatics, although many more statisticians are taking a research interest in aspects of this subject.

Perceptive professionals can discern the trends before they show up clearly in the literature. So, by and large, can the students. Indeed, the striking recent growth in Mathematical Finance has been largely fueled by student demand. Many faculty who were drafted into teaching courses in the area have subsequently developed a corresponding research interest.

The slow growth of the literature masks a slow decline in submissions to the key journals. This is proceeding in tandem with the creation of new journals, often of a quite narrow specialization. To quote some examples, the annual submissions to *Annals of Statistics* (AOS) have been 345, 344, 283, 292 over 1997-2000 compared with stable levels of 10+ years back of around 375. Those for *Annals of Probability* (AOP) have slipped from 242 and 219 in 1997 and 1998 to 200 in 2000. Those for *Annals of Applied Probability* slipped from 170 in 2000 to 136 in 2001. The journals *Journal of Applied Probability* and *Advances in Applied Probability* had its submission numbers peak around 450 in the early 1990s and these have slipped to 330 in 2001.

Sales of the key journals are also slipping. From 1998 to 2000 the total subscriptions of AOP dropped 11% from 2287 to 2039, while those of AOS dropped 7% from 3619 to 3358. The drop in each case has been worse amongst members than non-members.

We must note this slippage against a background of there being 200+ new PhDs a year in the USA in "Math."Statistics and perhaps half that number elsewhere. It is clear that many, perhaps most, of these people are essentially being lost to the statistical profession as it is now constituted. They are ending up in kindred disciplines for which their research skills are very relevant.

3. The future of university statistics

It is my feeling that statistics may need to be recast into a broader disciplinary setting. The autonomous university statistics department, as we now know it, could easily disappear in the forseeable future, but at the same

time there will be an increase in the number and diversity of university programs that have a serious statistical component. The amount of statistical work being done by those who do not identify with the statistical profession is increasing much more rapidly than that being done by those who do identify with the profession. However, the integrity of the profession cannot be sustained by a diverse collection of applications in the absence of adequate attention to the methodological core of the subject and the continuing development of deep and broadly applicable results.

When I joined academia 40 years ago it was an expansionary time for the tertiary education sector. Statistics departments were new, or being newly created, across the world. Statistics was acknowledged as an important discipline and a necessary component of a well rounded education in any discipline with a quantitative basis. Service courses were generally taught by statisticians and this fueled the growth of the new departments. But over the intervening years the report card has not been so good. Failure to provide client departments with what they wanted and competitive pressures on departments to generate revenue from student numbers have resulted in the loss of many of these courses. Statistics departments have contracted as a consequence and this has meant, not uncommonly, loss of autonomy through amalgamation with another department, usually Mathematics. Some departments have escaped through, for example, fortuitous location, institutional prestige, or niche marketing.

All disciplines now find themselves in a market driven environment. Internationally there has been a substantial decline in student numbers in the Physical Sciences. Mathematical Sciences has been significantly affected, but not as much as Physics and Chemistry. Some universities have closed Physics departments (three in the UK); there is no guaranteed core any more. The first private university in Australia closed (sold) its whole Science Faculty. Statistics departments will need to make their choices in a very competitive environment.

I would like to think that statistics can retain its traditional strengths and add to them with optional courses, differing from institution to institution, in reflection of important application areas and interests and expertise of the faculty. The universities have generally responded well to the recently emergent market niche in quantitative finance. We can, in principle, respond similarly to other market opportunities. Admittedly this is not painless against a background where staff numbers are often declining, for one has to take the risks, and offer the new courses, before the student numbers grow, and the payoff can be harvested. Adjuncts and other departments need to be enlisted as part of the process.

4. Historical roots and possible transformation

Our tradition is largely that of inference based on probabilistic models, as Professor Rao has noted. This places us squarely in the mathematical sciences. We have never sought a broad role in the information sciences, and even less in information technology. Indeed, computing has been a rather incidental statistical tool. But now much data is automatically recorded. There is no shortage of contexts where there are no real limitations of sample size. If one is studying network traffic models then 50,000 observations are routinely available. It is inevitable that fields such as Data Base Management have evolved. By default probably, rather than design, the statistical profession did not seek to embrace such things. The rationalizations are clear, the subject is fundamentally non-stochastic, it is technical rather than scientific etc etc. This attitude has been marginalizing the statistical profession in the marketplace, both commercial (and nowdays as a consequence) academic. There is no doubt in my mind that we are becoming less influential and that other data oriented sciences are taking customers, students and jobs away from us.

I see this very clearly at Columbia University where my principal role is that of Director of the Center for Applied Probability (CAP), a cross-campus, cross-disciplinary enterprise concerned with chance and risk in all their manifestations. Some 25 faculty across 10 university departments belong to this center, and I find that my closest colleagues there are based in disciplies outside Mathematics and Statistics, via Engineering and Business Schools. They do probability and statistics under other auspices and they have no real identification with the statistical profession, even though they have often been trained through mathematics or statistics departments. People like this can help to provide vital links to other kindred disciplines. On the whole, these links are tenuous. Half of the CAP postdocs have come from Mathematics and Statistics departments, have rebadged themselves into Quantitative Finance, and are likely to be lost to Statistics.

The term statistics has negative social connotations. So there may be a case for trying to reinvent and reinvigorate the subject under a new name which has a better image, and is suggestive of a broader spectrum of activity than hitherto. I do not have a really good name to suggest. Statistical Sciences seems better than Statistics. We need to embrace uncertainty, risk and chance as well as statistics. A new word in the spirit of "Stochastic" might be created. I think "data" should be kept out of the title. And what we do *is* a science rather than a technology. One difficulty is that for any prospective title there are likely to be other players who want to defend their perceived territory, and perhaps partnerships or amalgamations are the only way forward. There are certainly precedents, such as the merger of the Operations Research Society

of America and The Institute of Management Science in 1994/5.

REFERENCES

Heyde, C.C. (2002), Statistics in the international scene, Lecture to NSF Workshop *Statistics: Challenges and Opportunities for the 21st Century*, Wasington, D.C., May 6-8, 2002. (Available in .pdf form from the author via the e-mail address cch18@columbia.edu.)

Putnam, R.D. (2000), *Bowling Alone: The Collapse and Revival of American Community*, New York: Simon & Schuster.

Rejoinder

I want to thank Professor Heyde for his comments on my paper. He expressed some concern about the slow rate of growth of statistical societies (decline in some) in terms of membership, the ageing of societies, the slow rate of growth of literature in statistics judged by submission of papers to the core journals, and a general decline of student numbers in statistics. [In many advanced countries, the ratio of local to foreign students studying statistics in the universities is declining.]

The title of my paper may sound pessimistic. It was chosen to highlight what I see as a crisis in statistics, the way it is taught in the universities more as exercises in mathematics than demonstrating its usefulness in solving practical problems, and the way it is practiced by consultants using routine statistical methods, depending on available software, to answer specific questions rather than cross examining data to solicit whatever information is available on the problem under study.

I hope the reflections on the past and visions for the future as discussed in my paper and professor Heyde's comments, " we are becoming less influential and other data oriented sciences are taking customers, students and jobs away from us", will promote discussions at national and international conferences on the future of statistics and the way it should be developed to retain its status as a separate discipline of fundamental importance. [A Dean of the Pennsylvania University while addressing the statistical faculty remarked that if there is a budget cut, statistics will be one of the departments to be closed down.]

I would like to take this opportunity to provide some additional references expressing similar concerns.

Under Kolmogorov, Fisher and Neyman set up, specification of the probability model for given data is essential for data analysis. They have not, however, provided guidelines for the choice of an appropriate model. [A large number of papers are published in well known journals offering solutions to problems using unjustifiable or unrealistic models simply because they lead to easy solutions.] It is quite possible that the methodology based on some standard models which are easy to work with are somewhat robust against possible alternatives. It was believed that the inference based on the normal probability model will be valid for small to moderate departures from normality. It was shown by Tukey, Huber and others that even **a** slight departure from normality and contamination in data may lead to wrong analysis. However, there are widely different opinions ranging from completely discarding models [Breiman (2001) with the discussion], emphasizing the distinct roles of stochastic models in data analysis [Cox (1990), Lehmann (1990)], to the use of imprecise probability models [Walley (2002), Zadeh (2002)]. I believe that modeling should be an integral part of data analysis, and not something chosen a priori. Successful modeling might throw a new light on the problem. One should also recognize situations where model free analysis is called for. It may be of interest in this connection to refer to Tukey (1980), where he talked about exploratory and confirmatory data analysis and to Box and Draper (1987) who refer to two types of modeling, empirical (or interpolatory) used as a guide to action, often based on forecasts of what to expect from future observations, and explanatory (or mechanistic) involving the search for the basic mechanism underlying the process being studied leading to a fundamental understanding of the problem. [Do we have codified statistical methodology for explanatory modeling?]

There is some confusion between scientific and statistical methods. A recent paper by Mackay and Oldford (2000) clarifies the situation. They describe the statistical method as represented by five stages: Problem, Plan, Data, Analysis, Conclusion (PPDAC). This is a broadly acceptable description, but the questions to be addressed at each stage have to be properly formulated depending on the problem, on which there may be differences of opinion. [From where do the problems arise? The source is from other fields, but not statistics]. Analysis should not be confined to just tests of significance at a conventional level of significance and estimation of parameters as taught in the statistical courses. It should be aimed at extracting all the information contained in the data about the problem.

I have mentioned in my paper the need for statisticians to play a ma-

jor role in what has come to be known as data mining which is essentially statistical analysis of large data sets which are becoming available in commercial transactions, human genome studies and web pages consisting of textual and multimedia information that are used by millions of people. We have to evolve new statistical methodology of large data sets taking advantage of the available computing power. [See Table 4 in the paper]. Otherwise we may lose a large number of customers to computer scientists and engineers who are already actively engaged in data mining.

REFERENCES

Breiman, Leo (2001). Statistical modeling: the two cultures, *Statistical Science*, 16, 199-231.

Box, G.E.P. and Draper, N. (1987). Empirical Model - *Building and Response Surfaces*, Wiley, New York.

Cox, D.R. (1990). Models in statistical analysis, *Statistical Science*, 5, 160-168.

Lehmann, E.L. (1990). Model specification, *Statistical Science*, 5, 169-174.

Mackay, R.J. and Oldford, R.W. (2000). Scientific method, statistical method and the speed of light, *Statistical Science*, 15, 254-278.

Tukey, J. (1980). We need both exploratory and confirmatory, *Amer. Statist.*, 34, 23-25.

Walley, P. (2002). Reconciling frenquentist properties with the likelihood principle, *J. Statistical Planning and Inference*, 105, 35-66.

Zadeh, L.A. (2002). Toward a perception-based theory of probability reasoning with imprecise probabilities, *J. Statistical Planning and Inference*, 105, 233-264.

USING CONTINGENCY TABLES TO CONSTRUCT BOTH TRADITIONAL AND NEW NONPARAMETRIC TESTS

J.C.W. RAYNER AND D.J. BEST
School of Mathematics and Applied Statistics
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: john_rayner@uow.edu.au, jbest@uow.edu.au

A generic approach to the comparison of treatments in many experimental designs is to exhibit the data in a contingency table that is at least partially ordered, construct a model for the table of counts, and derive an omnibus test statistic and its components. For several important experimental designs the components are asymptotically chisquared distributed and assess moment differences between treatments. Often the initial components are the basis for well-known nonparametric rank tests such as the Kruskal-Wallis and Spearman tests. Further decompositions of the components to obtain LSD comparisons are easily derived. We usually recommend combining the higher moment components into a residual. The approach is demonstrated for three designs.

Keywords: balanced incomplete block design, completely randomised block design, orthogonal polynomials, partition, randomised block design.

1 Introduction

For many experimental designs the data many be entered into a contingency table that is at least partially ordered. The objective may be to simply compare treatment means, but nevertheless it will often be possible to formulate a general model and derive an omnibus test to compare treatments. In the designs we consider below, when there are no ties the test statistic is a simple multiple of Pearson's chisquared statistic. The omnibus statistic is then decomposed into components that are asymptotically independent and asymptotically have the chisquared distribution. These components assess the presence or abscence of location, dispersion, and higher univariate moment differences between treatments for singly ordered data, and of bivariate moment differences between treatments for doubly ordered data.

The distributional assumptions for the models proposed are minimal, and some of the lower order tests based on the components are well-known rank tests such as the Wilcoxon, Kruskal-Wallis, Friedman, Durbin and Spearman tests. The tests based on the higher order components are extensions of these

tests and are often new to the literature. Their convenient null distributions mean they are simple to use, and they are easy to interpret. Collectively they provide a depth of insight into the statistical analysis not previously possible. This general approach allows the development of new nonparametric tests, in particular for many of the more complicated experimental designs for which nonparametric analogues of the traditional parametric test statistics have not yet been developed.

In the following we focus on data analysis for three designs: the completely randomised design, the randomised block design, and a balanced incomplete block design. Details of the derivation of the analyses may be found in Rayner and Best (2001). A new book focusing on the analysis of sensory data, Rayner and Best (2002), is in preparation.

2 Overview of the analysis derivation

The analysis for the designs presented here assumes that it is possible to construct a partially or completely ordered contingency table of counts, $\{N_{ij}\}$ say. For such a table Pearson's X_P^2 may be expressed as a quadratic form, the matrix of which may be diagonalised to ultimately give $X_P^2 = C_1 + C_2 + ... + C_d$, where the number of components, d, depends on the design. This arithmetic decomposition makes no distributional assumptions, and has no distributional consequences. Typically the decomposition utilises the orthonormal functions $\{g_r(x_j)\}$ on the marginal proportions from the table, where $\{x_j\}$ are the *scores* assigned to the marginal classification. Common scores used are the *natural* scores $x_j = j$ for all j, and the midrank scores. Using orthonormal functions results in components that are asymptotically independent.

For example, for the completely randomised design with no ties, the data may be initially organised as ordered observations corresponding to each treatment, and then aggregated into a table of treatments and observations, with the jth ranked observation being in the (i, j)th position if it is an observation of treatment i. All other elements of this column are zero. From this table it is routine to construct a table of counts $\{N_{i,j}\}$, with N_{ij} being the number of times treatment i receives rank j. When there are no ties the column totals will all be one, and the row totals will be the number of times each treatment is observed. The column marginal proportions are all $\frac{1}{n}$, where n is the total numbers of observations. In this situation the well-known orthogonal polynomials for the discrete uniform-distribution $\{\frac{1}{n}\}$ may be used.

In such situations it follows that *constant* $*X_P^2 = Y^T Y$, where by the central limit theorem, Y is asymptotically $d-$variate normal with mean zero and unit covariance matrix. For the singly ordered designs the elements of Y

are of the form

$$constant * \sum_j N_{ij} g_r(x_j) / \sqrt{\sum_j N_{ij}}.$$

These are asymptotically standard normal, and the components are sums of squares of such terms.

For example, for the *balanced incomplete block design* when there are no ties, we assume that n judges or consumers rank k out of t products with each product being ranked r times. As usual $rt = nk$ as both quantities are the total number of data points. Put R_i = sum of the ranks for product $i, S_i = \sum_{j=1}^{k} j^2 d_{ij}$, in which d_{ij} is the number of times product i receives rank j. Durbin's statistic for testing equality of the product rank sums can be defined by

$$C_1 = \frac{12(t-1)\sum_{i=1}^{t} R_i^2}{rt(k^2-1)} - \frac{3r(t-1)(k+1)}{(k-1)}.$$

An accompanying statistic, which examines dispersion or quadratic differences between products, is

$$C_2 = \frac{180(t-1)\sum_{i=1}^{t}[S_i - (k+1)R_i + r(k+1)(k+2)/6]^2}{rt(k^2-1)(k^2-4)}.$$

The statistic C_2 is a new statistic not considered previously. Durbin's statistic is known to be approximately distributed as χ^2_{t-1}. The C_2 statistic has the same approximate distribution. Schach (1979, p. 43) defined an omnibus statistic, A say, which is defined in terms of the table of counts $\{d_{ij}\}$. Schach (1979) showed that

$$A = \frac{t-1}{t} X_P^2 = \frac{k(t-1)}{rt} \sum_{i=1}^{t} \sum_{j=1}^{k} (d_{ij} - \frac{r}{k})^2$$

has an approximate $\chi^2_{(t-1)(k-1)}$ distribution. It can be shown that C_1 and C_2 are components of A and that residual $A - C_1 - C_2$ has an approximate

$\chi^2_{(t-1)(k-3)}$ distribution. If $A - C_1 - C_2$ is large, then it may be important to look for cubic and higher order effects. Observe that if $k = t$ (each judge ranks every product) then C_1 becomes Friedman's rank test and A becomes the Anderson statistic for the randomised block design, to be discussed in section 4. If $k = 2$, then $A = C_1$ and C_2 is not available: it can be taken to be zero. We now show how to decompose C_1 and C_2.

Provided k is appropriately large we can define the polynomials $g_0(j), g_1(j)$ and $g_2(j)$, orthonormal on the discrete uniform distribution, by

$$g_0(j) = 1$$

$$g_1(j) = \sqrt{\left(\frac{12}{n^2-1}\right)}\left(j - \frac{n+1}{2}\right) \text{ and}$$

$$g_2(j) = \sqrt{\left\{\frac{180}{(n^2-1)(n^2-4)}\right\}}\left\{\left(j - \frac{n+1}{2}\right)^2 - \frac{n^2-1}{12}\right\}$$

each for $j = 1, \ldots, n$. If $r = nk/t$, the linear or location effect for the ith product, C_{1i} say, is defined as

$$C_{1i} = \sqrt{\left(\frac{t-1}{rt}\right)} \sum_{j=1}^{k} N_{ij} g_1(j),$$

and the quadratic or dispersion effect for the ith product is given by

$$C_{2i} = \sqrt{\left(\frac{t-1}{rt}\right)} \sum_{j=1}^{k} N_{ij} g_2(j).$$

Similarly the sth moment effect for the ith product can be defined by

$$C_{si} = \sqrt{\left(\frac{t-1}{rt}\right)} \sum_{j=1}^{k} N_{ij} g_s(j).$$

Then

$$A = \sum C_{1i}^2 + \sum C_{2i}^2 + \sum C_{3i}^2 + \text{ higher order terms.}$$

Table 1. Consultancies by ranks table.

	Rank									
Consultancy	1	2	3	4	5	6	7	8	9	10
F1	1	1	1	1	0	0	0	0	0	0
F2	0	0	0	0	1	0	0	1	1	0
F3	0	0	0	0	0	1	1	0	0	1

Typically we may isolate the first one, two or sometimes three terms in A and combine the rest into a residual. This approach is demonstrated in subsequent sections.

3 The completely randomised design

When introducing students to the Wilcoxon two-sample and Kruskal-Wallis tests, one approach is to order the observations for each treatment, and combine these to form a treatments by ranks table as in the example below.

Tasters Example. O'Mahony (1986, section 16.14, p. 341) cites an example where ten tasters are trained by three different consultancy firms, F1, F2, and F3. All the firms claim to have a superior psychological method for the sensory evaluation of flavour intensity. The ten tasters were tested for their flavour intensity measurement skills and then ranked, with 1 being best and 10 worst. The data, in the form of a treatments (consultancies) by ranks table, are given in Table 1.

On the basis of a Kruskal-Wallis test, O'Mahony (1986, section 16.14, p. 341) claims the consultancy means are significantly different at the 1% level, and that the sample sizes are too small to assess exactly which programmes are different from which via multiple comparison techniques. A 'by eye' inspection suggests that programmes F2 and F3 (average ranks 22/3 and 23/3) are equivalent and inferior to F1 (average rank 2.5).

A one-way ANOVA on the ranks has p-value 0.0103, while an LSD comparison of the means agrees with the suggestion above, that programme F1 is superior. This may seem to be a strange choice of analysis for these data, but we give it as the parametric competitor to the 'standard' nonparametric analysis.

A deeper nonparametric analysis may be based on $\left(\frac{n-1}{n}\right) X_P^2$ and its components. A feature of this analysis is that the first component is the Kruskal-Wallis statistic. See Table 2. Each column in Table 1 has a sum of one. The reader can quickly verify that $X_P^2 = 20$. In fact X_P^2 is independent of the

Table 2. Partition of $\left(\frac{n-1}{n}\right) X_P^2$

Source	df	SS	p-value
Kruskal-Wallis (Location)	2	6.56	0.04
Residual	16	11.44	-
$\left(\frac{n-1}{n}\right) X_p^2$ (Total)	18	18.00	-

data; any table of zeros and ones with column totals of one and row totals 4, 3, 3 will have $X_P^2 = 20$. This means that in Table 2, that gives details of the analysis, we cannot assign p-values to the residual or to $\left(\frac{n-1}{n}\right) X_P^2$.

The residual here is numerically small compared to its degrees of freedom. Intuition correctly suggests that further partitioning of $\left(\frac{n-1}{n}\right) X_P^2$ to obtain components assessing location, dispersion and a new residual involving skewness and higher order effects, will not find further significant effects.

It is possible to further decompose each component. Here we may calculate values c_{11}, c_{12}, c_{13} of the linear/location component such that the Kruskal-Wallis statistic takes the value c_1 is here given by

$$c_1 = c_{11}^2 + c_{12}^2 + c_{13}^2.$$

The statistic c_{i1} reflects a location effect for the ith treatment. We find $c_{11} = -1.98, c_{12} = 1.05$ and $c_{13} = 1.24$. An approximate LSD for the c_{1i} is $2\sqrt{2}$. Thus the 'average' ranks for consultancy B and C tasters do not differ, but the consultancy A 'average' rank does differ compared to the other two consultancies. This agrees with the ANOVA result reported above.

4 The randomised block design

When data for a randomised block design for comparing t treatments is put into a treatment by ranks table, Anderson (1959) showed that, when there are no ties, the usual Pearson statistic does not have the chisquared distribution, but a simple multiple of it, $A = \{(t-1)/t\}X_P^2$ does have the $\chi^2_{(t-1)^2}$ distribution under the null hypothesis of no distributional differences between treatments. Rayner and Best (2002) show how to decompose this statistic into the Friedman statistic, a new dispersion assessing statistic, and as many other components as may be informative. Usually these other components are combined into a residual.

Table 3. Consumer by breed table of midrank observations

	Breed		
Consumer	A	B	C
1	2.5	1	2.5
2	1.5	3	1.5
3	2.5	1	2.5
4	1	3	2
5	1.5	3	1.5
6	1	2.5	2.5
7	1	3	2
8	1.5	3	1.5

Table 4. Decomposition of the Anderson statistic for the geese meat flavour data

Statistic	Value	Degrees of freedom	p-value
Location (Generalised Friedman)	3.769	2	0.152
Dispersion	8.714	2	0.013
Total (Generalised Anderson)	12.484	4	0.014

When ties occur, as is frequently the case, recommendations include ignoring the tied data, randomly breaking the ties, or deriving an entirely new analysis, such as that given in Brockhoff, Best and Rayner (2002). Their generalised Anderson statistic has not been given previously. The approach is demonstrated in the following example.

Geese Meat Flavour Example. Suppose we have a taste-test involving meat from three breeds of geese. Flavour rankings were obtained and are given in Table 3. Note that consumers frequently cannot completely distinguish their preferences, so there are many ties in the data. The Anderson statistic, generalised to allow for ties, is decomposed in Table 4.

The overall difference in the meats from the various breeds, as assessed by the generalised Anderson statistic, is significant at the 5% level but not at the 1% level. The 15% p-value for the generalised Friedman statistic indicates that the location differences, reflected by the rank sums of 12.5 for A, 19.5 for B and 16 for C, are not significant at the usual levels. However there is a significant dispersion difference. Decomposition of this dispersion component permits us to analyse pairwise differences; see Table 5. These indicate a difference in dispersions for breeds B and C. Breed B has a large dispersion, with many large and small ranks, whereas there is strong agreement about

Table 5. Pairwise comparisons; statistic values (upper triangle), and p-values (lower triangle)

	Breed		
Breed	A	B	C
A	-	5.48	3.62
B	0.065	-	9.62
C	0.164	0.008	-

Table 6. ANOVA for the geese meat flavour data

Source	df	SS	F	p-value
Breed	2	3.0625	2.16	0.153
Consumers	7	0.0000	0.00	
Residual	14	9.3375		
Total	23	12.400		

breed C, with mainly intermediate ranks.

If we do a two-way or randomised block analysis on the Table 3 data, we find that, ignoring consumers, the variances seem to be homogeneous, but that the residuals are not normally distributed (the Anderson-Darling test has p-value 0.032). Thus some caution in basing inference on the ANOVA is advisable. Table 6 gives the ANOVA. As with other ranked two-way or randomised block data, the consumers sum of squares must be zero, because each consumer mean is the same when ranks are used. Of course the ANOVA, inasmuch as it is valid at all, only assesses location differences and fails to find the dispersion effect.

5 The balanced incomplete block design

As with the previous designs, for balanced incomplete block data it is possible to present the data in a treatments by rank table. Suppose, consistent with section 2, that each of b blocks contains k experimental units, that each of t treatments appears in r blocks and every treatment appears with every other treatment precisely λ times. As in section 2 we can show that an Anderson-type statistic for balanced incomplete block designs with no tied data, A, is such that $A = \{(t-1)/t\}X_P^2$ may be decomposed into $k-1$ statistics assessing the differences between the treatments in the first $k-1$ moments. In large samples A approximately follows the $\chi^2_{(t-1,k-1)}$ distribution. Each of the $k-1$

Table 7. Balanced incomplete block design off-flavour scores for dried egg taste-test

Consumer	Samples				Scores			
1	a	b	d	e	9.7	8.7	5.4	5.0
2	b	c	f	j	9.6	8.8	5.6	3.6
3	b	d	f	g	9.0	7.3	3.8	4.3
4	a	c	e	g	9.3	8.7	6.8	3.8
5	a	d	h	j	10.0	7.5	4.2	2.8
6	b	g	h	i	9.6	5.1	4.6	3.6
7	b	e	h	j	9.8	7.4	4.4	3.8
8	e	g	i	j	9.4	6.3	5.1	2.0
9	a	b	c	i	9.4	9.3	8.2	3.3
10	d	e	f	i	8.7	9.0	6.0	3.3
11	a	f	g	j	9.7	6.7	6.6	2.8
12	c	d	i	j	9.3	8.1	3.7	2.6
13	a	f	h	i	9.8	7.3	5.4	4.0
14	c	d	g	h	9.0	8.3	4.8	3.8
15	c	e	f	h	9.3	8.3	6.3	3.8

moment detecting components approximately follows the $\chi^2_{(t-1)}$ distribution in large samples. The first statistic in the decomposition assesses location differences between the treatments and is identical to Durbin's rank statistic for balanced incomplete blocks.

Dried Egg Example. In sensory evaluation there is evidence to suggest that tasting more than four or five samples at one sitting results in reduced acuity arising from confusion or fatigue associated with making a large number of comparative judgements and/or with taste-bud saturation. For this reason incomplete block designs are employed. Rayner and Best (1996) quoted results from a taste-test on ten dried egg samples, $a, b, ..., j$ where the design used was a balanced incomplete block design. Table 7 gives the design for 15 consumers and the corresponding scores for 'off-flavour'. If these scores are ranked 1, 2, 3, 4 and counts are made of how many times each sample received each rank, then Table 8 results. Here a rank of 1 is given to the highest score and a rank of 4 to the lowest score.

Our nonparametric analysis results in Table 9. We see that both linear and quadratic effects are significant and from Table 8 we see that as we go from product a to product j off-flavour worsens. However, also from Table 8, we see that the quadratic effects are due to clumping at one end only, or a middle rank clumping and not to some instances of two clumps. Thus

Table 8. Counts of ranks for the dried egg data

	Rank			
Product	1	2	3	4
a	6	0	0	0
b	4	2	0	0
c	3	2	1	0
d	0	5	1	0
e	2	2	1	1
f	0	2	3	1
g	0	2	3	1
h	0	0	4	2
i	0	0	2	4
j	0	0	0	6

Table 9. Partition of Anderson's statistic for the dried egg data

Statistic	df	SS	χ^2 p-value	Monte Carlo p-value
Location (Durbin)	9	39.12	< 0.001	< 0.001
Dispersion	9	22.80	0.007	0.002
Residual	9	10.08	0.344	0.337
$\{(t-1)/t\}X_P^2$ (Anderson)	27	72.00	-	-

there is no suggestion of market segmentation. Table 9 also gives Monte Carlo permutation test p-values as the chisquared approximation for balanced incomplete block statistics may not always be accurate.

REFERENCES

Anderson, R.L. (1959). Use of contingency studies in the analysis of consumer preference studies. *Biometrics* 15, 582-590.

Brockhoff, P.B., Best, D.J. and Rayner, J.C.W. (2002). Partitioning Anderson's statistic for tied data. Submitted.

O'Mahony, M. (1986). *Sensory Evaluation of Food - Statistical Methods and Procedures.* New York: Marcel Dekker.

Rayner, J.C.W. and Best, D.J. (1996). Extensions to some important nonparametric tests. *Proceedings of the A.C.Aitken Centenary Conference,*

Dunedin, 1995, Kavalieris, L., Lam, F.C., Roberts, L.A. and Shanks, J.A. (editors). Dunedin: University of Otago Press.

Rayner, J.C.W. and Best, D.J. (2001). *A Contingency Table Approach to Nonparametric Testing.* Boca Raton: Chapman & Hall/CRC.

Rayner, J.C.W. and Best, D.J. (2002). More informative nonparametric testing with sensory evaluation application. In preparation.

Schach, S. (1979). An alternative to the Friedman test with certain optimality properties. *Annals of Statistics* 7, 537-550.

MAXIMUM LIKELIHOOD ESTIMATION OF CORRELATION MATRIX UNDER INEQUALITY CONSTRAINTS USING GIBBS SAMPLING

TOSHIO SAKATA
Department of Industrial Design
Kyushu Institute of Design
Minami-ku Shobaru Fukuoka, Japan
E-mail: Sakata@kyushu-id.ac.jp

RYUICHI SAWAE
Department of Applied Mathematics
Okayama University of Science
Ridai-cho, Okayama, Japan
E-mail: sawae@xmath.ous.ac.jp

KENTARO NOMAKUCHI
Department of Mathematical Science
Kochi University
Akebono-cho, Kochi, Japan
E-mail: nomakuti@cc.kochi-u.ac.jp

The maximum likelihood estimation of correlation matrix under order restrictions among correlations is treated. Two maximization process:(A) maximization with respect to correlation matrix and (B) maximization with respect to variance are considered. For the maximization process (A), we generate uniformly distributed random correlation matrices on the hypothesis space by Gibbs sampling. In the maximization process (B), we show that the maximum point is the fixed point of the iterative application of a certain non-linear function and the convergence of the process (B) is proved. A simulation result is given which compares the relative errors of the m.l.e. and other competitors.

Keywords: maximum-likelihood estimation, correlation matrix, Gibbs sampling.

1 Introduction

Motivation

We consider the maximum likelihood estimation of the correlation matrix of a multivariate normal distribution under order restrictions among component correlations. For example, we want to estimate the correlation matrix of the four variables X_1, X_2, X_3 and X_4 where X_1 denotes the test score in an

entrance examination acquired by students and X_2, X_3 and X_4 denote respectively the test score acquired by the same student in the pre-tests executed in one month, two months and three months ago respectively. Putting ρ_{ij} the correlation of the variable X_i and X_j, it seems to be natural to assume the following two types of monotonicity among component correlations,

$$H_c : \rho_{12} > \rho_{13} > \rho_{14} \quad and \quad \rho_{23} > \rho_{24} \tag{1}$$

and

$$H_r : \rho_{13} < \rho_{23} \; and \; \rho_{14} < \rho_{24} < \rho_{34}. \tag{2}$$

Thus the situation we consider seems to be considerably realistic. Here in general by H_c we denote the column decreasing hypothesis, that is, $\rho_{i,j} > \rho_{i,j+1}$ of all $1 \leq i \leq p$ and $1 \leq j \leq p-1$ and by H_r the row increasing hypothesis, that is, $\rho_{i,j} < \rho_{i+1,j}$ of all $1 \leq i \leq p-1$ *and* $1 \leq j \leq p$. Further we put $H_{rc} = H_r \cap H_c$. We denote by H_{all} "no hypothesis", that is, the space is the set of all correlation matrices, and by H_{rc}^+ we denote $H_{rc} \cap \{\rho_{ij} > 0$ of all $\{i,j\}$. Thus we have the following decreasing sequence of hypotheses

$$H_{all} \rightarrow \{H_c, H_r\} \rightarrow H_{rc} = H_r \cap H_c \rightarrow H_{rc}^+. \tag{3}$$

The problem is to obtain the maximum likelihood estimator (m.l.e) for the correlation matrix under each hypothesis. Though for simplicity we treat the hypothesis H_{all} and H_{rc}^+ in this paper other hypotheses are treated in a similar way. This work was inspired by the work of Gelfand et al. (1995) which first used Gibbs sampling for several inferential problems under order restrictions. Lie et al. (1995) used Gibbs sampling in estimating a covariance matrix and Calvin and Dykstra (1995) treated the problem of the maximum likelihood estimation of covariance matrices with constrained parameter spaces in the context of EM algorithm. The problem treated in this paper seems to be new as far as the authors know. The paper of Sakata and Sawae (1998) is closely related to this work. A general reference for ordered restricted inference is Robertson et al. (1988).

Likelihood function

In this paper we assume that the vector $X = (X_1, X_2, ..., X_p)$ has a p dimensional normal distribution $\mathcal{N}_p(0, \Sigma)$ and that we have a sample of size $N(> p)$ from this distribution. Then the sufficient statistic for Σ is the sample covariance matrix $S = (s_{ij})$ where $s_{i,j} = \sum_{k=1}^{k=N} X_{ik}X_{jk}/N$ and X_{ik} is the k-th sample of the variable X_i. Let $W = NS$. Then W has a Wishart distribution, $W(N, p, \Sigma)$. The density of $W(N, p, \Sigma)$ is given by

$$\frac{1}{2^{Np/2}\Gamma_p(\frac{N}{2})|\Sigma|^{\frac{N}{2}}} etr(-\frac{1}{2}\Sigma^{-1}W)|W|^{(N-p-1)/2}. \tag{4}$$

where $etr(\cdot)$ denotes $exp(trace(\cdot))$. Now let R denote the sample correlation matrix, that is,

$$R = (r_{ij}), \quad r_{ij} = \frac{w_{ij}}{\sqrt{w_{ii}w_{jj}}}. \tag{5}$$

Let $v_i = w_{ii}$ for $i = 1, 2, ..., p$ and $V = diag(v_1, v_2, ..., v_p)$.

Then the statistic (R, V) has the joint density given by

$$\frac{1}{2^{Np/2}\Gamma_p(\frac{N}{2})|\Sigma|^{\frac{N}{2}}} etr(-\frac{1}{2}\Sigma^{-1}V^{1/2}RV^{1/2})|W|^{(N-p-1)/2}\prod_{i=1}^{p} v_i^{(p-1)/2}.$$

Here we consider the corresponding decomposition of the population covariance matrix Σ into the population correlation matrix Γ and the population variances $\Delta = diag(\sigma_{11}, \sigma_{22}, ..., \sigma_{pp})$, that is,

$$\Sigma = diag(\sqrt{\sigma_{11}}, \sqrt{\sigma_{22}}, \ldots, \sqrt{\sigma_{pp}})\ \Gamma\ diag(\sqrt{\sigma_{11}}, \sqrt{\sigma_{22}}, \ldots, \sqrt{\sigma_{pp}}) \tag{6}$$

$$= \Delta^{1/2}\Gamma\Delta^{1/2}. \tag{7}$$

Then we have the following log likelihood function for Γ and Δ

$$-\frac{N}{2}\log|\Gamma| + N\log\prod_{i=1}^{p} t_i - N\log\prod_{i=1}^{p}\sqrt{v_i} - \frac{1}{2}trace(\Gamma^{-1}TRT), \tag{8}$$

where $t_i = \sqrt{\frac{v_i}{\sigma_{ii}}} = \sqrt{\frac{Ns_{ii}}{\sigma_{ii}}}$ and $T = \Delta^{-1/2}V^{1/2} = diag(t_1, ..., t_p)$.

As far as the maximization, it suffices to consider the following equivalent function with the likelihood function

$$\ell(\Gamma, t) = -\frac{N}{2}\log|\Gamma| + N\log\prod_{i=1}^{p} t_i - \frac{1}{2}trace(\Gamma^{-1}TRT). \tag{9}$$

Maximization of the function $\ell(\Gamma, t)$

We need to maximize $\ell(\Gamma, t)$, that is, to take

$$\sup_{\Gamma\in H_*}\ \sup_{t>0}\ell(\Gamma, t), \tag{10}$$

where H_* denotes any hypothesis and $t > 0$ means that $t_i > 0$ for all i. In order to take the above maximization we execute the following two maximization processes iteratively.

(A) Correlation maximization process

For a given $t^{(i)}$, the i-th step value of t, we take the maximum with respect to correlations, that is,

$$\sup_{\Gamma\in H_*}\ell(\Gamma, t^{(i)}) = \ell(\Gamma^{(i+1)}, t^{(i)}). \tag{11}$$

(B) Variance maximization process
For a given $\Gamma = \Gamma^{(i+1)}$, we take the maximum with respect to variances, that is,

$$\sup_{t>0} \ell(\Gamma^{(i+1)}, t) = \ell(\Gamma^{(i+1)}, t^{(i+1)}). \tag{12}$$

The maximization process (A) is done by a random search. That is, we generate a certain number, say k, of random correlation matrices on the hypothesis H_* and find a Γ which gives the maximum value of $\ell(\Gamma, t_i)$ among them. Note that the computing time increases linearly in k. Though it does not take so much time to obtain a maximum likelihood estimator for an arbitrary k, it takes much time to repeat the simulations to see the performance of the m.l.e. And so we used $k = 1$ in this paper. This means that the maximization about Γ is done through only one iteration of (A) and (B). The maximization process is done by solving a certain non-linear equation numerically. It is shown that this process converges.

We iterate both processes and stop at a predetermined number of times. Finally we adopt a pair of (Γ, t) as the maximum likelihood estimator which gives the maximal value of the likelihood among all pairs that appear in the iteration.

2 Gibbs sampling and uniform random correlations

Here we consider to generate uniformly distributed random elements over the set of the correlation matrices. We use the Gibbs sampling method. The Gibbs sampling is one of the methods to generate random vectors obeying the target distribution in the limit through a certain Markov chain. In general the one step of the Markov chain is composed of the iterative generations of one dimensional random variable from the conditional distribution given all other coordinate variables. For the definition of Gibbs sampling and Markov chain Monte Carlo methods see Bernardo and Smith (1994), chapter 5. For its applications to restricted inference see Gelfand et al. (1995). In our Gibbs sampling we cyclically generate correlation coefficients with other correlation coefficients fixed. The following Lemma and the statement after the lemma state how the method works.

Lemma 1 *Let $\Gamma = (\rho_{ij}) > 0$ be a correlation matrix, and ρ_{-ij} be the set of all upper half components of Γ other than ρ_{ij}. For fixed ρ_{-ij}, in order that $\Gamma = (\rho_{ij})$ is a positive definite correlation matrix, ρ_{ij} must be in an interval.*

Proof: Without loss of generality we assume $i = 1$ *and* $j = 2$. The condition that Γ is a positive definite correlation matrix means that all the

$q(q = 1, 2, \ldots, p)$-th main diagonal square matrix from the right corner have a positive determinant. Since the q-th $(q \leq p-1)$ main diagonal matrix is unrelated to the element ρ_{12}, the condition is equivalent to that the determinant of Γ is positive. We now show that ρ_{12} is in some interval. Let us expand the determinant $|\Gamma|$ with respect to the second column. We have

$$(-1)\rho_{12}\begin{vmatrix} \rho_{12} & \rho_{23} & ** & \rho_{2p} \\ \rho_{13} & 1 & ** & \rho_{3p} \\ * & * & ** & * \\ \rho_{1p} & \rho_{3p} & ** & 1 \end{vmatrix} + \begin{vmatrix} 1 & \rho_{13} & ** & \rho_{1p} \\ \rho_{13} & 1 & ** & \rho_{3p} \\ * & * & ** & * \\ \rho_{1p} & \rho_{3p} & ** & 1 \end{vmatrix}$$

$$-\rho_{32}\begin{vmatrix} 1 & \rho_{13} & ** & \rho_{1p} \\ \rho_{12} & * & ** & \rho_{2p} \\ * & * & ** & * \\ \rho_{1p} & \rho_{3p} & ** & 1 \end{vmatrix} + \cdots + (-1)^p \begin{vmatrix} 1 & \rho_{13} & ** & \rho_{1p} \\ \rho_{12} & * & ** & \rho_{2p} \\ * & * & ** & * \\ \rho_{1,p-1} & \rho_{3,p-1} & ** & \rho_{p-1,p} \end{vmatrix} > 0.$$

The left hand side of the above equation is a polynomial of degree two in $x = \rho_{12}$ and the coefficient of x^2 is the main diagonal matrix obtained from Γ by deleting first and second columns and rows from Γ, and so is positive.

Hence the condition is the inequality

$$-ax^2 + bx + c \;>\; 0, \quad a > 0.$$

This proves ρ_{12} is in the interval (α, β), where α and β are two real roots of the equation

$$-ax^2 + bx + c = 0.$$

Note that the iterative generation of the random numbers on 1-dimensional intervals appearing in the Lemma generates an uniform random correlation matrix $\Gamma \in H_{all}$. Further if we add the restriction such as $\rho_{i,j-1} > \rho_{ij} > \rho_{i,j+1}$ to the interval for generating ρ_{ij}, it would generate the uniform random correlation matrices on the hypothesis H_c and it is easy to generate the uniform distributed elements on H_r, H_{rc} and H_{rc}^+ similarly.

In the maximization process (A) we generate uniformly distributed random correlation matrices over the hypothesis, say, H_{all} by the Gibbs sampling to seek the correlation matrix maximizing $\ell(\Gamma, t)$ for a given t. A seemingly easier way to generate random correlations on H_{all} is to generate uniform random elements in the box $[-1, 1]^{\frac{p(p-1)}{2}}$ and adopt the element only if it is a positive definite correlation matrix, that is, in H_{all}. However the relative size of H_{all} to the box is very small even for $p \geq 4$ as is shown in Table 1. This is a main reason why we propose Gibbs sampling.

Table 1. Relative volume of Hypothesis H_{all}

p	3	4	5	6	7
Relative Volume	61.7	18.3	2.2	0.095	0.0013

We note that it is an interesting but difficult problem to obtain theoretically the marginal distribution of each correlation coefficient under the uniform distribution on any hypothesis space considered in this paper.

3 Maximization process (B)

In maximization process (B) for a given $\Gamma \in H$, we seek $t = (t_1, t_2, ..., t_p)$ which maximizes the likelihood $\ell(\Gamma, t)$.

Theorem 2 *For a given $\Gamma \in H$ $t = (t_1, t_2, ..., t_p)$ which maximizes the likelihood is the solution of the following non-linear equation,*

$$N/t = Bt, \tag{13}$$

where

$$B = (b_{ij}) = (\rho^{ij} r_{ij}) \text{ and } 1/t = (1/t_1, \ldots, 1/t_p).$$

Proof: The equation is obtained by taking the partial derivative of the likelihood with respect to t.

Our second main result is the following theorem.

Theorem 3 *The iterative method described below converges to a solution in the positive orthant.*

Description: *Take a starting point $t^{(0)} = (t_1^{(0)}, t_2^{(0)}, \ldots, t_p^{(0)})$ in the positive orthant. First we solve the equation (13) as one of degree two in t_1 and put the positive solution as $t_1' = f_1(t_2^{(0)}, t_3^{(0)}, \ldots, t_p^{(0)})$. Next we solve the same (13) as the equation of degree two in t_2 with t_1' in place of $t_1^{(0)}$ and put the positive solution as $t_2' = f_2(t_1', t_3^{(0)}, \ldots, t_p^{(0)})$. Likewise we have $t_3' = f_3(t_1', t_2', t_3^{(0)}, \ldots, t_p^{(0)}), \ldots, t_p' = f_p(t_1', \ldots, t_{p-1}')$. We set $t^{(1)} = (t_1', t_2', ..., t_p')$. By repeating the process with $t^{(1)}$ as a starting point, we have $t^{(2)}$. Continuing this process we have the sequence $t^{(0)}, t^{(1)}, ...,$.*

Proof: First we show that the sequence given by the description above converges to a fixed point of a certain map if it converges. By f we denote the

map f from $\Re^p$ to itself by $(t_1^{(j)}, t_2^{(j)}, \ldots, t_p^{(j)}) \to (t_1', t_2', \ldots, t_p')$ in the description. Our iteration method thus can be described as the iterative applications of the same mapping to the starting point. That is, the sequence of the points is given by

$$t^{(n)} = f^{(n)}(t^{(0)}) = f(f(\cdots f(t^{(0)})\cdots)). \tag{14}$$

If these sequence of points converge to some point $t^{(\infty)}$ it is necessarily the fixed point of the mapping. Since $t^{(\infty)}$ is the fixed point we have

$$t^{(\infty)} = f(t^{(\infty)}).$$

Then we have $t_1^{(\infty)} = f_1(t_2^{(\infty)}, t_3^{(\infty)}, \ldots, t_p^{(\infty)})$, $t_2^{(\infty)} = f_2(t_1^{(\infty)}, t_3^{(\infty)}, \ldots, t_p^{(\infty)})$, $\ldots$, $t_p^{(\infty)} = f_p(t_1^{(\infty)}, t_2^{(\infty)}, \ldots, t_{p-1}^{(\infty)})$. This shows that the fixed point $t^{(\infty)}$ is the solution of the non-linear equation (13). Moreover the process shows that the solution is in the positive orthant.

Next we show the convergence. Let us consider the following function:

$$h(t) = g(TR^{-1}T, \Gamma),$$

where

$$g(V, W) = \log|VW| - \text{tr } VW.$$

Then we have easily that

$$\frac{\partial h}{\partial t} = Nt^{-1} - 2Ut,$$

$$\frac{\partial^2 h}{\partial t \partial t'} = -NT^{-2} - 2U,$$

where $t^{-1} = (1/t_1, 1/t_2, .., 1/t_p)'$, $T = diag(t_1, ..., t_p)$ and

$$U = (u_{ij}), \quad u_{ij} = (R^{-1})_{ij}\Gamma_{ij} \text{ for } i = 1, 2, ..., N, \ j = 1, 2, ..., N.$$

From Rao (2000) it follows that the Hessian matrix $\partial^2 h/\partial t \partial t'$ is negative definite. Hence, the function h is strictly concave in every orthant without boundary, especially in the positive orthant. Noting that h is equivalent to the likelihood function this enables us to apply the convergence theorem of Warga (1963) for our procedure. In the conditions of his theorem, the concavity is essential and others can be easily checked, and hence we have the convergence of our iterative procedure.

It should be noted that from our experience the non-linear equation can be solved by at most 20 iterations in the process (B).

Table 2. True $\Gamma_* \in H_{rc}^+$

	X_1	X_2	X_3	X_4
X_1	1.0000	0.9258	0.8452	0.7559
X_2	*	1.0000	0.9129	0.8165
X_3	*	*	1.0000	0.8944

4 Comparisons of relative errors of m.l.e. and others

In this section we compare the relative error, the mean absolute error divided by a true value, among three estimators of the population correlation matrix $\Gamma_* \in H_{rc}^+$, R_{gibbs}, R_{pava} and R. R_{gibbs} is the pseudo m.l.e. obtained by Gibbs sampling, and R_{pava} is the estimator obtained by Dykstra's matrix order PAVA algorithm, which calculates the isotonic regression of the rows and the columns itelatively (see the page 27 of Robertson et al. (1988) for the definition of matrix order PAVA algorithm). Note that both R_{gibbs} and R_{pava} are calculated based on the usual sample correlation matrix R.

In order to run a Gibbs sampling on H_{rc}^+, first we choose an initial correlation matrix Γ_0 in H_{rc}^+ and an initial variance vector V_0. After 50000 times repetitions of the process (A) and (B) we take out R_{gibbs} which gives the highest likelihood value among the 50000 Gibbs samples. By 100 times simulations we obtained the relative error for each estimator. We should note that in the above simulation Γ_0 and $\Gamma_* \in H_{rc}^+$ are obtained as

$$\rho_{ij} = \sqrt{\frac{\kappa_1 + \kappa_2 + ... + \kappa_i}{\kappa_1 + \kappa_2 + ... + \kappa_j}} \tag{15}$$

with $\kappa_1 = 0.8, \kappa_2 = 0.6$ and $\kappa_3 = 0.4$ for Γ_* and $\kappa_1 = 0.5, \kappa_2 = 0.7$ and $\kappa_3 = 0.8$ for Γ_0. It is easy to show that the matrix whose components are given by (15) is a correlation matrix. Table 2 gives correlation matrix $\Gamma_* \in H_{rc}^+$.

Below we present the simulation results about the relative errors.

From Table 3 we see that R_{gibbs} is slightly efficient than R and there is no difference between R_{gibbs} and R_{pava}.

REFERENCES

Bernardo, J.M. and Smith, A.F.M. (1994). *Bayesian Theory.* New York: John Wiley.

Calvin, J.A. and Dykstra, R.L. (1995). RELM estimation of covariance matrices with restricted parameter spaces. *J. Amer. Stat. Ass.* 90, 321-329.

Table 3. Relative errors $\times 100$ of $R(\text{upper})(n = 10)$

	X_1	X_2	X_3	X_4
X_1	0.00	8.93	20.44	26.05
	0.00	9.1	18.6	24.7
	0.00	8.90	19.74	25.56
X_2	*	0.00	9.62	15.50
	*	0.00	9.35	15.10
	*	0.00	9.16	15.26
X_3	*	*	0.00	6.68
	*	*	0.00	5.77
	*	*	0.00	6.60

Gelfand, A.E., Smith, A.F.M. and Lee, T.M. (1995). Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *J. Amer. Stat. Ass.* 87, 523-532.

Lui, J.S., Wong, W.H. and Koung, A. (1995). Covariance structure and convergence rate of the Gibbs sampler with various scans. *J. Roy. Stat. Soc. Ser.* B.57, 157-169.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory.* New York: John Wiley.

Rao, C.R. (2000). Statistical proofs of some matrix inequalities. *Linear Algebra and its Applications* 321, 307-320.

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference.* New York: John Wiley.

Sakata, T. and Sawae, R. (1998). Maximum likelihood estimation of correlation matrix under inequality constraints and Gibbs sampling. *Statistical Research Report SRR 005-98*, Centre for Maths and its Applications, Australian National University.

Warga, J. (1963). Minimizing certain convex functions. *J. Soc. Indust. Appl. Math.* 11, 583-93.

RESERVICING SOME CUSTOMERS IN $M/G/1$ QUEUES, UNDER TWO DISCIPLINES

M.R. SALEHI-RAD
The Ferdowsi University of Mashhad, Iran
E-mail: salehi_rad@hotmail.com

K. MENGERSEN
University of Newcastle, NSW 2308, Australia
E-mail: Kerrie.Mengersen@maths.newcastle.edu.au

Consider a production line in which a production item is failed with probability p and it is then repaired. We consider two repair disciplines I and II. For each discipline we find the p.g.f. of the steady state size of the system at the moment of departure of the customer. For discipline II, we show that the closed form of the p.g.f. depends on solving a functional equation that is the Laplace-Stieltjes transform of the service time in the main queue. Through an example, we compare the two disciplines in terms of specified performance measure.

Keywords: busy period, idle period, imbedded Markov chain, M/G/1 queue, probability generating function, reservicing.

1 Introduction

In this paper we consider an $M/G/1$ queueing model in which some customers, after the completion of their service, require reservice with probability p. For example in a production line, some items may be failed with probability p and must be repaired. Here,the production line is considered as the main queue (MQ) and the failed items join a "failed queue" (FQ). We consider two disciplines I and II for repairing the failed items. For each, we find a steady-state probability generating function (p.g.f.) for the size of the system, and show that it depends on the Laplace-Stieltjes transform (for convenience LST) of the distribution functions of service and repair times. Through an example, we then compare the two disciplines in terms of the probability (or proportion of time) that the server is idle. Customers arrive in the system in a Poisson process with rate λ. Service (s) and repair ($\tilde{s}$) times are independent and have general distributions, denoted by $B_1(t)$ and $B_2(t)$ respectively, with means $1/\mu_1$ and $1/\mu_2$.

In discipline I, if the server is in MQ, then when a service is completed, if the item served has failed, then the server switches to repair it; otherwise

the server continues to service MQ customers if any are present, or if not the server is idle.

In discipline II, if the server is in MQ then on completion of the service, the server continues to service any other MQ customers, or if there are no more MQ customers but if there are FQ customers, the server attends to these. If,however,there are no more MQ and no more FQ customers,the server is idle. If the server is in FQ, then the server continues in FQ until FQ is empty,and then switches to MQ customers if any are present; otherwise, the server becomes idle. If the server is idle, service starts again as soon as customers arrive in MQ.

As mentioned, our aim is to find analytic expressions for the steady-state p.g.f. of the system size under the two disciplines at the moment of service completion. Since we are studying an $M/G/1$ model, we may consider an imbedded Markov chain in which the size of the system is the number of customers remaining in the system at the time of service completion for each customer. The method of the imbedded Markov chain allows us to simplify analysis, since it converts a non-Markovian into a Markovian problem[see Kleinrock (1976)]. In Discipline I, the imbedded Markov chain is $X(t_n)$ (or for convenience, X_n), the number of customers remaining in MQ at the completion of the nth customer's service time. In Discipline II, we require the bivariate imbedded Markov chain $(X(t_n), Y(t_n))$ (or for convenience (X_n, Y_n)), where Y_n is the number of customers remaining in FQ at the completion of the nth customer's service time.

This problem has been considered from a slightly different perspective by Trivedi (1982). In his problem, as the items come to be serviced, the failed item returns to the queue with probability q and servicing is completed with probability p. In this paper, we do not allow the item to return to the queue, but store the item with probability p if it requires reservice. We suppose that there is no set-up time, but this condition is the subject of ongoing research by the authors.

We find the required expressions for the p.g.f. under discipline I and II in sections 2 and 3 respectively. At the end we give an example, and by discussion we compare a performance measure of the two disciplines.

2 The queueing discipline I

Let $t_1, t_2, \ldots$, be times at which a service is completed. Let $X(t_i) \equiv X_i$ be the number of customers remaining in the MQ at t_i. Then we have

$$X_{n+1} = (X_n - 1)_+ + A_{n+1}(s) + \left[\tilde{A}_{n+1}(\tilde{s}) - 1\right]_+ I_{\{X_n=0\}}, \qquad (1)$$

where $(a-1)_+ = \max\{a-1, 0\}$. In other words, X_{n+1} is given by

$$\begin{cases} X_n + A_{n+1}(s) - 1 & \text{if } X_n \geq 1 \text{ and departure at } t_{n+1} \text{ occurs} \\ X_n + A_{n+1}(s) + \tilde{A}_{n+1}(\tilde{s}) - 1 & \text{if } X_n \geq 1 \text{ and departure at } t_{n+1} \text{ is failed} \\ A_{n+1}(s) & \text{if } X_n = 0 \text{ and departure at } t_{n+1} \text{ occurs} \\ \left[\tilde{A}_{n+1}(\tilde{s}) - 1\right]_+ + A_{n+1}(s) & \text{if } X_n = 0 \text{ and departure at } t_{n+1} \text{ is failed} \end{cases} \tag{2}$$

where $A_{n+1}(s)$ is the number of arrivals during the servicing of the $(n+1)$th customer in MQ, and $\tilde{A}_{n+1}(\tilde{s})$ is the number of arrivals during repair in FQ.

Remark 1: $[\tilde{A}_{n+1}(\tilde{s}) - 1]_+$ means: when $X_n = 0$ and the serviced item is failed, it must be repaired.Then,either there are no arrivals during the time spent reservicing failed item, in which case the server then becomes idle until an arrival does occur,or there is a first such arrival and service on this customer starts immediately the server finishes on the failed item.Then to cover both these cases in the one equation, we use the $[\tilde{A}_{n+1}(\tilde{s}) - 1]_+$ function.

Remark 2: $A_{n+1}(s)$ and $\tilde{A}_{n+1}(\tilde{s})$ are independent of X_{n+1} and also of each other and of n, and depend only on s and $\tilde{s}$, respectively. Thus we denote these by $A(s)$ and $\tilde{A}(\tilde{s})$ in the following. Three Lemmas become useful for finding the required p.g.f.

Lemma 1 *If $A(s)$ and $\tilde{A}(\tilde{s})$ are the number of arrivals during service and reservice times, respectively, then their p.g.f.s are $Q_i(u) = B_i^*[\lambda(1-u)], i = 1, 2$, where $B_1^*(\cdot)$ and $B_2^*(\cdot)$ are the LSTs of the service and repair times.*

Proof. We obtain the expression of the p.g.f. of $A(s)$, noting that the proof for $\tilde{A}(\tilde{s})$ is similar.

$$Q_1(u) = E\left[u^{A(s)}\right] = \sum_{j=0}^{\infty} u^j \int_0^{\infty} \frac{e^{-\lambda s}(\lambda s)^j}{j!} dB_1(s) = \int_0^{\infty} e^{-\lambda(1-u)s} dB_1(s) \ ,$$

that is, $B_1^*[\lambda(1-u)]$.

Lemma 2 *By Lemma 1 we have $E[\Pr\{\tilde{A}(\tilde{s}) = 0\}] = B_2^*(\lambda)$.*

Proof. $E[\Pr\{\tilde{A}(\tilde{s}) = 0\}] = E\left[e^{-\lambda \tilde{s}}\right] = \int_0^{\infty} e^{-\lambda \tilde{s}} dB_2(\tilde{s}) = B_2^*(\lambda)$.

Lemma 3 *If X is a non-negative integer valued r.v. with p.g.f. $P(u)$, then*

$$E\left[u^{(X-1)_+}\right] = u^{-1}[P(u) - (1-u)\Pr(X=0)].$$

Proof. $E\left[u^{(X-1)_+}\right] = E\left[u^0 I_{\{X=0\}} + u^{(X-1)} I_{\{X \geq 1\}}\right] =$

$$\Pr\ (X=0) + u^{-1}\left[P(u) - \Pr(X=0)\right] = u^{-1}[P(u) - (1-u)\Pr(X=0)] \ .$$

Proposition 1 *According to (1) and (2) the p.g.f. of the number of customers in the system at the moment of departure is*

$$P(z) = \frac{(1-z)\, B_1^*\left[\lambda(1-z)\right]\left[1 - p[1 - B_2^*(\lambda)]\right]}{(1-p)B_1^*\left[\lambda(1-z)\right] + pC^*\left[\lambda(1-z)\right] - z}\ \pi_0 \ , \tag{3}$$

where $C^*(\cdot)$ is the LST of the convolution of service and reservice times, and π_0 is the probability that the MQ is empty and is equal to

$$\pi_0 = (1-\rho)\{1-p[1-B_2^*(\lambda)]\}^{-1} . \tag{4}$$

where $\rho = \rho_1 + p\rho_2$, $\rho_1 = \lambda/\mu_1$ and $\rho_2 = \lambda/\mu_2$.

Proof. $P(z) = E(z^{X_{n+1}}) =$

$$E\left[z^{X_n+A(s)-1}I_{\{X_n\geq 1\}}\right](1-p) + E\left[z^{X_n+A(s)+\tilde{A}(\tilde{s})-1}I_{\{X_n\geq 1\}}\right]p$$
$$+E\left[z^{A(s)}I_{\{X_n=0\}}\right](1-p) + E\left[z^{A(s)+[\tilde{A}(\tilde{s})-1]_+}I_{\{X_n=0\}}\right]p =$$
$$E\left[z^{A(s)}\right]\left\{\frac{1}{z}\left[E\left(z^{X_n}I_{\{X_n\geq 1\}}\right)(1-p) + E\left(z^{X_n}I_{\{X_n\geq 1\}}\right)E\left(z^{\tilde{A}(\tilde{s})}\right)p\right]\right.$$
$$\left.+\pi_0\left[(1-p) + E\left(z^{[\tilde{A}(\tilde{s})-1]_+}\right)p\right]\right\}$$

Now, by using Lemmas 1 to 3, and since $E\left(z^{X_n}I_{\{X_n\geq 1\}}\right) = P(z) - \pi_0$,

$$P(z) = B_1^*[\lambda(1-z)]\left\{\frac{1}{z}[(1-p)(P(z)-\pi_0) + p(P(z)-\pi_0)B_2^*[\lambda(1-z)]]\right.$$
$$\left.+\pi_0\left[(1-p) + \frac{p}{z}\left[B_2^*[\lambda(1-z)] - (1-z)E\left[\Pr\left(\tilde{A}(\tilde{s})=0\right)\right]\right]\right]\right\}$$

After simplification, we reach (3).

Two special cases are as follows:

a) If $p = 0$, then $P(z) = \{(1-z)\,B_1^*[\lambda(1-z)]\}\{B_1^*[\lambda(1-z)] - z\}^{-1}\pi_0$ (5)

with $\pi_0 = 1 - \lambda/\mu_1 = 1 - \lambda E[$ service time $] = 1 - \rho_1$. This is similar to the $M/G/1$ queue without any conditions [see Gross and Harris (1985)].

b) If $p = 1$, so that each customer is served twice then

$$P(z) = \{(1-z)B_1^*[\lambda(1-z)]\,B_2^*(\lambda)\}\{C^*[\lambda(1-z)] - z\}^{-1}\pi_0 \tag{6}$$

with $\pi_0 = (1-\rho)[B_2^*(\lambda)]^{-1}$.

Note that, in order to find π_0 in (3), (5) and (6), we use Hôpital's rule and the fact that $P(1) = 1$.

Remark 3: We can think of the service time of an individual as being the service time in MQ plus any 'reservice time' (conditional on requiring such reservice), times the probability p of the item being failed. Then E[service time]$=1/\mu_1 + p/\mu_2 = 1/\mu$. Now, we have an $M/G/1$ queue with mean service time $1/\mu$. On the other hand, the mean of busy period in $M/G/1$ is found by taking the derivative of the functional equation $\Gamma(u) = B^*[u + \lambda(1-\Gamma(u))]$ [see Takacs (1955)], with $u = 0$, in which $\Gamma(u)$ and $B^*(\cdot)$ are the LST of the

distribution functions of busy period in MQ and service time, respectively, and is equal to $1/(\mu - \lambda)$[see Gross and Harris(1985)]. Therefore, E[busy period]$=\rho/\lambda(1-\rho)$. By using Little's law[see Stidham(1974)] and the mean of idle period$(1/\lambda)$, that is exponentially distributed, we have E(busy period)$/E$ (idle period)$=(1-\pi_0^*)/\pi_0^*$, which yields $\pi_0^* = 1-\rho$, where $\pi_0^* = \Pr$(idle period).

3 The queueing discipline II

In this case, since we store the failed items and then repair them, we require two variables. One of these, X_n, is the number of customers in the MQ at the epochs $\{t_n\}$,which are departure times, and the other, Y_n, is the number of failed items in store (FQ), again at the epochs $\{t_n\}$. We now have a bivariate imbedded Markov chain (X_n, Y_n). Thus (X_{n+1}, Y_{n+1}) is

$$\left((X_n - 1)_+ + A(s) + \left[\tilde{A}(\Sigma_{i=1}^{Y_n}\tilde{s}_i) - 1\right]_+ I_{\{X_n=0\}}, Y_n\ I_{\{X_n>0\}} + U_{n+1}\right) \quad (7)$$

in which $U_{n+1} = 1$ if departure has failed, and is otherwise zero. Since $\tilde{A}(\cdot)$ is a counting measure, then $\tilde{A}\left(\Sigma_{i=1}^{Y_n}\tilde{s}_i\right) = \Sigma_{i=1}^{Y_n}\tilde{A}(\tilde{s}_i)$ in which $\tilde{A}(\tilde{s}_i)$ is the number of arrivals during repair of the (i)th failed item in FQ. The $\tilde{A}(\tilde{s}_i)$ are i.i.d. r.v.'s. To evaluate the joint p.g.f. (X_n, Y_n) in the steady-state, we consider by $P(u,v) = E(u^{X_n}v^{Y_n})$, and develop the proposition below.

Proposition 2 *The joint p.g.f. of* (X_n, Y_n) *is*

$$P(u,v) = \{(1-p+pv)\,B_1^*[\lambda(1-u)] - u\}^{-1}\{(1-p+pv)B_1^*[\lambda(1-u)]\times [R(v) - G^*(u,p,\lambda) + (1-u)G^*(0,p,\lambda)]\}\pi_{0\bullet} \quad (8)$$

in which

$$R(v) = \Sigma\pi_{j|0}v^j\ ,\ j \geq 0 \quad (9)$$

$$\pi_{j|0} = \Pr\{Y_n = j \mid X_n = 0\} \quad (10)$$

$$\pi_{0\bullet} = \Pr(X_n = 0) \quad (11)$$

$$G^*(u,p,\lambda) = \psi\left(1-p+pB_2^*[\lambda(1-u)]\right) \quad (12)$$

$$\psi(u) = uB_1^*[\lambda(1-\psi(u))] \quad (13)$$

where $\psi(u)$ *is the p.g.f. of the number of served customers (departures) in a busy period (here in MQ) [see (Takacs (1955) and Saaty (1961))]. It is clear that the number of failed items in FQ is distributed as a binomial distribution with parameters p and K, where K is the number of served customers in a busy period in MQ and has p.g.f.* $\psi(u)$. *Expression (13) is a functional equation. Takacs has proved the existence and uniqueness of an analytic solution of* $\psi(u)$ *for* $|u| \leq 1$ *subject to* $\psi(0) = 0$. *In addition, he has shown that* $\lim\psi(u)$, *where*

$u \to 1$, equals the smallest positive real root of the equation $B_1^*[\lambda(1-x)] = x$. By solving Expression (13) we can find the p.g.f. of Y_n given $X_n = 0$, denoted by $R(\cdot)$, in terms of $\psi(u)$.

Proof. We start with (7) and use (9) – (13).

$$P(u,v) = E(u^{X_{n+1}} v^{Y_{n+1}}) =$$
$$E\big(u^{(X_n-1)+I_{\{X_n>0\}}+A(s)+\left[\sum_{i=1}^{Y_n}\tilde{A}(\tilde{s}_i)-1\right]_+ I_{\{X_n=0\}}} v^{Y_n I_{\{X_n>0\}}+U_{n+1}}\big)$$
$$= E(u^{A(s)})\Big\{\tfrac{1-p+pv}{u} E\big[u^{X_n} v^{Y_n} I_{\{X_n>0, Y_n\geq 0\}}\big] + \pi_{00}(1-p+pv) +$$
$$(1-p+pv) E\big[u^{\left[\sum_{i=1}^{Y_n}\tilde{A}(\tilde{s}_i)-1\right]_+ I_{\{X_n=0, Y_n>0\}}}\big]\Big\}$$

By Lemma 3,using $I_{AB} = I_A \times I_B$ and some computations, $P(u,v)$ is

$$(1-p+pv)B_1^*[\lambda(1-u)]\Big\{\tfrac{1}{u}\big[P(u,v) - R(v)\pi_{0\bullet}\big] + \pi_{00} +$$
$$\tfrac{\pi_{0\bullet}}{u}\big\{E\big(\big[B_2^*[\lambda(1-u)]\big]^{Y_n}\big) - (1-u)E\big(\big[B_2^*(\lambda)\big]^{Y_n}\big)\big\} - E\big(I_{\{X_n=0,Y_n=0\}}\big)\Big\}$$

Now, by simplifying, it yields (8).

Remark 4: Using $\lim P(u,1)$, where $u \to 1$, an ergodicity argument, Remark 3, Little's law, the means of the busy periods in MQ and FQ and the idle period, we can find $\pi_{0\bullet}$ and π_{00}^* which are Pr(MQ is empty) and Pr(idle period), respectively. On the other hand, we have

$$E[\text{busy period}] = E[\text{busy period in } MQ] + pE[\text{busy period in } FQ] \quad (14)$$

The first expression is equal to $1/(\mu_1 - \lambda)$. To find the second expression, π_0 and π_{00}, we use the results **a-g** below.

a $E(K) = \psi'(1) = (1-\rho_1)^{-1}$.

Proof. The proof is straightforward and is left to the reader.

b The LST of the distribution function of a busy period in FQ is

$$F^*(u) = G^*(0,p,u) = R(B_2^*(u)) \quad (15)$$

Proof. First we find $R(u)$ as follows.

$$R(u) = E(u^{Y_n}) = E\big[E[u^{Y_n}|K]\big] = E\big[(1-p+pu)^K\big] = \psi(1-p+pu)\ .$$

On the other hand, the repair time of failed items is $T = \Sigma_{i=1}^{Y_n}\tilde{s}_i$, where the $\tilde{s}_i$ are i.i.d and distributed as $B_2(t)$. We find the distribution of T as follows:

$$F(t) = \Pr(T \leq t) = \Pr\big[\Sigma_{i=1}^{Y_n}\tilde{s}_i \leq t\big] = E\big\{\Pr\big[\Sigma_{i=1}^{Y_n}\tilde{s}_i \leq t|Y_n\big]\big\}$$

If we take the LST of both sides of this equation, use the fact that the LST is a Linear Operator and exploit the LST convolution property and by $R(\cdot)$, then $F^*(u)$ is equal to

$$E\left[\mathrm{L}^*\left\{\Pr\left[\Sigma_{i=1}^{Y_n}\tilde{s}_i \le t | Y_n\right]\right\}\right] = E[(B_2^*(u))^{Y_n}] = \psi(1-p+pB_2^*(u))$$

Hence $F^*(u) = R(B_2^*(u)) = G^*(0,p,u)$.

c E [busy period in FQ] $= p[\mu_2(1-\rho_1)]^{-1}$. (16)

Proof. By taking the derivative of (15), with $u=0$, we find

$$F^{*'}(u)\big|_{u=0} = -E[\text{busy period in } FQ] = pB_2^{*'}(u)\psi'(1-p+pB_2^*(u)\big|_{u=0} .$$

Since $B_2^{*\prime}(0) = -1/\mu_2$ and using $\psi'(1)$ from **a**, the proof is completed.

d E [busy period] $= (\mu_2+p^2\mu_1)[\mu_2(\mu_1-\lambda)]^{-1}$.

Proof. By using (14), (16) and $1/(\mu_1-\lambda)$, we obtain the result.

e $\pi_{00}^* = (1-\rho_1)(1+p^2\rho_2)^{-1}$. (17)

Proof. From Little's Law, we have

E [busy period]/E [idle period] $= (1-\pi_{00}^*)/(\pi_{00}^*)$ which yields (17).

f $\pi_{0\bullet} = (1-\rho_1)^2[p\rho_2 + (1-\rho_1)G^*(0,p,\lambda)]^{-1}$.

Proof. We use Hôpital's rule and the fact that $\lim P(u,1) = 1$, where $u \to 1$.

g If $p=0$ and $v=1$, $P(u) = (1-u)B_1^*[\lambda(1-u)]\{B_1^*[\lambda(1-u)]-u\}^{-1}\pi_0$. This is similar to the special case (**a**) in section 2.

4 Concluding remarks

This paper has addressed the problem of M/G/1 queues with the possible repair of failed items. Two queuing disciplines have been discussed: in discipline I, the server repairs a failed item immediately, whereas in discipline II, failed items are stockpiled and repaired only after all customers in the main queue are served. For each discipline, explicit formulae are obtained for the steady state probability generating functions of the imbedded Markov chains involved.

By developing the corollary below, the two disciplines are compared in terms of the probability(proportion of time) that the server is idle, denoted by π_0^* and π_{00}^* for disciplines I and II, respectively.

Corollary 1 $\pi_0^* \le \pi_{00}^*$. *If* $p \neq 0$, *the inequality is strict.*

Table 1. Values of (π_0^*, π_{00}^*) for selected λ and p, with $\mu_1 = 2$, $\mu_2 = 6$

$p\backslash\lambda$	1/2	9/10	1	4/3
0.1	(.742,.749)	(.535,.549)	(.483,.499)	(.311,.333)
0.3	(.725,.744)	(.505,.543)	(.450,.490)	(.267,.327)
0.5	(.708,.735)	(.475,.530)	(.417,.480)	(.222,.316)
0.7	(.692,.721)	(.445,.512)	(.383,.462)	(.178,.301)
0.9	(.675,.703)	(.415,.491)	(.350,.441)	(.133,.283)

Proof. Suppose that the inequality is not true. Then we should have $\pi_0^* > \pi_{00}^*$, which means

$$-p^3\frac{\lambda^2}{\mu_2^2} + p^2(1 - \frac{\lambda}{\mu_1})\frac{\lambda}{\mu_2} - p\frac{\lambda}{\mu_2} > 0$$

By considering the cases $p = 0$, $p \neq 0$ and $p(1-\rho) \leq 1$, the proof is completed.

Hence, if the goal is to minimise this measure, discipline I is superior. This is illustrated numerically in table 1 for selected values of p, λ, μ_1, μ_2 and degenerated service and reservice times distribution.

Acknowledgments

We wish to thank Professor Daryl Daley for hosting the first author as a visitor at ANU and for seminal and ongoing discussions. We also thank Jan Garnsey for her typing assistance.

REFERENCES

Gross, D. and Harris, C.M. (1985). *Fundamentals of Queueing Theory.* New York: John Wiley and Sons.

Kleinrock, K. (1976). *Queueing Systems*, Vol. 2. New York: John Wiley.

Saaty, T.L. (1961). *Elements of Queueing Theory.* New York: McGraw-Hill.

Stidham, S. Jr. (1974). A last word on $L = \lambda W$. *Opers. Res* 22, 417-421.

Takacs, L. (1955). Investigation of waiting time problems by reduction to Markov Processes. *Acta. Math. Acad. Sci. Hung* 6, 101-129.

Trivedi, Kishor S. (1982). *Probability and Statistics with Reliability, Queueing and Computer Science Applications.* Englewood Cliffs: Prentice-Hall.

BIOINFORMATICS: STATISTICAL PERSPECTIVES AND CONTROVERSIES

PRANAB KUMAR SEN
Departments of Biostatistics and Statistics
University of North Carolina
Chapel Hill, NC 27599-7420, USA
E-mail: pksen@bios.unc.edu

In bioinformatics with special emphasis on computational biology, genomic science, polygenic models, and computational sequence analysis, principles of molecular genetics (biology) provide room for stochastics to comprehend the basic differences between mathematical exactness and biological diversity. With a large number of sites or loci having mostly categorical (qualitative) responses and imprecise dependence patterns, standard (discrete or continuous) multivariate statistical modeling and analysis may encounter roadblocks of various kinds. Limitations of likelihoods and their variants are appraised and contrasted with the inadequacy of the knowledge discovery and data mining approach that exploits mainly computational algorithms. Alternative approaches that take into account underlying biological implications to a greater (and parametrics to a lesser) extent are appraised in the light of validity and robustness considerations.

Keywords: biostochastics, DNA strand, genomics, hidden Markov models, high-dimensional data, KDDM, likelihoods, nonparametrics, stochastic evolutionary force.

1 Introduction

With the advent of modern information technology and scientific curiosity in human genomics, we are at the dawn of bioinformatics; it is not precisely known what really constitutes the domain of bioinformatics: pharmaceutical researchers, molecular biologists, computer scientists, biomathematicians, and statisticians differ in their definitions and interpretations, though to a much lesser extent than in their basic emphasis on methodologic versus computational aspects. Faced with this dilemma, let me quote a few lines from a very recent statistics text on bioinformatics by Ewens and Grant (2001):

We take bioinformatics to mean the emerging field of science growing from the application of mathematics, statistics, and information technology, including computers and the theory surrounding them, to study and analysis of very large biological and in particular, genetic data sets. The field has been fueled by the increase in the DNA data generation. In a similar manner, an image of bioinformatics has been drawn earlier by Waterman (1995), with

more emphasis on computational biology.

At the current stage, gene scientists can not scramble fast enough to keep up with the genomics, emerging at a furious rate and in astounding detail. This is particularly significant in view of the nearly completed status of the mapping of human genome sequence (IHGSC 2001, Venter et al. 2001). Faced with this monumental genomic task, it is more important to see how sound methodolgy could be tied to superb computational schemes to unfold the mysteries of the molecular activities in our body system. But, that way, we need to have the dusts settled before finding the right path for the desired scientific findings. Bioinformatics, at least at this stage, as a discipline, does not aim to lay down fundamental mathematical laws that govern biological systems parallel to those laid down in physics. There is, however, at this stage, mathematical utility in the creation of tools (mostly, in the form of computer graphics and algorithms) that investigators can use to analyse exceedingly large data sets arising in bioinformatics studies. Because of underlying stochastic evolutionary forces, such tools may generally involve statistical modeling of (molecular) biological systems, which in turn, requires incorporation of probability theory, statistics, and stochastic processes. Although *knowledge discovery and data mining* (KDDM) procedures dominate computational biology, and bioinformatics in general, from the above discussion we may gather that it would be improper to jump on statistical conclusions based on data analysis alone (even under the fancy umbrella of KDDM). Elements of *statistical learning* (Hastie et al. 2001) have therefore emerged as an endeavor to impart statistical reasoning to data mining; *bagging, boosting, bootstrapping* and *additive trees* are emerging as useful tools in KDDM explorations. Yet there is a genuine need to grasp the genetic and molecular biologic bases of bioinformatics to sort out the stochastic aspects from purely empirical computational aspects, and to lead to some resolutions that synchronize methodology and algorithms towards a common goal. Primarily driven by this motivation, we use the following terminology (Sen 2001): *Biostochastics to deal primarily with stochastic modeling and analysis (i.e., stochastics) for very large biological (including genetic and genomic) data sets.* In this formulation, biostochastics attempts to cover large biological systems which may not have predominant genetic factors; neuronal spatio-temporal models are noteworthy examples. In this scenario, we embark on an excursion of biostochastics from the traditional biometry to modern biostatistics to the evolutionary field of bioinformatics, without being confronted solely in the molecular genetics frontier. Yet identifying that molecular biology and large genetic models are important members, biostochastics is charged with the development of sound methodologic support for valuable computer intensive

algorithms that are currently widely used in genomic studies. It is in this sense somewhat complementary to statistical learning that places more emphasis on computational aspects.

Section 2 deals with an introduction to the basic need of statistical reasoning in bioinformatics including the budding field of pharmacogenomics. Section 3 is devoted to certain aspects of polygenic models where also statistical needs are genuine. Variants of likelihoods, such as the quasi-likelihood and pseudolikelihood, are discussed in Section 4. Section 5 deals with some nonparametrics. Some general remarks are appended in the last section.

2 Bioinformatics : a preamble

Given the background of bioinformatics evolution, and the relevance of statistical reasoning, let us start with the current state of art. Computer Science, in conjunction with information technology, seem to have stolen the limelight of current developments in this field, and KDDM is indeed a hot cake. While most of the adopted algorithms are advocated in the name (or game) of data mining, there appears to be a propensity, in this way, to choose a model from the observed data, and quite often, scores of possible models can be viewed as competitors. When a single model is thus chosen as the most appropriate one, there may be others that may give as good a fit. On the other hand, even when a model is specified on some other (mostly, subjective) grounds, it might not have been formulated with due statistical as well as biological motivations and justifications, or even, a true model may not exist. Thus, under such uncertainty, preliminary checking on *model selection* (as is needed in a computational scheme) merits critical statistical appraisal. Hence, statistical perspectives in KDDM in the context of bioinformatics can not be overlooked. A similar picture holds for statistical impacts in other aspects of bioinformatics, some of which are outlined below. It is in this spirit, we like to appraise statistical methodology that plays a basic role in bioinformatics.

Let us also comment on the impact and role of the triplet: *pharmacokinetics, pharmacogenetics* and *pharmacogenomics* in the evolutionary field of bioinformatics. Pharmacology is the science of drugs including materia medica, toxicology and therapeutics, dealing with the properties and reactions of drugs, especially with relation to their therapeutic values. In the same vein, pharmacodynamics is a branch of pharmacology dealing with reactions between drugs and living structures; pharmacokinetics relates to the study of the bodily absorption, distribution, metabolism, and excretion of drugs. It is no wonder that in many cases, the reaction of drugs on living organism may have genotoxicity; pharmacogenetics deals with the study of interrelation of

hereditary constitution and response to drugs.

Genome is the sum of all the genetic material in any organism; the precise sequence of the four component chemicals $\{A, C, G, T\}$ determines who we are as well as how we function. Each human cell has 46 chromosomes, 23 from each parent, and genes are precise sequences of $\{A, C, G, T\}$ arrayed at definite sites or *loci* on chromosomes. DNA is the carrier of genetic information; it is double helic model, made up of the four nucleotides where A pairs with T and G with C; A and G are purines while C and T are pyrimidines. Like the DNA, RNA and *proteins* are macromolecules of a cell, though differ in their forms and constitution; RNA is single-stranded (with T replaced by U), and proteins are complex chains of 20 amino acids that carry out tasks necessary for life, while *enzymes* are proteins that take other proteins apart or put them together. Both RNA and proteins are made from instructions in the DNA, and new DNA molecules are made from copying existing DNA molecules (in a process called synthesis). DNA has a nearly constant diameter with regularly spaced and repeated structures, irrespective of the base composition or the order of the four bases. Recent researches in human genome analysis have revealed that each gene produces, on an average, three proteins, and sometimes, as many as five. In this way, pharmacogenetics have paved the way for the evolution of pharmacogenomics.

The *Central Dogma* states that once (genetic) information passes into protein, it can not get out again. The transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein is possible, but transfer from protein to protein or protein to nucleic acid is not possible. The loop from DNA to DNA is called replication, from DNA to RNA is called transcription, and the loop from RNA to protein is called translation. RNA that is translated into protein is called messenger RNA or ${}_mRNA$, and the transfer RNA or ${}_tRNA$ translates the genetic code into amino acids. The central dogma has been extended in later years; *retroviruses* can copy their RNA genomes into DNA by a mechanism called *reverse transcription.*

If we accept the basic hypothesis that DNA is the blueprint for a living organism then it becomes natural to conclude that *molecular evolution* is directly related to changes in DNA; during the course of molecular evolution, *substitutions* occur. Recall that $A-T$ and $G-C$ are pairs formed by hydrogen bonds. As such, the substitution $A \leftrightarrow G$ or $C \leftrightarrow T$ are called transitions, while $A \leftrightarrow C$, $A \leftrightarrow T$, $G \leftrightarrow C$, $G \leftrightarrow T$ are called transversions. Thus, two purins or two pyrimidines are said to differ by a transition while a purine and a pyrimidine are said to differ by a transversion. Also, it may be noted that amino acids are encoded by triplets of nucleotides of DNA called *codons.* Let us define $\mathcal{N}_R = \{A, C, G, U\}$ as the set of nucleic acids, and let

$$\mathcal{C} = \{(x_1, x_2, x_3) : x_j \in \mathcal{N}_R,\ j = 1, 2, 3\} \tag{1}$$

be the codon. Finally, let $\mathcal{A}$ be the set of aminoacids and termination codons. Then the *genetic code* can be defined as a map:

$$g : \mathcal{C} \to \mathcal{A},\ g \in \mathcal{G}. \tag{2}$$

Thus, $\mathcal{G}$ is the set of all genetic codes.

Stochastic evolutionary forces act on genomes (molecular evolution). Probability models have been advocated recently by a host of researchers; we refer to Ewens and Grant (2001) for a up-to-date account. However, genes are not simple: the very high dimensionality and yet unknown nature of the battery of activities, specially in the evolutionary phase, have created an enormous task for molecular biologists and geneticists in the years to come. With the nearly completed picture of the human genome project, there are other formidable statistical tasks too. We shall appraise some of them later on.

3 Genetic undercurrents

The genome sequence is a classical illustration of a large molecular biological system with persistent genetic undercurrents. High-resolution linkage maps of genetic markers play a basic role in human genome project where an increasing number of molecular genetic markers are being identified, creating a genuine need for genetic maps to describe the location of genetic markers along chromosomes as well as centrosomes. Genetic researchers have recently reached a consensus that it takes just about 30,000 to 40,000 genes to make, maintain and repair a human being. *Gene- environment interaction* has emerged as one of the most active research areas. It is conceived that there are certain genes that are associated with disease phenotype, side-effects, and drug efficacy, so that pharmacogenomics occupies a focal position that deals with the study of variability in distribution, absorption, metabolism and reaction of humans to drugs which are geared to interact at the molecular level in relation to toxicologic and therapeutic impacts. Because of inherent variations and an enormously large number of genes as well as a very large pool of diseases and disorders, there is a genuine need of statistical methods to assess the *genetic mapping of disease genes.*

In that way, appropriate drugs can be designed to identify disease related susceptibility genes, study their mutation and mode of penetration, and thereby possible means of controlling or eliminating some target disease or

disorder. In spite of this being a very appealing motivation, it may not be that easy to secure the desired information network. This is why pharmacogenomics is stepping in in this complex genetic modeling and analysis. Usual genetic approach may not often be amenable to *high throughput screening.* In pharmacogenomics, data mining of sequences plays a basic role from operational as well as application points of view. Yet targets need much functional validation to relate genomics to a disease process; generally, this way they are amenable to high throughput screening. Another approach is the *functional validation* method wherein differential gene expressions, proteomics etc., are emphasized. Either way, taking into account the enormous variability in this biological process, as well as, the underlying stochastic evolutionary forces, it is clear that statistical reasoning is essential in this context.

The genes inherited by an offspring can be of differing *allelic* type; if there are only two possible types, the locus is termed a *diallelic* locus. The two genes at each locus determine the *genotype* at that locus. a person is *homozygote* or *heterozygote* according as the two allels are identical or not. The *phenotype* of an individual for a trait refers to the observable characteristics of the trait; the frequencies of outcomes for the various possible crosses are commonly termed in genetic analysis as *segregation ratios.* In contrast to loci on nonhomologous chromosomes, genes at two physically close loci on a common chromosome exhibit interference, and it has mostly been observed that the *recombination fraction* behaves as a monotone (but nonlinear) function of their physical distance. In this respect, the *law of linkage equilibrium* relates to those recombination fractions (under the usual regularity assumptions that underlie the classical Hardy-Wienberg equilibrium). In this respect, infinite population (without migration), discrete generations, random mating, and equal initial maternal and paternal genotype frequencies are assumed, albeit one or more of these assumptions are likely to be vitiated.

In the classical genetic models involving only a finite number of loci, the usual *Mendelian likelihoods* can be formulated under suitable regularity assumptions. Such likelihoods are based on three submodels: (i) Probability distribution of the parential genotypes, (ii) penetrances, defined as the likelihood of an observed phenotype, given an unobserved genotype, and (iii) transition probabilities depicting the flow of genetic information from the parents to an off-spring. Though such a formulation can be worked out for a small number of loci, the task becomes prohibitively laborious as the number of loci becomes large (as is the case in genomes). The main complication may arise due to linkage. In biometrical genetics, one has quantitative traits determined by a large number of loci; in a standard polygenic model, it is assumed that the loci are acting independently and additively, so that adopting a Mendelian

likelihood approach may lead to considerable conceptual as well as computational complications. There is also a need to incorporate environmental effects which may not interact in an additive way.

Identifying gene(s) that cause a specified disease is done by assigning the location of the gene(s) on a chromosome. In simple Mendelian inheritance models this has been possible by the so called *positional cloning* approach. Many of the common genes may have a genetic component, though their mode of inheritance may not be precisely known. Moreover, they are generally believed to be under the influence of multiple environmental and genetic factors. The genetic basis of a trait may be established by demonstrating that the trait runs in families; there may be, in addition to environmental effects, some infection effects which could lead to *familial aggregation* of affected individuals. Also, the genetic effects may be the consequence of a single gene to a Medenlian mode of inheritance, a small number of genes in *oligogenic inheritance*, or due to a large number of genes in a polygenic inheritance. It is usually very difficult to distinguish between the three situations, and as a result, genetic analysis is, by no means, very simple. Lange (1997) contains a very lucid treatise of currently available statistical tools for conventional genetic analysis, with due emphasis on Markov chain monte carlo (MCMC) methods, as well as, some interesting Poisson approximations. In genomics studies, some of the transitions may not be observable, so that usual Markov models may not apply well. *Hidden Markov models* (HMM) have been advocated (Durbin et al. 1998) to adjust for such incompleteness, though there remains some concern for the proper interpretation and appropriateness of such models, particularly with a large number of sites. At the wake of bioinformatics, genetics is undergoing an evolutionary phase wherin pharmacogenetics and pharmacogenomics are the most visible and dominant factors. In the light of these factors, we need to appraise the evolution in classical genetic analysis for its much needed adaptation in bioinformatics. Probability theory, stochastic processes and statistical reasoning are all indispensible in this context. However, the basic task requires incorporation of software development, process design, project management, protein structure prediction, disease specialities, and biotechnology. At the present the principal areas of focus include

(a) gene discovery (disease gene mapping),
(b) pharmacogenetics and pharmacogenomics,
(c) biological target identification and validation,
(d) biological network evolution, and
(e) chemical lead screening.

This bioinformatics venture relies heavily on computationally incentive statistical methodology in conjunction with statistical genetics. However, much of the molecular biological developments are taking place at a fast pace, and allowing the seeding time to absorb the fundamental outcomes to emerge in an interpretable and identifiable manner, statistical reasoning is somewhat lagging behind the (bio-)technological advances. The dominant computational aspects have tilted the bulk of statistical work load to computer science, and related KDDM, raising some serious questions on the validity of prescribed statistical approaches.

There is good evidence, acquired by now, that there are certain genes that are associated with disease phenotype, side-effects, and drug efficacy. On the other hand, the pool of diseases and disorders in human is quite large, albeit there are the subclasses, namely, dominant, recessive, and oligogenic families. Oligogenes with smaller effects in segregating families can mostly be detected by linkage, while polygenes have effects so small that linkage has little power to detect them. The simplicity of monogenetic traits vanishes for oligogenic and polygenic inheritance. Vast observational discovery of genomic phenomena has made it necessary to work on the whole genomes wherein comparison of gene families and data mining for disease candidates have received greater attention. Human geneticists having good evidence of a target gene proceed directly to genetic linkage analysis to localize disease genes. However, in view of the enormously large pool of genes, drug discovery and target evaluation could be a very delicate task, lacking complete scientific precision. Microarrays have swept the investigational routes for gene and protein expression analysis. In the analysis of microarray (gene chip) data sets, the raw data come in such a vast and unmanageable fashion that some preprocessing is needed to reduce the data set to a manageable form where the data mining and KDDM algorithms can be used to apply classification rule, cluster analysis, and interrelationship methodology. In this respect, the meta-analysis part in KDDM needs proper statistical appraisal before being accepted as a valid and efficient tool in such purely empirical data dredging exercise. Phylogenomics has stepped into the task of combining sequence information with attempts to eliminating seeming inconsistencies. Of course, superb computational abilities underlie all these developments.

4 Likelihoods and their limitations

Conventional likelihood function based statistical inference procedures, though appropriate, may encounter prohibitively laborious computational

complexities as well as lack of robustness prospects in large parameter space models. In genomics we typically have data sets on a large number (K) of sites or positions, where in each site, there is a purely qualitative (i.e., categorical) response with 4 to 20 categories depending on the DNA or the protein sequence. Neither these sites can be taken to be stochastically independent nor their spatial dependence or association pattern may be precisely known. Moreover, as typically K is large, there are roadblocks to implementing simple patterns in this complex setup. On the other hand, the embedded variability in these responses and the nearly identical structures of the DNA molecules suggest that alternative variational studies should be more appropriate from statistical modeling and analysis perspectives. For example, in the context of judging whether or not mutations at different sites take place independently of each other, consider a reduction in modeling based on the count of whether or not there is a mutation in position j at a given time, for $j = 1, \ldots, K$. If we let $Y_j = 1$ or 0, according as there is a mutation or not in position $j(= 1, \ldots, K)$, and define the stochastic vector $\mathbf{Y} = (Y_1, \ldots, Y_K)'$, then the probability law of $\mathbf{Y}$ is defined on $\mathbf{\Omega} = \{(i_1, \ldots, i_K) : i_j = 0, 1; j = 1, \ldots, K\}$ (whose cardinality is 2^K). Even for moderately large values of K, the number of parameters $(2^K - 1)$ associated with this probability law is quite large, and for such a large parameter space, conventional likelihood approaches stumble into computational as well as conceptual impasses. Replacement of the classical likelihood function by suitable conditional, partial, profile, pseudo or quasi-likelihood functions may generally lead to more severe nonrobustness properties of associated statistical tests and estimates.

If we let $P(\mathbf{y}) = P\{\mathbf{Y} = \mathbf{y}\}$, $\mathbf{y} \in \mathbf{\Omega}$, and define

$$Q(\mathbf{y}) = \log\{P(\mathbf{y})/P(\mathbf{0})\}, \ \mathbf{y} \in \mathbf{\Omega} \tag{3}$$

then, we obtain by routine computation that

$$P(\mathbf{y}) = e^{Q(\mathbf{y})} / \sum_{\mathbf{z} \in \mathbf{\Omega}} e^{Q(\mathbf{z})}. \tag{4}$$

One can then use the Bahadur (1961) representation of multivariate dichotomous random variables (see Liang, Zeger and Qaqish (1992) for further extensions), and write

$$P(\mathbf{y}) = \exp\Big\{ \sum_{k=1}^{K} u_k y_k + \sum_{1 \le s < k \le K} y_s y_k u_{sk} + \ldots + y_1 \cdots y_K u_{1 \cdots K} \Big\} \tag{5}$$

where the u_k are the conditional *logits*, u_{sk} are the conditional *log odd-ratio* etc. In this setup, if we assume further that the $u_{i_1 \cdots i_l}$ for $l \geq 3$ are all null, we end-up with a pairwise dependence model:

$$Q(\mathrm{y}) = \sum_{k=1}^{K} \alpha_k y_k + \sum_{1 \leq s < k \leq K} \gamma_{sk} y_s y_k, \tag{6}$$

wherein the α_k and γ_{sk} represent the main effect and first-order interactions.

This pairwise dependence model has been incorporated in a *pseudolikelihood function* approach, albeit with insufficient theoretical justifications without a conditional, partial, or even profile likelihood interpretation (Besag 1974). The specific form of this function (based on n independent sequences $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$) is given by

$$\prod_{i=1}^{n} \prod_{k=1}^{K} \left\{ \frac{e^{Y_{ik}(\alpha_k + \sum_{j=1}^{K} \gamma_{kj} Y_{ij})}}{1 + e^{Y_{ik}(\alpha_k + \sum_{j=1}^{K} \gamma_{kj} Y_{ij})}} \right\}; \tag{7}$$

this model is also termed the *autologistic model.* Although, highly computational incentive Markov chain monte carlo (MCMC) methods can be prescribed for finding the maximum pseudo-likelihood estimator (MPLE) of the associated parameters, their robustness and efficiency properties may not be usually tenable, and moreover, simulated likelihood ratio techniques using the Gibbs sampling or the Metropolis-Hastings algorithm may stumble into computational roadblocks when K is large (compared to n), as may be the case in the present context. We refer to Sen (2002) for some discussion.

As mentioned before, in genomic sequence analysis, we have sequences of data sets on a large number of sites, and we may like to know about their interrelations as well as possible lack of homogeneity over different groups of subjects; this is known as *computational sequence analysis* (CSA). Often, we need to test for homogeneity of $G(\geq 2)$ independent groups of sequences, each group having in turn a number of presumably independent sequences. For this external CSA problem, analogous to the classical multivariate analysis of variance (MANOVA) proble, several approaches have been advocated (Sen 2001). First, we may consider the classical likelihood ratio type test. But, in view of the unassessed nature of the total (or conditional, or partial) likelihood function, such a procedure is difficult to formulate. As such, we may consider the autologistic model, as described above, where for the gth group, we denote the associated parameters by $\theta_g = (\alpha_g', \gamma_g')'$, for $g = 1, \ldots, G$. Based on the PMLE of the θ_g one may consider then a Wald-type test. But, it could be quite cumbersome to obtain a good estimator of the (asymptotic) covariance matrices of the MPLE (specially when K is large), and this in turn may require

a very large sample size (compared to K and G) in order that asymptotics may yield reasonably good approximations. Moreover, no (asymptotic) optimality properties for such tests have been precisely formulated, nor they are likely to be true. For this reason, it might be more attractive to use suitable aligned scores statistics (Sen 2002), albeit in a permutation model to come up with more robust and reasonably simple tests for homogeneity. Empirical Bayes and hierarchical Bayes procedures have also been advocated. These procedures also depend on suitable likelihood formulations, and in addition, on the choice of appropriate priors on the associated parameters. There is still ample room for further developments in this direction.

As a second example, we consider a *quantitative trait loci* (QTL) model involving a (large) number of loci and quantitative phenotypic observable variables. It is typically assumed that the observable random variable (say, Y) depends on a number of loci with quantitative traits (say $Q_1, \ldots, Q_m$) and possibly under the surveillance of genetic markers at some of these loci, and in view of the multiplicity of the traits and markers, it is generally taken for granted that Y has a (mixed) normal distribution, given these extraneous variables; since some of these are not observable, there is a need to use suitable conditional normal laws, and on integrating on the unobservable variables one can then arrive at the appropriate likelihood function. No matter how we proceed, there could be a very large number of parameters associated with such a likelihood formulation. Although EM algorithm can be used for computational facilities, there remains the basic concern: For a large parameter space with (moderately large) n, number of observations, what could be said about properties of derived maximum likelihood estimators? Such estimates are generally not efficient (even asymptotically), are biased, and in some extreme cases, may even be inconsistent; the Neyman-Scott problem is a glaring example of this type. The same criticism may be labelled against the likelihood ratio (or allied Wald-type or Rao's scores) tests. A greater concern is how robust would be a likelihood based test or estimate for model departures (e.g., contamination by heavy tail distribution to assumed normal ones)? As additional examples, we may consider any other problem that crops up in testing for independence of mutations at multiple sites, genetic mapping of disease genes, gene-environment interaction, and other problems referred to in the preceding section. Although, for most of these genetic models, some (pseudo-)likelihood formulations have been advocated, they are generally far from being ideal from statistical modeling and analysis perspectives. For lack of space, we shall not enter into detailed discussions on individual problems; rather, in the next section, we shall discuss some alternative procedures that attach less emphasis on likelihood formulations and more on suitable non-

parametrics.

5 Whither nonparametrics

In order to illustrate the relative merits and demerits of likelihood based approaches and some alternative ones, let us consider the genomic (external) CSA problem, treated in the preceding section. Instead of binary response variables, treated there, we consider here a more general model that arises in genomics. We consider K positions or sites where at each position the response variable is purely qualitative with C possible outcomes, so that we have a full model involving C^K possible response vectors. Thus essentially, we are to test for the homogeneity of G high-dimensional contingency tables. In a single-site model, this reduces to a conventional C-category multinomial law, so that we have a (categorical) CATANOVA model, treated nicely by Light and Margolin (1971, 1974), and more elaborately in the genomic context by Pinheiro et al. (2000), where the *Hamming distance* in a simple formulation has been exploited. We consider here the general case, and formulate some alternative procedures.

Let $\mathbf{X}_i = (X_{i1}, \ldots, X_{iK})'$ be a random vector where the coordinate X_{ik} stands for the category outcome $c(= 1, \ldots, C)$ for the ith sequence at site k, $i = 1, \ldots, n$; $k = 1, \ldots, K$.Note that the responses are purely categorical variables, and hence, conventional norms or distances may not be usable here. For a pair (i, i') of sequences with responses $\mathbf{X}_i$ and $\mathbf{X}_{i'}$ respectively, the Hamming distance is defined as

$$D_{ii'} = \frac{1}{K} \sum_{k=1}^{K} I(X_{ik} \neq X_{i'k}). \tag{8}$$

This leads us to the sample measure

$$\bar{D}_n = \binom{n}{2}^{-1} \sum_{1 \leq i < i' \leq n} D_{ii'}, \tag{9}$$

which is a U-statistic (Hoeffding, 1948) with a kernel of degree 2. Note that this formulation does not assume that the coordinates of the $\mathbf{X}_i$ are all stochastically independent. Moreover, $\bar{D}_n$, being a U-statistic, unbiasedly estimates

$$\Delta_H = \frac{1}{K} \sum_{k=1}^{K} P\{X_{ik} \neq X_{i'k}\}$$

$$= \frac{1}{K} \sum_{k=1}^{K} I_{GS}^{(k)}, \tag{10}$$

where $I_{GS}^{(k)}$ is the Gini-Simpson index of biodiversity, as adapted in the present context (Pinheiro et al. 2000).

Consider now G independent groups, where the gth group consists of n_g independent sequences $\mathbf{X}_{gi}$, $i = 1, \ldots, n_g$, for $g = 1, \ldots, G$. We let $n = \sum_{g=1}^{G} n_g$. For the gth group we denote the sample and population Hamming distances as $\bar{D}_{n,g}$ and $\Delta_{H,g}$ respectively, for $g = 1, \ldots, G$. Also, we denote the pooled sample measure by $\bar{D}_n$ and its population counterpart as $\bar{\Delta}_H$. Basically, we are interested in testing the homogeneity of the $\Delta_{H,g}$ based on their sample counterparts $\bar{D}_{n,g}$ which are all U-statistics. Pinheiro et al. (2000) followed the conventional ANOVA approach based on U-statistics and their estimated variance-covariance, with the main emphasis on the partition of the total sum of squares into within and between groups components. Since in the present context, we do not have a (generalized) linear model, their suggested test procedure encounters some complex distributional problems, even under the hypothesis of homogeneity. Note that for each $k(= 1, \ldots, K)$ and $g(= 1, \ldots, G)$, $X_{gi}^{(k)}$ has a probability law on the C-simplex, and on top of that, for different k, the $X_{gi}^{(k)}$ are not necessarily independent. This renders some degeneracy in the null hypothesis distribution theory, and instead of anticipated asymptotic chi-squared distribution, we end up with more complex Cramér - von Mises type distributions involving a linear combination of independent chi-squared variables with one degree of freedom. It is to be noted that the pooled sample $\bar{D}_n$ can be decomposed into two nonnegative quantities, representing the between and within group components (see Chatterjee and Sen (2000) for some more general results), so that one may compare the between groups component with the within group component to draw statistical conclusions. We end up with similar Cramér - von Mises type distributions. We refer to Pinheiro et al. (2000) for some details.

Let us consider here two related tests based on the within group estimates $\tilde{G}_{n,g}$, $g = 1, \ldots, G)$, under appropriate regularity assumptions. First, we consider here a test for the homogeneity of the $\Delta_{H,g}$, in a little bit more stringent form of homogeneity of the G probability laws defined on the same probability space; in this setup, of course the alternative hypotheses relate to the part of the parameter space where the $\Delta_{H,g}$ are not the same. Recall that under this null hypothesis, all the G probability laws are the same (though unknown), so that conventional permutation tests can be formulated. We could generate all possible $N = \{(n!)/\prod_{g=1}^{G}(n_g)!\}$ partitioning of the n sequences (in the

pooled group) into G subsets of sizes $n_1, \ldots, n_G$ respectively, and under the null hypothesis, all these permutations are equally likely; we denote this conditional (permutational) probability measure by $\mathcal{P}_n$. It is then easy to verify that for each $g(= 1, \ldots, G)$,

$$E\{\bar{D}_{n,g}|\mathcal{P}_n\} = \bar{D}_n, \tag{11}$$

which is the same measure based on the pooled set,

$$Var\{\bar{D}_{n,g}|\mathcal{P}_n\} = \frac{4(n_g-2)(n-n_g)}{n_g(n_g-1)(n-3)}\zeta_{1,n} + \binom{n_g}{2}^{-1}\left[1 - \frac{\binom{n_g-2}{2}}{\binom{n-2}{2}}\right]\zeta_{2,n}, \tag{12}$$

and for $g \neq h(= 1, \ldots, G)$,

$$Cov\{\bar{D}_{n,g}, \bar{D}_{n,h}|\mathcal{P}_n\} = \frac{-4}{n-3}\zeta_{1,n} - \binom{n}{2}^{-1}\zeta_{2,n}, \tag{13}$$

where

$$\zeta_{1,n} = \frac{1}{n(n-1)(n-2)} \sum_{1 \leq i \neq j \neq l \leq n} D_{ij}D_{il} - \bar{D}_n^2,$$

$$\zeta_{2,n} = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} D_{ij}^2 - \bar{D}_n^2, \tag{14}$$

and the D_{ij} refer to the Hamming distance between the ith and jth observations in the pooled sample. A little more algebraic manipulations lead us to conclude that

$$Cov\{\bar{D}_{n,g}, \bar{D}_{n,h}|\mathcal{P}_n\} = \frac{4(n\delta_{gh}n_g)}{n_g(n-3)}\zeta_{1,n} + O(n_g^{-2}), \tag{15}$$

for $g, h = 1, \ldots, G$, where δ_{gh}, the Kronecker delta, is equal to 1 or 0 according as $g = h$ or not. All these results follow from the general theory of U-statistics under simple random sampling without replacement scheme (from the finite population of size n), and a detailed account of this computation and its approximations may be found in Nandi and Sen (1963) This permutation variance can be well approximated by its jackknifed version. Based on the above computation we consider a quadratic form in the $\bar{D}_{n,g}$ with their permutation means and variance-covariance terms, and arrive at the following test statistic:

$$\mathcal{L}_n^* = \Big[\sum_{g=1}^{G}(n_g - 1)\{\bar{D}_{n_g} - \bar{D}_n\}^2\Big]/4\zeta_{1,n}. \tag{16}$$

Note that $\zeta_{1,n}$, the pooled sequence estimator, is invariant under any permutation of the n sequences among themselves, and so is $\bar{D}_n$. On the other hand, for each of the N permuted set, we could compute the corresponding values of $\bar{D}_{n,g}, g = 1, \ldots, G$, and hence, generate a version of $\mathcal{L}_n^*$. With all these N realizations, one could order them and generate the exact permutation distribution of $\mathcal{L}_n$. Though this procedure works out well for small values of the n_g, it becomes prohibitively laborious as the n_g become large.

Fortunately, using the results of Sen (1981, Ch. 3), it is possible to verify that under the permutation law, asymptotically, $\mathcal{L}_n^*$ has the central chi square distribution with $G-1$ degrees of freedom (DF). This suggests that even if we do not use the permutation principle, the test statistic $\mathcal{L}_n^*$ can be used in an asymptotic setup for testing the homogeneity of the G groups with respect to the parameters Δ_{H_g}.

Next, we note that the $\Delta_{H,g}$ may be equal without requiring that the G probability laws are all the same. Hence, it may be more desirable to formulate suitable tests for the homogeneity of the $\Delta_{H,g}$ without assuming the homogeneity of the underlying probability laws. Since the within group measures $\bar{D}_{n,g}$ are all U-statistics, we could do it by estimating consistently the variance of each measure, and exploiting the asymptotic normality of the estimators $\bar{D}_{n,g}$. Toward this end, we let $\mathbf{X}_{gi} = (X_{gi}^{(1)}, \ldots, X_{gi}^{(K)})', \ i = 1, \ldots, n_g; g = 1, \ldots, G$. Let then

$$\pi_g(k,c) = P(X_{gi}^{(k)} = c); \ \pi_g(k,c;l,d) = P(X_{gi}^{(k)} = c, X_{gi}^{(l)} = d), \tag{17}$$

for $c(\neq d) = 1, \ldots, C, \ k(\neq l) = 1, \ldots, K, \ g = 1, \ldots, G$. Let then

$$\gamma_g^2 = \frac{1}{K^2} \sum_{k=1}^{K} \sum_{l=1}^{K} \sum_{c=1}^{C} \sum_{d=1}^{C} \pi_g(k,c)\pi_g(l,d)\{\pi_g(k,c;l,d) - \pi_g(k,c)\pi_g(l,d)\}, \tag{18}$$

for $g = 1, \ldots, G$. Then using the Hoeffding projection result, it follows (Sen 1981, Ch.3) that as n_g increases,

$$\sqrt{n_g}(\bar{D}_{n,g} - \Delta_{H,g}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 4\gamma_g^2), \tag{19}$$

for each $g(= 1, \ldots, G)$. Note that even under the null hypothesis, the γ_g^2 may not be all equal. However, for each g, we may estimate γ_g^2 by using the same formula as in $\zeta_{1,n}$ (in the pooled sample case, treated earlier) but solely using the n_g sequences in the group g; we denote these estimators by $\hat{\gamma}_g^2, \ g = 1, \ldots, G$. Let then

$$\hat{D}_n^* = \frac{\sum_{g=1}^G n_g \bar{D}_{n,g}/\hat{\gamma}_g^2}{\sum_{g=1}^G n_g/\hat{\gamma}_g^2}. \tag{20}$$

Let us then define

$$\mathcal{L}_n^{o*} = \sum_{g=1}^G \frac{n_g}{\hat{\gamma}_g^2}\{\bar{D}_{n,g} - \bar{D}_n^*\}^2. \tag{21}$$

It follows from the Cochran theorem (along with the Slutzky theorem) (cf. Sen and Singer 1993, Ch.3) that under the null hypothesis of homogeneity of the $\Delta_{H,g}$, $\mathcal{L}_n^{o*}$ has asymptotically chi square distribution with $G-1$ degrees of freedom (DF). This is then used in the determination of the asymptotic critical level for the test based on the test statistic $\mathcal{L}_n^{o*}$. Note that when the γ_g^2 are not all the same, the test statistic $\mathcal{L}_n^*$ may not have asymptotically chi-square distribution with $G-1$ DF (but a Cramér - von Mises type distribution), so that it might not have the correct (asymptotic) significance level, although both the tests will be consistent against any possible heterogeneity of the $\Delta_{H,g}$.

6 Concluding remarks

In the preceding section, we have considered the usual external CSA MANOVA problem and discussed suitable nonparametric procedures. If we would have considered some internal CSA problems, such as independence of mutations at different sites, a formulation of a suitable likelihood function could be in general quite complex. Karnoub et al. (1999) have considered some tests for the hypothesis of independence of mutations at a pair of sites, discussed the roadblocks for a conventional likelihood approach, and in the light of that, formulated a conditional likelihood approach that takes into account the consensus pair as a pivotal point, and leads to suitable binomial test. Whereas for a multivariate normal population, the concept of pairwise correlation has been successfully extended to canonical correlations, there are genuine impasses for such extensions in the present context; these are not only due to the categorical nature of the responses but also due to the fact that the number of sites is in general very large. This large dimension problem is also perceived in the neuronal spatio-temporal model where the simultaneous spike-trains relate to some multivariate counting processes (Sen 2001). Unlike the multinormal case, pairwise dependence measures do not characterize

the higher-factor interactions or dependence patterns, so that the classical canonical analysis may not generally apply to count data or qualitative data models. As of now, mapping of disease genes, microarray analysis, gene-environmental interaction etc., all relating to high dimensional qualitative data sets, rely heavily on computer graphics and biological intuitions. It is indeed a challenge to fathom out the stochastic evolutionary forces underlying these models, so that stochastic modeling and statistical analysis can be formulated in an objective manner. The scope for conventional likelihood approaches in these problems is often obscured by the biological undercurrents that may not as yet have a complete scientific explanation.

There has been a systematic development in statistical genetics dealing with quantitative traits (Lange 1997). Even if such quantitative trait models are conceived in bioinformatics, there remains clouds over the plausible dimensionality of the traits, as well as, the validity of the conventional multinormality assumption (sans which all statistical models in use at the present time may lose their rationality and validity). It may be quite appealing to adopt suitable location-scale family of distributions, and to examine, to what extent, the robustness of assumed normal models for such alternative situations. Again, the high dimensionality of the model arising in bioinformatics can create a complete impasse for such multinormality based statistical approaches. In this context too, there is a genuine need to explore alternative nonparametric approaches that are more viable under the biological and genetic setups. Formulation of such alternative approaches, though is a challenging task, should be given due consideration.

As has been referred to in Section 3, pharmacogenetics occupies a very focal point in bioinformatics, the main impetus being the tremendous scope for genomics in drug research and marketing perspectives. The principal difficulty in the implementation of phramacogenomics stems from the fact that experimental evidence to justify clinical conclusions often precludes human subject (due to medical ethics and basic clinical considerations). In the name of KDDM, simulation studies have therefore permeated the drug research arena. These ventures are mostly based on some algorithms, and often, without much statistical insights. In such a case, it might be wistful to conceive of simpler probabilistic models where such conventional algorithms can be validly and efficiently applied with full likelihood appreciation. It would be therefore more appropriate to conceive of alternative statistical approaches that put lesser emphasis on likelihoods and greater emphasis on alternative reasoning that adapts well to the biological and genetic explanations. Nonparametrics seems to have advantages in this setup, and we advocate the use of biologically motivated, genetics based, nonparametrics in bioinformatics.

It is a challenge, and with the steady flow of research in nonparametrics in all its horizons, the success should be in the reach of statisticians' theory and methodology basket.

REFERENCES

Bahadur, R. R. (1961). A representation of the joint distribution of responses to n dichotomous items. In *Studies in Item Analysis and Prediction* (ed. H. Solomon), 158-176. Calif.: Stanford Univ. Press.

Besag, J. (1974). Spatial interaction and the statistical analysis of life systems (with discussion). *J. Roy. Statist. Soc. B48*, 192-236.

Chatterjee, S. K. and Sen, P. K. (2000). On stochastic ordering and Gini-Simpson type poverty indexes. *Calcutta Statist. Assoc. Bull. 50*, 137-156.

Durbin, R., Eddy, S., Krogh, A. and Mitchison, G. (1988). *Biological Sequence Analysis: Probabilistic Models for Proteins and Nucleic Acids.* UK:Cambridge Univ. Press.

Ewens, W. J. and Grant, G. R. (2001). *Statistical Methods in Bioinformatics: An Introduction.* New York: Springer-Verlag.

Geyer, C. J. and Thompson, E. A. (1995). Annealing Markov chain monte carlo with applications to ancestral inference. *J. Amer. Statist. Assoc. 90*, 909-920.

Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer-Verlag.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* 19, 293-325.

International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

Karnoub, M. C., Seillier- Moiseiwitsch, F., and Sen, P. K. (1999). A conditional approach to the detection of correlated mutations. *Inst. Math. Statist. Lect. Notes Monogr. Ser.* 33, 221-235.

Lange, K. (1997). *Mathematical and Statistical Methods for Genetic Analysis.* New York: Springer-Verlag.

Liang, K., Zeger, S. L. and Qaqish, B. (1992). Multivariate regression analysis for categorical data (with discussion). *J. Roy. Statist. Soc. B* 54, 3-40.

Light, R. H. and Margilin, B. H. (1971). An analysis of variance for categorical data. *J. Amer. Statist. Assoc. 66*, 534-544.

Light, R. H. and Margolin, B. H. (1974). An analysis of variance for categorical data, II : Small sample comparisons with chi square and other

competitors. *J. Amer. Statist. Assoc. 69*, 755-764.

Nandi, H. K. and Sen, P. K. (1963). On the properties of U-statistics when the observations are not independent. Part Two: Unbised estimation of the parameters of a finite population. *Calcutta Statist. Assoc. Bull. 12*, 124-148.

Pinheiro, H., Seillier-Moiseiwitsch, F., Sen, P. K. and Eron, J. (2000). Genomic sequence analysis and quasi-multivariate CATANOVA. In *Handbook of Statistics, Volume 18: Bioenvironmental and Public Health Statistics* (eds. P. K. Sen and C. R. Rao), 713-74.6 Amsterdam: Elsevier.

Sen, P. K. (1981). *Sequential Nonparametrics: Invariance Principles and Statistical Inference.* New York: John Wiley.

Sen, P. K. (2001). *Excursions in Biostochastics: Biometry to Biostatics to Bioinformatics.* Lecture Notes, Inst. Statist. Sci. Academia Sinica, Taipei, Taiwan.

Sen, P. K. (2002). Computational sequence analysis: Genomics and statistical controversies. In *Recent Advances in Statistical Methods* (ed. Y. P. Chaubey), World Scientific Publ. in press.

Sen, P. K. and Singer, J. M. (1993). *Large Sample Methods in Statistics: An Introduction with Applications.* UK: Chapman-Hall.

Simpson, E. H. (1949). The measurement of diversity. *Nature* 163, 688.

Speed, T. and Waterman, M. S. (eds.) (1996). *Genetic Mapping and DNA Sequencing.* New York: Springer-Verlag.

Venter, J. C. et al. (2001). The sequence of the human genome. *Science 291*, 1304-1351.

Thompson, E. A. (1986). *Pedigree Analysis in Human Genetics.* Baltimore: Johns Hopkins Press.

Waterman, M. S. (1995) *Introduction to Computational Biology: Maps, Sequences and Genomes.* UK: Chapman-Hall.

Watson, J. D. and Crick, F. H. C. (1953). Genetical implications of the structure of deoxyribonucleic acid. *Nature 171*, 964-967.

MONITORING PAVEMENT CONSTRUCTION PROCESSES

ROSS SPARKS
CSIRO Mathematical & Information Sciences
Locked Bag 17, North Ryde, NSW 2113, Australia
E-mail: Ross.Sparks@csiro.au

JAMES OLLIS
Roads & Traffic Authority
Kembla Street, Wollongong, NSW 2500, Australia
E-mail: Jim_OLLIS@rta.nsw.gov.au

Monitoring processes with data collected from spatial systems is a common need in industry. This paper outlines an approach commonly used in geostatistics, namely Universal Kriging (Cressie, 1993), for modelling spatial trends. The fitted spatial models together with their standard errors are then used to establish control limits for monitoring changes in spatial trends. Also the QQ-plot and related tests are used to signal processes that are out-of-control. These methods are applied to the process of constructing concrete road pavements.

Keywords: statistical process control, universal kriging, control charts.

1 Introduction

The application presented in the paper involves monitoring the construction of concrete road pavements at the correct height and thickness. The Roads and Traffic Authority (RTA) have stated specifications for concrete pavement height because it is related to the desired design life of the pavement. Pavements constructed less than the design thickness have significantly reduced pavement life. Latest RTA research suggests that concrete pavements that are constructed 10 mm less than their design thickness will reduce the design life from 40 years to 18 years. This is more than 50% reduction in design life for a reduction of 5% in thickness. Construction engineers should therefore sets paving machines to construct pavement surfaces "slightly" higher than the design height. Their task is to control pavement thickness variation during the construction process to ensure the height does not go below the design height. Therefore, a critical quality characteristic is the difference between the constructed height and the design height of pavements.

The monitoring process differs slightly depending on whether you are the contractor or the RTA. Companies contracted to construct road pavements are

more concerned about controlling the construction process variation within a contract, therefore within pavement variation is critical. They have the potential to use real-time controlling measurements to feedback information that helps reduce variation in construction height. The RTA (the customer) treats a contracted pavement as a single sampling unit and is interested in controlling across pavement variation. They want to know which pavements are soundly constructed and meet specification. This paper is written primarily from the RTA's point of view, but it does provide useful information for contractors indicating changes in performance within a contracted pavement. Generally stakeholders are concerned about the average height, variation in height and local smoothness of pavements.

RTA has data on past constructed pavements. These data are used to define a reference or training pavement which has average smoothness, its mean construction height is close to that of the average of all pavements, and its variance in construction height is close to the average of that of all pavement variances. New constructed pavements will be compared to this training pavement. Comparisons presented in this paper are only valid if each future pavement is measured using a grid identical to that of the training pavement.

One approach to monitoring pavements is to present RTA decision makers with an image that will represent the quality of construction. This approach starts by fitting a Universal Kriging model to training data. This is used to establish the expected 'in-control' levels of the variation of construction height from the designed height. Then control limits for monitoring the construction process of new pavements are developed.

2 Universal Kriging

In this paper, we are interested in models that interpolate measurements Z (the differences between the measured height and the design height) onto a 2-dimensional surface defined by coordinates $s = (x, y)$ given data collected on a regular grid over a pavement. The Universal Kriging model is used for this purpose, that is, $Z(s)$ is modeled as follows

$$Z(s) = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 x.y + d(s) \tag{1}$$

where: $d(s)$ is a zero-mean random process with stationary, isotropic variogram, that is

$$\text{var}(Z(s) - Z(s')) = 2g(|h|) \tag{2}$$

and $|h|$ is the distance between two points s and s' in space, i.e., $h = s - s'$ (see Cressie, 1993).

Note that this error term $d(s)$ is spatially dependent and this is not equivalent to the random error term ε of a bivariate polynomial regression model below $Z(s) = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 x.y + \varepsilon$. For Universal Kriging, the predicted value of $d(s)$ is not zero as is the case for the random error term ε. Therefore, in most spatial settings, Universal Kriging (using model (1)) improves on the predictions established from bivariate polynomial regression model. The relative improvement depends of the strength of the spatial relationships in these error terms $d(s)$ which can be determined from the variogram model for $var(Z(s) - Z(s'))$. The fitted model is given by $\hat{Z}(s) = \hat{b}_0 + \hat{b}_1 x + \hat{b}_2 y + \hat{b}_3 x^2 + \hat{b}_4 y^2 + \hat{b}_5 x.y + \hat{d}(s)$, where $\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4, \hat{b}_5$,and $\hat{d}(s)$ are estimated parameters and values. Also, let τ be the standard error for this fit. The fitted value is significantly different from the mean of Z, μ_z , when either

$$\hat{Z}(s) > \mu_z + 3\tau(s) \quad \text{or} \quad \hat{Z}(s) < \mu_z - 3\tau(s). \tag{3}$$

These could be used as control limits for the contractor in controlling within pavement variation.

3 Sampling design and data considerations

Sampling was on a "regular" grid as indicated in Figure 1. The values recorded are the offset values (x) and chainage (y) as in (1). Recording the location of points as offset (across pavement distances) and chainage (down pavement distances) has effectively transformed the pavement into a rectangular space, and this has consequences when modelling. Large radius curves on RTA roads minimise distortion on the rectangular grid when compared with the required accuracy of the location of the sampling points. Therefore, for practical applications, it is assumed that the whole pavement is transferred into a rectangular grid, including where it is constructed around curves. In future, recording the values of sampling locations, as determined by survey measurements, would enhance the spatial information.

The differences between the constructed and design height did not differ significantly from a normal distribution and therefore at times in this paper we will assume a normal distribution. It is convenient if modelling of the variogram is carried out as a function of distance only, but before doing this we checked the anisotropic behaviour of the variogram. This was done visually by fitting the variogram in various directions down the pavement. A visual comparison of the variogram models for distances in various directions

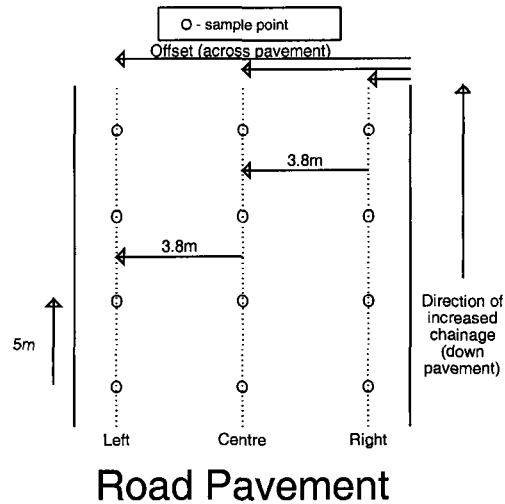


Figure 1. Sample grid for Road Pavements

indicated no sign of anisotropic behaviour for each data set considered in this paper.

This means that it is reasonable to assume that the variogram is only a function of distance points are apart, i.e., $\text{var}(Z(s) - Z(s^{'})) = 2g(|h|)$ where $h = s - s^{'}$. The fitted variogram model is necessary for fitting the Universal Kriging model. Fitted models were used to interpolate values onto a finer regular grid within the pavement. The spherical variogram model fitted well and it was used to model the semi-variogram of pavements in this paper.

4 Signalling unusual pavement construction

The design height for the pavement is the lower specification limit for the construction height, and the RTA's wishes to reject pavement sections with measured height below the design height. If contractors can produce a pavement with its height equal to the design height, with no variation, then they will reduce their cost of materials and thus maximise profit. Therefore, we signal whether a pavement remains within specifications but also simultaneously we signal when the construction process is out-of-control. Data were provided for a pavement that was classified as being in-control by experts at the RTA. Given this training data set we now wish to assess and signal problems with new pavement construction processes. In addition we wish to signal when a pavement has a mean and variance that is significantly different

Table 1. Assessing whether PAV1 or PAV2 have the same distribution as the training pavement (TPAV)

	PAV1	PAV2	Conclusion
Intercept($\hat{\mu}_0$)	−0.0008	−0.0055	Means for PAV1 and PAV2 are
Standard errors for $\hat{\mu}_0$	0.0002	0.0001	< that of the training TPAV
Slope($\hat{\sigma}_1$)	1.1147	0.7674	PAV1 varies more than TPAV
Standard errors for $\hat{\sigma}_1$	0.0101	0.0086	PAV2 varies less than TPAV

from the mean and variance of the training pavement.

The QQ-plot is used to compare distributions of construction heights for new pavements to the training sample. Let $\hat{Z}_q(s)$ and $T_q(s)$ be the qth quantile for the measured construction and designed heights for a pavement being monitored and training pavement, respectively. Once constructed, the whole pavement is judged to have the identical distribution to the training pavement if the relationship between the quantiles of both pavements is a simple linear regression model with slope equal to one and intercept equal to zero, i.e.,

$$\hat{Z}_q(s) = \mu_0 + \sigma_1 T_q(s) + e, \tag{4}$$

where $\mu_0 = 0$ and $\sigma_1 = 1$. If either μ_0 is significantly less than 0 or σ_1 is significantly greater than 1, then the pavement will be classified as significantly different (out-of-control) in terms of location or variation to the training pavement.

Two examples of new pavements are considered, denoted PAV1 and PAV2. Grid of measured values for these two pavements were identical to that of the training sample. The results of the fitted models for the quantiles are summarised in Table 1.

In Figure 2, notice that (at least for PAV1) the linear model is inadequate at describing the relationship between the quantiles. The tails of the distributions of PAV1 and training pavement differ. Several quantiles are below zero on both axes in Figure 2, thus raising concerns whether either the training pavement or test pavement are within specification.

The QQ-plot does give us a visual image of how the pavement is out-of-control and how badly it fails to meet the specification limit on construction height. However here it is worth noting that the pavement could be on average in-control (no signal using the QQ-plot), but locally out-of-control, e.g., all low pavement heights are measured in one region of the pavement. Therefore this approach fails to consider the spatial aspect of pavement construction.

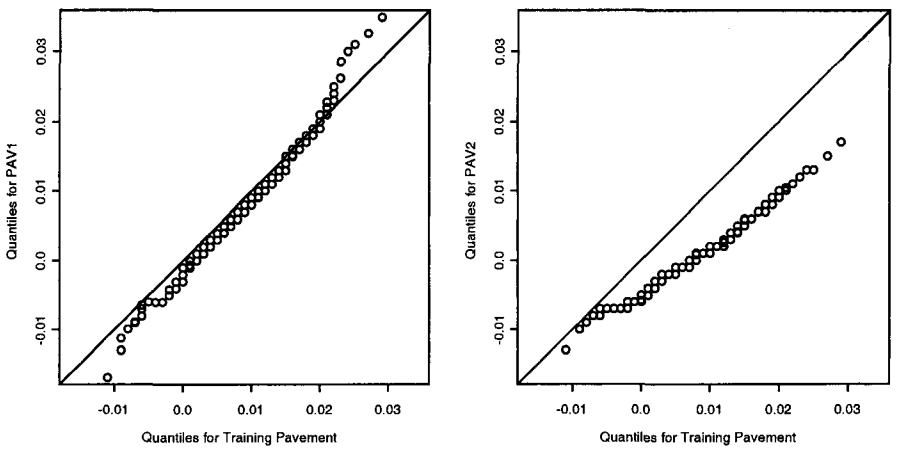


Figure 2. QQ-plot of the differences between pavement construction and design heights

Signaling local changes is location and variance

Non-overlapping groups of measurements (say three across and 12 down the pavement) could be used to signal local changes in mean and variance of the surface. This gives 36 sample heights. Sort these in order of magnitude, and for region G let these be defined as $Z_{[1]}(s), Z_{[2]}(s), ..., Z_{[36]}(s)$. Fit the model

$Z_{[j]}(s) = \alpha_0 + \alpha_1 T_{(j-0.5)/36}(\varsigma) + \varepsilon$ for $j = 2, 3, .., 35$

where ς is the whole training sample. The values for $j = 1$ and 36 are ignored to make the procedure robust to outliers. The fitted values $\hat{\alpha}_0$ and $\hat{\alpha}_1$and their estimated standard errors $\hat{\sigma}_{\alpha_0}$ and $\hat{\sigma}_{\alpha_1}$, respectively signal no change in location or variance if

$0 \in [\hat{\alpha}_0 - 3\hat{\sigma}_{\alpha_0}, \hat{\alpha}_0 + 3\hat{\sigma}_{\alpha_0}]$ or $1 \in [\hat{\alpha}_1 - 3\hat{\sigma}_{\alpha_1}, \hat{\alpha}_1 + 3\hat{\sigma}_{\alpha_1}]$

respectively. Otherwise a change is signalled. Note that this approach is distribution free.

Signaling changes in mean height

Note that $\hat{\alpha}_0 = 0$ indicates the construction height, on average, equals the design height. The centre line of the two graphs in Figure 3 indicates when $\hat{\alpha}_0 = 0$. If the estimated value of the intercept is below zero ($\hat{\alpha}_0 < 0$) for any part of that new pavement, then this part is estimated as having a lower mean than the training pavement. In Figure 3, the only parts of PAV2 that have a mean that is higher than the training pavement mean is chainage 46195 to 46555 (12% of PAV2); the remainder has sample means smaller than that of the training pavement. The estimated intercepts signal significantly

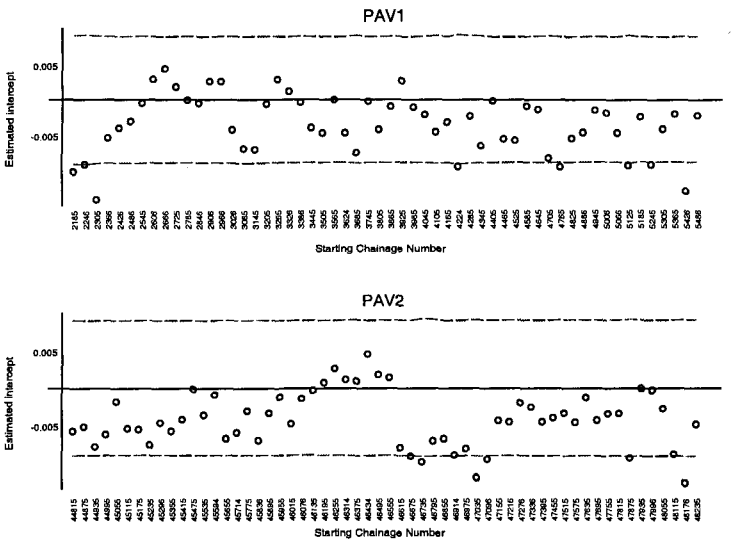


Figure 3. Control chart for signaling location changes

lower means than that of the training pavement on 5 occasions, (see groups with chainage numbers starting with 46735, 47035, 47096, 47875 and 48176 in Figure 3). PAV1 is estimated to have mean height greater than that of the training pavement for only 10 of the 56 groups (18% of PAV1), and it also gives 7 out-of-control signals on the low side. These occur at chainage starting numbers 2185, 2305, 4224, 4765, 5125, 5245 and 5426. PAV1's and PAV2's construction processes are out-of-control in several parts.

Signaling changes in variance

In Figure 4, PAV1 signals significantly higher variance than the training pavement variance for 21 of the 56 groups (38% of PAV1). PAV1 is often out-of-control in terms of its variance. PAV2 has only 4 groups that fail to signal lower variance than the training pavement variance. Therefore we can conclude that the control of variation in the height for this pavement is exceptionally good - significantly better than the training sample.

The local QQ-plots are good at signalling problems down the pavement but provides no information across the pavement. However an image plot will now be demonstrated as useful for signaling both changes down and across pavements thus giving an improved spatial perception of process control. The images Figures 5,6,7 & 9 of the paper are grey scale images, however if the

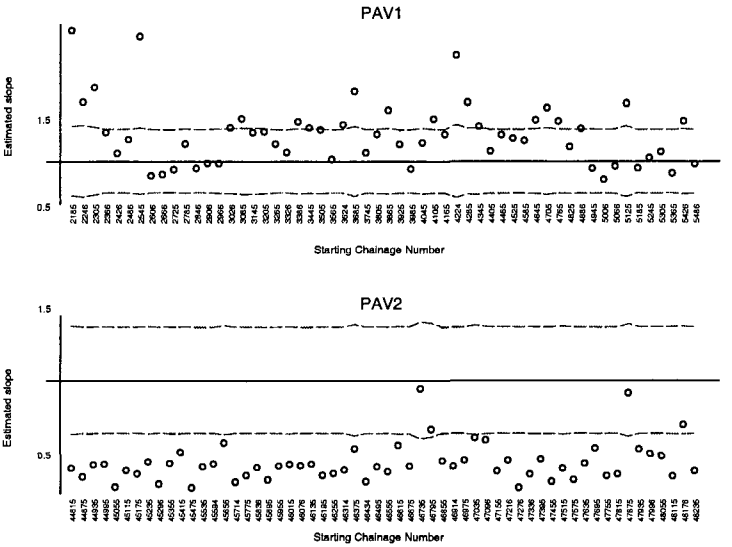


Figure 4. Control chart for signaling changes in variation

reader goes to web address www.cmis.csiro.au/Ross.Sparks (click on RTA paper) they can observe the colour images.

5 Image indicating pavement problem areas

A simple way of signaling whether a constructed pavement fails specification limits is to colour each pixel which corresponds to $\hat{Z}(s) < 0$ black. Pavements within specification will have images with no black pixels. The remaining pixels of the image are coloured as follows.

- Dark blue for pixels that corresponds to $\hat{Z}(s) > \mu_T + 3.\tau(s)$. This colour signals areas of the pavement pixels as out-of-control from the contractor's perspective (construction height is significantly higher than designed height).

- Red for pixels that corresponds to $0 < \hat{Z}(s) < \mu_T - 3.\tau(s)$. This colour signals areas of the pavement as out-of-control from the RTA perspective (construction height is significantly lower than designed height).

- Blue and orange for each pixel which corresponds to $\mu_T + 2.\tau(s) < \hat{Z}(s) < \mu_T + 3.\tau(s)$ or $\mu_T - 3.\tau(s) < \hat{Z}(s) < \mu_T - 2.\tau(s)$, respectively.

- Light blue and yellow for each which corresponds to $\mu_T + \tau(s) < \hat{Z}(s) < \mu_T + 2.\tau(s)$ or $\mu_T - 2.\tau(s) < \hat{Z}(s) < \mu_T - \tau(s)$, respectively.
- Very light blue for each pixel which corresponds to $\mu_T - \tau(s) < \hat{Z}(s) < \mu_T + \tau(s)$.

Note that new pavements should have 2.1% of pixels coloured blue, 2.1% coloured orange, 0.3% coloured dark blue and 0.3% coloured red, 13.6% coloured light blue, 13.6% coloured yellow, and 66% very light blue to be equivalent to the training pavement is distribution.

6 Interpreting the image plots

Once all pixels are coloured then the resultant image displays information on the constructed pavement. Some examples are presented below.

In-control pavements: Examples from the RTA's perspective follow.

1. A pavement that has near all pixels very light in colour in its image, e.g., with roughly half the pixels light yellow and the remainder very light blue, is very much in-control.
2. A pavement with colours nearly all pixels blue or light blue.

In-control pavements from the contractor's perspective is a pavement with pixel colours nearly all very orange, yellow, and light blue, but with no pixels red or black.

Out-of-control pavements: Examples from the RTA's perspective follow:

1. *Location shifts.* A pavement that have pixels with colours significantly more light yellow, orange and red colours than expected, will be signaled as out-of-control by producing pavements significantly lower than expected.
2. *Increased variation.* A pavement having significantly more dark blue, blue, orange and red colored pixels than expected, has variation larger than expected.

The resulting images for the training data, and the two pavements being monitored (PAV1 and PAV2) are included in Figures 5, 6 and 7 respectively. Notice that the training pavement is not always within specification, and

down the left-hand side it does not appear to be in-control (top row of pixels in Figure 5 has 7.9% black and 16.6% red coloured pixels). This was not known to the RTA prior to this analysis. PAV1 has lower construction heights on average than the training pavement by having less blue pixels and relatively more shades of orange pixels in Figure 6 relative to Figure 5. Also from chainage number higher than 5000 the pavement appears both out-of-control(most pixels are black, red or a shade of orange) and not in specification (several pixels are black). Notice that this pavement is more out-of-control down the left-hand side (top row of pixels in Figure 5 has 18.8% black and 26.9% red coloured pixels). PAV2 is clearly not within specification and it is out-of-control by having nearly no pixels shaded blue.

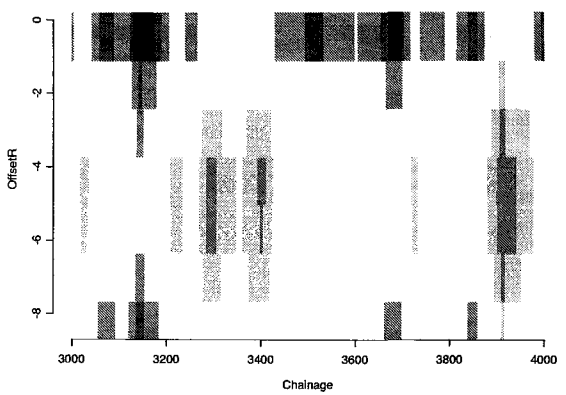


Figure 5. The image control plots from the fitted Universal Kriged interpolations for the training pavement

The image plot discussed so far present visual information on large scale pavement variation. However, we now explore procedures that will display and control this better.

7 Signaling problem with pavement smoothness

The surface level of the pavement could wonder locally but still on average have good distributional properties as measured by the QQ-plot. The image plot in Figure 8 gives some information on smoothness, but it is only able to

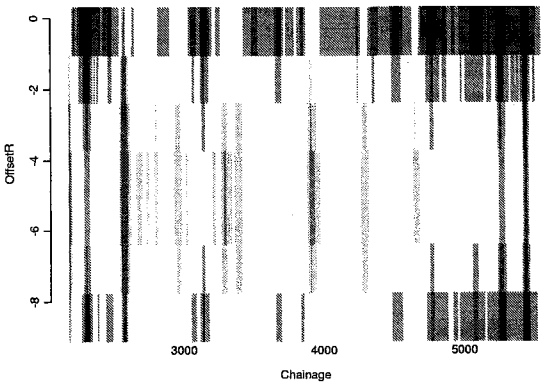


Figure 6. The image control plots from the fitted Universal Kriged interpolations for PAV1

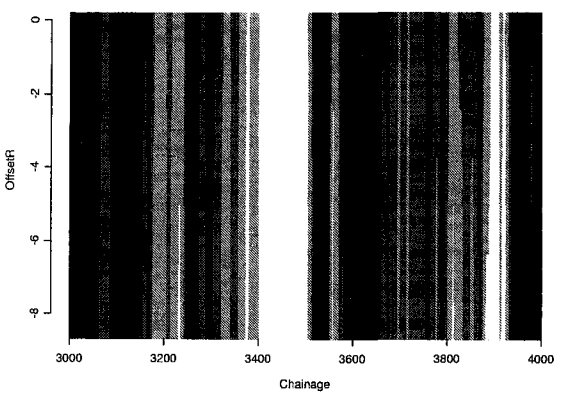


Figure 7. The image control plots from the fitted Universal Kriged interpolations for PAV2

detect large discrepancies in smoothness. Therefore, we wish to have a method that efficiently but simply measures how the pavement changes locally in height. This will be achieved by looking at the distribution of local differences.

Across the pavement variation: This variation can be examined by looking at differences between left-hand side and centre measurements, and centre and right-hand side measurements. The simultaneous plot of these differences versus chainage is useful for assessing across pavement variation as moving along the chainage values. Three standard deviation control limits can be placed on these differences to signal unacceptable variation (red dashed lines).

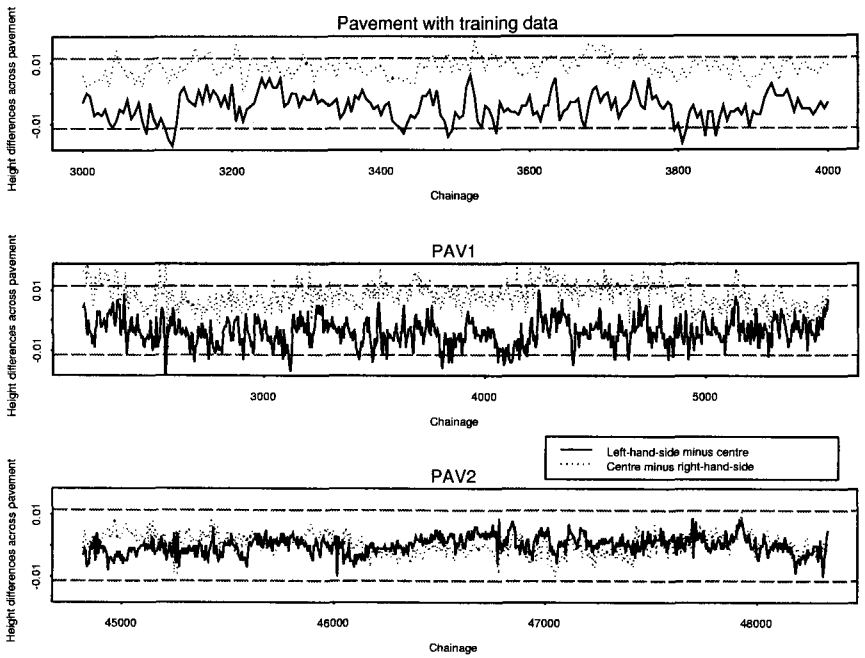


Figure 8. Across pavement differences versus chainage scatterplot

Figure 8 indicates that the training data is higher in the centre than on either of the sides of the pavement. The similar is true of PAV1 but this is not evident for PAV2. The heights across PAV2 is more consistent than the other two pavements.

Down the pavement variation: Here we take adjacent measurement pairwise

differences down the chainage value of the pavement, but here we present the information as a image. We could fit a Universal Kriging model to these differences, however the empirical variogram indicates near spatial randomness and therefore we let σ_D be the standard deviation of all down pavement adjacent differences for the training data. Establish the quantlies $q_{i,20}$ for probabilities $\frac{i}{20}$ for $i = 1, 5, 15, 19$. The pixels for new pavements corresponding to differences that are:

- smaller than $q_{1,20}$ are coloured purple,
- larger than $q_{1,20}$ but less than $q_{15,20}$ are coloured light purple,
- larger than $q_{15,20}$ but less than $q_{19,20}$ are coloured light green,
- larger than $q_{19,20}$ are coloured green,

while all others pixels are not coloured. Pavement that correspond to image plots with mostly light colours are relatively smooth compared with the training pavement. Images with many more darker colours than expected indicate pavements that are rougher than the training pavement.

Figure 9 compares PAV2 and the training data for smoothness. PAV2 appears significantly smoother than the training pavement by having much lighter colours. Testing whether this is statistically significant is achieved by using tests similar to earlier tests relating to the QQ-plot. In this situation, however, we examine the QQ-plot of adjacent differences as we move down the pavement. Figure 10 illustrates tests for comparing PAV2 to the training pavement. Notice that there is very little difference in the conclusions for the Left, Centre and Right hand-side of the pavement, even when there is a substantial outlier in the regression sense. This confirms what was already evident from the image in Figure 9, that PAV2 is significantly smoother than the training pavement.

An alternative to this approach is to perform a multivariate EWMA chart on adjacent differences moving down the pavement (see Lowry et al., 1992, and Sparks, 1992). However, this will not be explored here because the solution supplied is very visual, intuitive and informative.

8 Conclusion

We have illustrated how to signal whether the construction process of a pavement is simultaneously in-control and within specification, particularly when a reference data set is made available. The control process looks at both

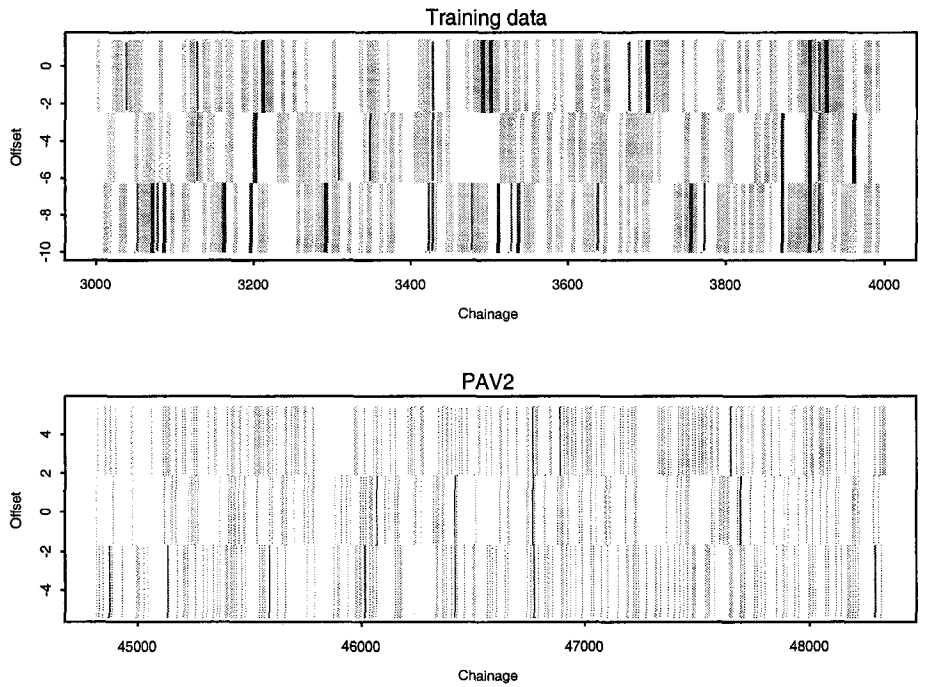


Figure 9. Along pavement differences versus chainage image plot

overall construction performance (long range variation) and local pavement performance (short range variation). Also an image representation of height variation is used to supplement these methods to give a visual image of where things are going wrong. It was clear that problems are encountered on the edges rather than in the centre of the pavement. This is very useful information for trying to correct pavement construction problems. The approach used is very different from the classical approach, and is fairly generic because in most cases it fails to rely on any distributional theory or spatial correlation structure. Therefore it has wide applicability.

We also have distinguished between control from the perspective of the contractor, who is more interested in control within a pavement, and the perspective of RTA (customer) is more interested in control pavement to pavement variation.

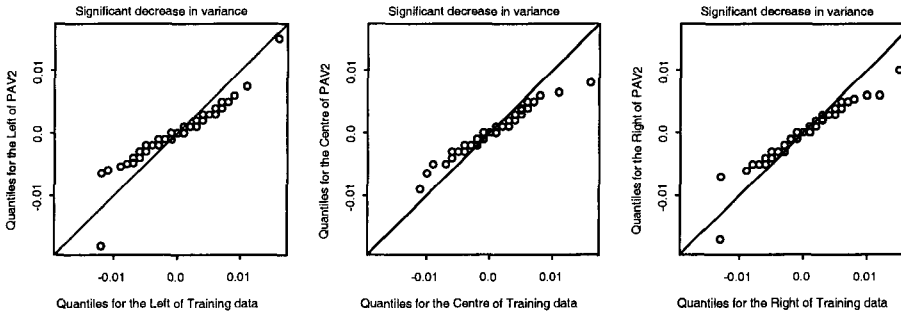


Figure 10. QQ-plot for testing for smoothness of the pavement

REFERENCES

Cressie, N.A. (1993). *Statistics for Spatial Data.* New York: John Wiley & Sons.

Lowry, C.A., Woodall, W.H., Champ, C.W., and Rigdon, S.E. (1992). A Multivariate Exponentially Weighted Moving Average Control Chart. *Technometrics 34, 46-53.*

Montgomery, D.C. (1991). *Introduction to Statistical Quality Control, 2nd ed.* New York: John Wiley & Sons.

Ollis, J.E. (1997). *Statistical tolerance for concrete road pavement surfaces,* Unpublished Masters of Total Quality Management (Honours) thesis at the University of Wollongong.

Sparks, R.S. (1992). Quality control with multivariate data. *Austral. J. Statist.* 34, 375-390.

IN SEARCH OF A GOOD ROTATION PATTERN

D.G. STEEL
Institute for Mathematical Modelling and Computational Systems
University of Wollongong
Wollongong NSW 2522, Australia
E-mail: dsteel@uow.edu.au

C.H. MCLAREN
Australian Bureau of Statistics
Belconnen ACT 2617, Australia
E-mail: craig.mclaren@abs.gov.au

A key element in the design of a repeated sample survey is the rotation pattern, which affects the variability of the time series of survey estimates and the seasonally adjusted and trend estimates produced from them. This paper considers the choice of rotation pattern for seasonally adjusted and trend estimates obtained from a repeated survey, using X11 based methods.

Keywords: sampling error, seasonal adjustment, repeated surveys, rotation pattern, trend, X11.

1 Introduction

Many time series are based on sample surveys repeated over time. A key element in the design of a repeated survey is the rotation pattern, which is the pattern of units' inclusion in the survey over time. The rotation pattern affects the degree of sample overlap and hence the correlation between the sampling errors at different lags.

Seasonally adjusted and trend estimates can be produced from the original estimates (ABS, 1993). A widely used seasonal adjustment package is the X11 program (Shiskin *et al.*, 1967) and its variants X11ARIMA (Dagum, 1988) and X12ARIMA (Findley *et al.*, 1998). Since seasonally adjusted and trend estimates are obtained from the original series, they are also influenced by the sampling error and therefore the rotation pattern.

Some repeated series are based on independent samples at each time period, but usually the samples used have a degree of overlap from period to period to reduce costs and the standard errors associated with the estimates of change between consecutive time periods (Kish, 1998). The rotation patterns currently used have been designed for estimates of the level and/or movements

in the original and seasonally adjusted series (Binder and Hidiroglou, 1988).

Rotation patterns vary in the number of times a unit is included in the survey and the time interval between inclusions. A general class of rotation patterns is defined by selected units being included in the survey for a consecutive months, removed for b months and included again for a further a months. The pattern is repeated so that units are included for a total of c occasions. These rotation schemes are denoted a-b-$a(c)$ (Rao and Graham, 1964). Setting $b = 0$ gives an *in-for-c* rotation pattern. Examples include the U.S. Current Population Survey, which uses a 4-8-4(8) rotation pattern (Fuller *et al.*, 1992). The Australian Monthly Labour Force Survey(MLFS) uses an *in-for*-8 rotation pattern (ABS, 1998) and the Canadian MLFS uses an *in-for*-6 rotation pattern (Singh *et al* 2001). These designs produce high sample overlap for consecutive months, reducing the sampling errrors of estimates of monthly change in the original and seasonally adjusted series. However, Mclaren and Steel (2000) suggested that rotation patterns with no monthly overlap were better for trend estimation.

McLaren and Steel (2000) considered how the design of the rotation pattern affected the sampling variance of the level and one-month change in seasonally adjusted and trend estimates, for a selection of possible rotation patterns. This paper considers a more extensive search over different rotation patterns for the sampling variance of the level, month-to-month and three-month movements of the seasonally adjusted and trend estimates. It also shows how the effect of a rotation pattern is equivalent to the effect of a filter applied to a series obtained from an independent design.

2 Filter representation of a rotation pattern

The seasonally adjusted and trend estimates obtained from X11 can be expressed, approximately, as the result of applying linear filters to the series (Dagum *et al.* 1996, McLaren and Steel, 2000). Consider a filtered value at time t

$$\tilde{y}_t = \mathbf{w}_t' \mathbf{y}_T$$

where $\mathbf{y}_T$ is a vector of estimates up to time $T \geq t$, obtained from a repeated sample survey and $\mathbf{w}_t$ is a vector of weights. The corresponding series of true values is $\mathbf{Y}_T$ and $\tilde{Y}_t = \mathbf{w}_t' \mathbf{Y}_T$ is the filtered value that would be obtained if there were no sampling errors. The weights can be chosen to produce estimates of the level and change in seasonally adjusted and trend estimates.

Let $V(\mathbf{y}_T|\mathbf{Y}_T)$ be the covariance matrix of the sampling errors associated with $\mathbf{y}_T$. Different rotation patterns produce different $V(\mathbf{y}_T|\mathbf{Y}_T)$. The

sampling variance of the filtered value at time t is

$$V(\tilde{y}_t) = \mathbf{w}_t' V(\mathbf{y}_T|\mathbf{Y}_T)\mathbf{w}_t \tag{1}$$

One way to understand the impact of different rotation patterns is to consider how they produce a correlation structure that is equivalent to the application of a weight matrix, $\mathbf{B}$, to an independent design. Using this concept, linear filters can be found that approximate the impact of the chosen rotation patterns.

Consider two applications of (1)

$$V^0(\tilde{y}_t) = \mathbf{w}_t' \mathbf{V}^0 \mathbf{w}_t$$
$$V^1(\tilde{y}_t) = \mathbf{w}_t' \mathbf{V}^1 \mathbf{w}_t$$

where $\mathbf{V}^0$ is the variance matrix for an independent design, and $\mathbf{V}^1$ is the variance matrix for a more complex rotation pattern. These matrices are symmetric and of full rank and can be decomposed using the Cholesky decomposition

$$\mathbf{V}^0 = \mathbf{A}_0' \mathbf{A}_0 \qquad \mathbf{V}^1 = \mathbf{A}_1' \mathbf{A}_1 \tag{2}$$

For an independent design $\mathbf{A}_0 = \text{diag}(\sqrt{V(y_t)})$.

Define $\mathbf{B} = \mathbf{A}_0^{-1}\mathbf{A}_1$ then $\mathbf{V}^1 = \mathbf{B}'\mathbf{V}^0\mathbf{B}$. The matrix $\mathbf{B}$ transforms the variance of an independent design to the variance for a specified rotation pattern. Each column of $\mathbf{B}$ contains weights which approximate the effect of the chosen rotation pattern. That is, it represents the results of an application of a filter to a series with an independent error structure.

A gain function $G(x)$ can then be plotted for the filter corresponding to a column in $\mathbf{B}$. Gain functions can be used to look at the effect of a linear filter on the amplitude of a cycle in a given series at a given frequency. The gain function associated with a filter shows those frequencies which are reduced ($G(x) < 1$) and those which are amplified ($G(x) > 1$) by the application of the filter. Frequencies (x) can be converted to months using the relationship, $1/x$ for $x \neq 0$. All columns of $\mathbf{B}$ will exhibit similar gain function properties.

Figures 1(a) to 1(f) give average gain functions for linear filters representing selected rotation patterns. A correlation model for the sampling error of employment estimates given in McLaren and Steel (2000) was used. For example, Figure 1(b) is obtained by choosing a column from $\mathbf{B}$ after setting $\mathbf{V}^1$ to be a variance matrix for a 1-2-1(5) rotation pattern and then calculating the gain function. This is done for each column of $\mathbf{B}$ and the average gain plotted. This shows that, by using this type of rotation pattern, cycles of approximate length three are produced. This is because a 1-2-1(c) rotation pattern induces correlations 3 months apart. This rotation pattern has

little effect on the longer term cycles, but reduces the effect of cycles in the 4 to 12 month band. Figure 1(c) shows that filters approximating the 1-1-1(6) rotation pattern have a similar gain function with a peak at a cycle of 2 months.

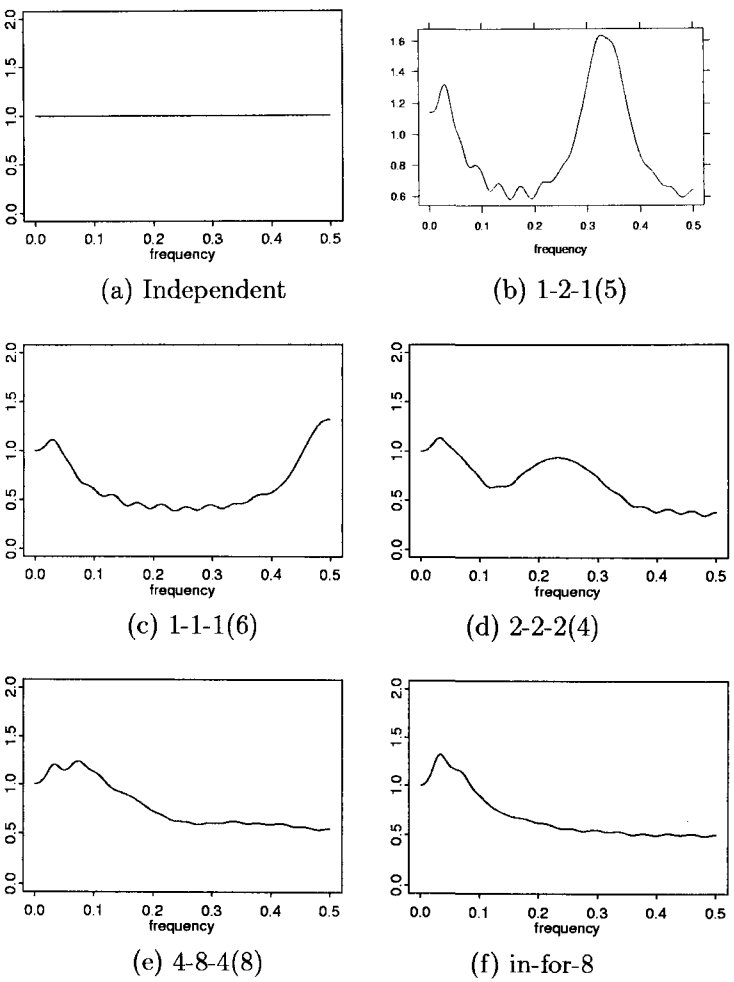


Figure 1. Gain function representation of different rotation patterns for proportion employed (MLFS)

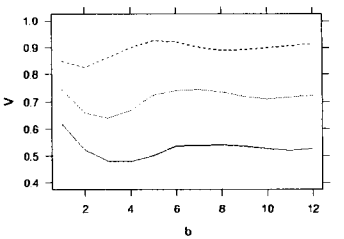
Figure 1(e) is the gain function of a filter representing a 4-8-4(8) rotation pattern. In this case cycles of length 9 months and onwards are increased in amplitude, whereas cycles from approximately 2 to 6 months are reduced. Filters approximating the rotation patterns *in-for*-8, 2-10-2(4), 3-3-3(6), and 6-6-6(12) have similar gain functions.

This filter representation of rotation patterns explains the impact that they have on the variance obtained by the subsequent application a different filter, $\mathbf{w}_t$, to obtain seasonally adjusted or trend estimates. Consider the 1-2-1(5) rotation pattern and $\mathbf{w}_t$ corresponding to a trend filter. Trend filters reduce cycles of random noise (typically 1 to 3 months) and allow cycles indicative of the trend to pass through the filter unchanged (typically 12 months onwards). The 1-2-1(5) pattern increases cycles of 3 months, where the trend level filter reduces these cycles. The 1-2-1(5) pattern reduces cycles from approximately 4 to 12 months where the effect of the trend level filter is to amplify these cycles. The overall effect is to allow these longer term cycles to pass through relatively unchanged. In contrast, the 4-8-4(8) rotation pattern reduces cycles around 2 to 10 months, where the trend level filter has already reduced cycles from 2 to 6 months. The 4-8-4(8) actually *increases* cycles from the 10 month period on top of an *increase* by the trend level filter. It is a combination of the 1-2-1(5) rotation pattern slightly reducing longer term cycles and the 4-8-4(8) type rotation patterns increasing the longer term cycles that leads the 1-2-1(5), and rotation patterns with similar gain functions, to be superior for reducing the variance of the trend level estimates.

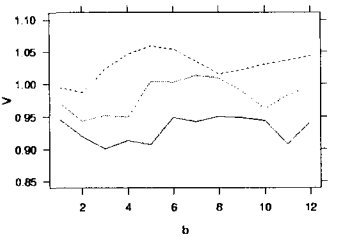
3 Effects of a-b-$a(c)$ rotation patterns on sampling variance

McLaren and Steel (2000) suggested that rotation patterns such as 1-2-1(8) out-perform other widely used rotation patterns in reducing the variance of the estimates of the level and three-month difference in seasonally adjusted estimates and the level, one and three-month difference in trend estimates. However, only a selected range of rotation patterns was considered. Correlation models for the sampling errors for employment and unemployment were used. To examine the impact of the gap between inclusion in the survey, a systematic search is conducted here for the 1-b-1(8), 2-b-2(8) and 4-b-4(8) rotation patterns. Figures 2(a) to 2(b) show the resulting variances for the estimates of level and one and three-month difference in the seasonally adjusted and trend estimates for the variable employment. Figures 3(a) to 3(b) give the results of a search over values of b for 1-b-1(m) rotation patterns where $b = 1, \ldots, 12$ and $c = 2, \ldots, 12$. This will show whether $b = 2$ is the best option. An additive linear approximation to X11 was used with no outliers

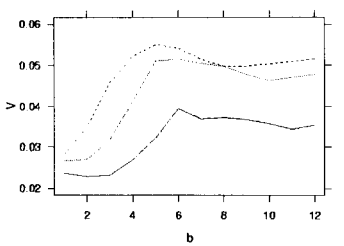
(Dagum et al., 1996). The trend estimates considered were those at the end of the series, $t = T$. Similar results were obtained for unemployed persons.



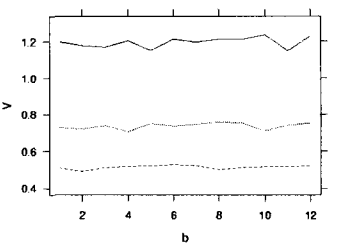
(a) Trend level



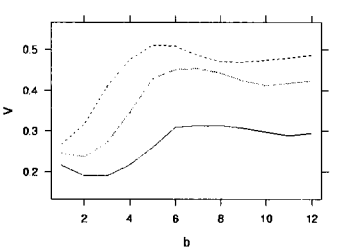
(b) Seas. adj. level



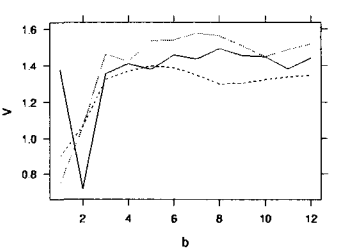
(c) One month difference in trend



(d) One month difference in seas. adj.



(e) Three month difference in trend



(f) Three month difference in seas. adj.

Figure 2. Variances for 1-b-1(8) (——), 2-b-2(8) (· · ·) and 4-b-4(8) (- - - -) rotation patterns for proportion employed (MLFS).

Figures 2(a), 2(c) and 2(e) show that for the level, one and three-month differences for trend estimates, the 1-b-1(8) rotation patterns out-perform the 2-b-2(8) and 4-b-4(8) rotation patterns for all choices of b. The choice of $b = 2$

or 3 provides the minimum for each type of rotation pattern.

Figures 2(b), 2(d) and 2(f) show that for seasonally adjusted estimates the 1-b-1(8) rotation pattern performs well for both level and three-month differences. The 4-b-4(8) rotation patterns are good for reducing variance of the one month difference as they have a high sample overlap from month to month. The 4-8-4(8), 2-10-2(8), and 1-11-1(8) patterns all induce correlation at lag 12 which Figure 2(b) shows is beneficial when looking at the level of seasonally adjusted estimates.

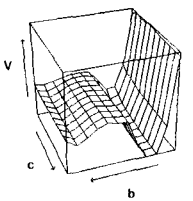
For the 1-b-1(c) rotation patterns, Figure 3(a) shows that for trend level the sampling variance is reduced by choosing $b = 3$ or 4 with c as large as possible. Figures 3(c) and 3(e) show that for the one and three-month difference of the trend estimates the 1-2-1(c) and 1-3-1(c) patterns produce the lowest variances, with no noticeable difference between the two. There is a marked increase in sampling variance for $b \geq 4$. These results suggest that setting $b = 2$ or 3 is best for trend estimation.

Figure 3(b) shows the effect on the level of the seasonally adjusted estimates. A choice of $b = 3$ provides greater gains along with c as large as possible. The advantage of correlations at lags 4, 6 and 12, which correspond to seasonal peaks, is evident for choices of $b = 3, 5$ and 11. Results for the one-month difference of the seasonally adjusted shown in Figure 3(d). This illustrates the effect of having the sample overlap at a seasonal frequency, ie. 1-11-1(8) and 1-5-1(8) which put correlations at lag 12 and 6 respectively. Figure 3(b) shows that for the three month difference, 1-2-1(c) produces the lowest variances.

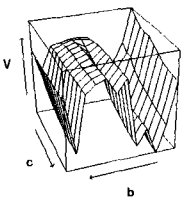
4 Trade-off between seasonally adjusted and trend estimates

The results suggest that the use of 1-b-1(c) with $b = 2$, or 3 is a good choice, except for the one-month change in the seasonally adjusted estimates. The trade-off in variance between designing for the level and one-month difference in trend estimates and the level and one-month difference in seasonally adjusted estimates is considered by presenting the results as scatterplots in Figures 4(a) to 4(d). On each Figure the o's represent an *in-for-c* rotation pattern for $c = 1, 2, 3, \ldots$ and 1 represents 1-2-1(5), 2 represents 1-2-1(8), 3 represents 1-1-1(6), 4 represents 2-10-2(4), 5 represents 2-2-2(8), 6 represents 3-3-3(6), 7 represents 4-8-4(8) and 8 represents 6-6-6(12).

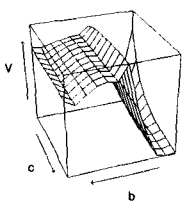
Figure 4(b) plots the variance of trend level estimates against the seasonally adjusted level estimates for different rotation patterns and shows that 1-2-1(8) gives the lowest variances for both. As overlap increases the variances of both trend level and seasonally adjusted level estimates increase.



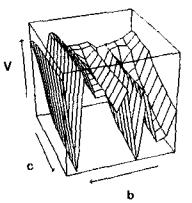
(a) Trend level



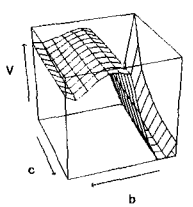
(b) Seas. adj. level



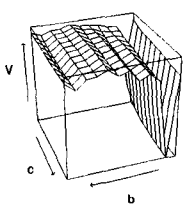
(c) One month difference in trend



(d) One month difference in seas. adj.



(e) Three month difference in trend

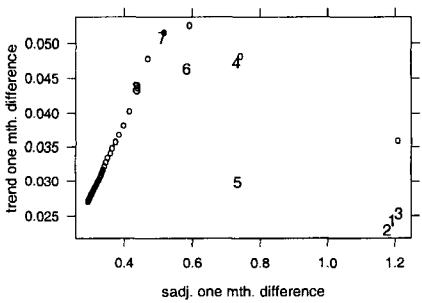


(f) Three month difference in seas. adj.

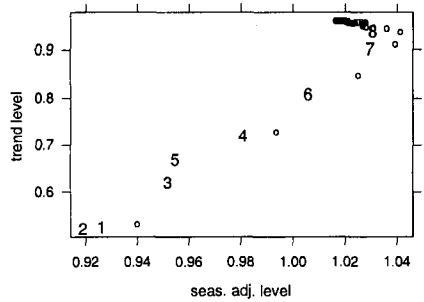
Figure 3. Variances for 1-b-1(c) rotation patterns for proportion employed (MLFS).

Figure 4(a) plots the variance of the one-month difference in trend against that of the one-month difference in seasonally adjusted estimates. The *in-for-c* rotation patterns perform well for both provided c is very large, say $c \geq 24$. The 1-*b*-1(c) rotation patterns perform well for trend but not well for the seasonally adjusted estimates. A compromise is the 2-2-2(8) rotation pattern.

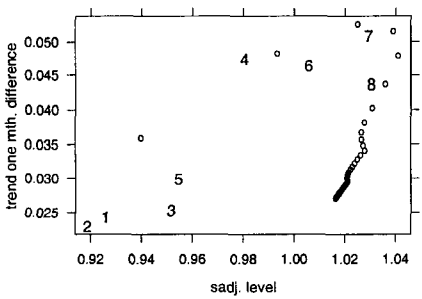
Figure 4(c) shows that the 1-2-1(c) pattern gives lower variance for both



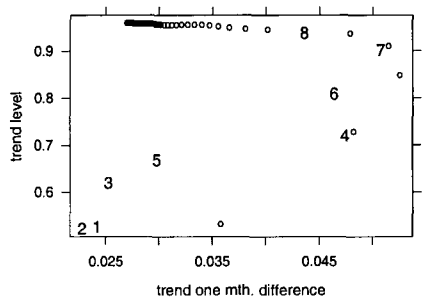
(a) One month diff. in trend vs. one-month diff. in seas. adj.



(b) Trend level vs. seas. adj. level



(c) One month diff. in trend vs. seas. adj. level



(d) Trend level vs. one month diff. in trend

Figure 4. Trade-off between variances of trend and seasonally adjusted estimates assuming different rotation patterns for proportion employed (MLFS).

the one-month difference in trend and seasonally adjusted level. Rotation patterns with high monthly overlap have high variances for the one-month difference in trend estimates, unless the overlap is very high, $c \geq 24$. The results for trend level and one-month difference in trend in Figure 4(d), show that the 1-2-1(c) patterns again out-perform the other rotation patterns.

Similar comments apply to the trend level vs. three-month difference in seasonally adjusted, seasonally adjusted level vs. three-month difference in seasonally adjusted, and three-month difference in trend vs. three-month difference in seasonally adjusted. In all cases the 1-2-1(c) out-performs the other chosen rotation patterns. Similar results are obtained for different choices of correlation model for the sampling errors.

5 Conclusion

The choice of rotation pattern depended on the estimates that are considered important. It was found that the 1-2-1(c) rotation patterns reduced the sampling variance of trend and seasonally adjusted level estimates, sampling variance of the movements in trend estimates, and the movements in the seasonally adjusted estimates except for the one-month movement, for which rotation patterns with high monthly sample overlap perform well.

Most rotation patterns are designed based on criteria involving the change between consecutive periods in the original and seasonally adjusted series. There is a trade-off when designing for the movement estimates for the seasonally adjusted and trend estimates. There is a case to use 1-2-1(c) rotation patterns, unless the one-month change in the seasonally adjusted estimates is the key statistic to be analysed.

Acknowledgments

This research was supported by the Australian Research Council and the Australian Bureau of Statistics. The views expressed do not necessarily reflect those of either organisation.

REFERENCES

Australian Bureau of Statistics (1993). *A Guide to Interpreting Time Series, Monitoring "Trend" - an Overview.* Australian Bureau of Statistics, cat. 1348.0, Canberra, Australia.

Australian Bureau of Statistics (1998). *The Labour Force - Australia.* Australian Bureau of Statistics, cat. 6203.0, Canberra, Australia.

Binder, D.A. and Hidiroglou, M.A. (1988). Sampling in time. In *Handbook of Statistics, Vol. 6*, Krishnaiah, P.R. and Rao, C.R. (editors) . Amsterdam: North Holland, 187-211.

Dagum, E.B. (1988). *The X11ARIMA/88 Seasonal Adjustment Methods - Foundations and User's Manual.* Ottawa: Statistics Canada.

Dagum, E.B., Chhab, N. and Chiu, E. (1996). Derivation and properties of the X11ARIMA and census X11 linear filters. *Journal of Official Statistics* 12, 329-347.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen, B. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program. *Journal of Business and Economic Statistics* 16, 127-177.

Fuller, W.A., Adam, A. and Yansaneh, I.S. (1992). Estimators for longitudinal surveys with application to the U.S. current population survey.

Proceeding of Statistics Canada Symposium 1992, Design and Analysis of Longitudinal Surveys.

Kish, L. (1998). Space/time variations and rolling samples. *Journal of Official Statistics* 14, 31-46.

McLaren, C.H. and Steel, D.G. (2000). The impact of different rotation patterns on the sampling variance of seasonally adjusted and trend estimates. *Survey Methodology* 26, 163-172.

Rao, J.N.K. and Graham, J.E. (1964). Rotation designs for sampling on repeated occasions. *Journal of the American Statistical Association* 69, 492-509.

Shiskin, J., Young, A.H. and Musgrave, J.C. (1967). The X-11 variant of the census method II seasonal adjustment program. *Technical Paper 15, Bureau of the Census, U.S. Department of Commerce, Washington D.C.*

Singh, A.C., Kennedy, B. and Wu, S. (2001). Regression composite estimation for the Canadian labour force survey with a rotating panel design. *Survey Methodology* 27, 33-44.

SIMULATING SEAWATER INTRUSION IN AQUIFERS USING THE MODIFIED FOKKER-PLANCK EQUATION AND BOUSSINESQ EQUATION SUBJECT TO PHASE-MODULATED TIDAL WAVES

[1]NINGHU SU, [1,2]FAWANG LIU AND [1]VO ANH

[1]*School of Mathematical Sciences*
Queensland University of Technology
Gardens Point Campus
Brisbane Qld. 4001, Australia
[2]*Department of Mathematics*
Xiamen University
Xiamen 361005, China

This paper is concerned with the interactions between periodic tidal waves and coastal aquifers. The analysis is based on a set of equations governing the hydrodynamic interactions near coastal regions consisting of a modified Fokker-Planck equation for modelling the transport of salts, a Boussinesq equation for groundwater flow and a wave equation for tidal waves. The modified Fokker-Planck equation is featured with a time- and scale-dependent dispersivity, and solutions of the Boussinesq equation are obtained subject to phase-modulated tidal waves as the boundary condition which is a solution of the wave equation. The phase-modulated periodic solution of the wave equation is shown to satisfactorily simulate the uneven twin peaks of semi-diurnal tides. Fourier series solutions of the Boussinesq equation subject to periodic boundary conditions reveal that the tidal waves damp towards land, and the half amplitude of the tide above the mean sea level is greater than that below it. Seawater-freshwater interface is also analysed using the modified Fokker-Planck equation.

Keywords: modified Fokker-Planck equation, Boussinesq equation, diffusion equation.

1 Introduction

In recent years, more attention has been paid to the interactions between coastal aquifers and tidal waves in order to reveal detailed hydrodynamics, and their impact on beaches and coastal aquifers (see Bear (1972) and Li and Barry (2000) for reviews). In these works, in addition to the use of wave equation, diffusion equation or Laplace equation for modelling groundwater waves, some researchers also presented new mathematical models (Li and Barry, 2000; Nielsen et al., 1997; Parlange and Bruttsert, 1987) to analyse tide-aquifer interactions across a more realistic geometry on the beaches. The majority of

the works is concerned only with fixed boundary conditions except for Li et al. (2001a) who recently used a moving boundary to simulate beach watertable fluctuations due to spring-neap tides. These presentations concentrate on the influence of tidal waves on groundwater on a horizontal base ignoring aquifers overlaying sloping bases which impose a different influence on the tide-aquifer interactions. Recently Su et al. (2001) examined tide-induced groundwater waves in coastal aquifers overlaying a sloping base subject to a periodic boundary condition.

The interactions between periodic tidal waves and coastal aquifers are analysed in this paper. The analysis is based on a set of equations governing the hydrodynamic interactions near coastal regions. A modified Fokker-Planck equation (MFP) is presented for modelling the transport of salts, Boussinesq equation (BE) for groundwater flow and wave equation (WE) for tidal waves. The MFP is featured with a time- and scale-dependent dispersivity, and solutions of BE are subject to phase-modulated tidal waves as the boundary condition (BC) which is a solution of WE. The development of the MFP is achieved by combining two components of the dispersivity, one is a fractional power function of time, and the other a power function of fractal scale. It is shown that the phase-modulated periodic solution of WE satisfactorily simulates the uneven twin peaks of semi-diurnal tides as observed in a study area on the east coast of Queensland, Australia. Furthermore numerical analyses show that the Fourier series solutions of BE subject to a periodic BC, which is a phase-modulated periodic solution of WE, reveal two important features of the tidal waves. First, the tidal waves damp towards land, and second, the half amplitude of the tide above the mean sea level is greater than that below it. These two features have been successfully modelled using the BE. Seawater-freshwater interface is analysed using the MFP, and the simulated results are realistic compared with field data.

2 Problem formulation

The mathematical formulation and analysis start with one-dimension groundwater flow on a sloping base. The configuration of an aquifer in the coast coupled with tidal waves is illustrated in Figure 1.

The basic equation governing one-dimensional groundwater flow on a sloping impervious base is the Boussinesq equation that is written as (Werner, 1957; Mariño and Luthin, 1982)

$$\frac{\partial h}{\partial t} = \frac{K}{S}\left[h\frac{\partial^2 h}{\partial x^2} + \left(\frac{\partial h}{\partial x}\right)^2 - \Theta\frac{\partial h}{\partial x}\right], \tag{1}$$

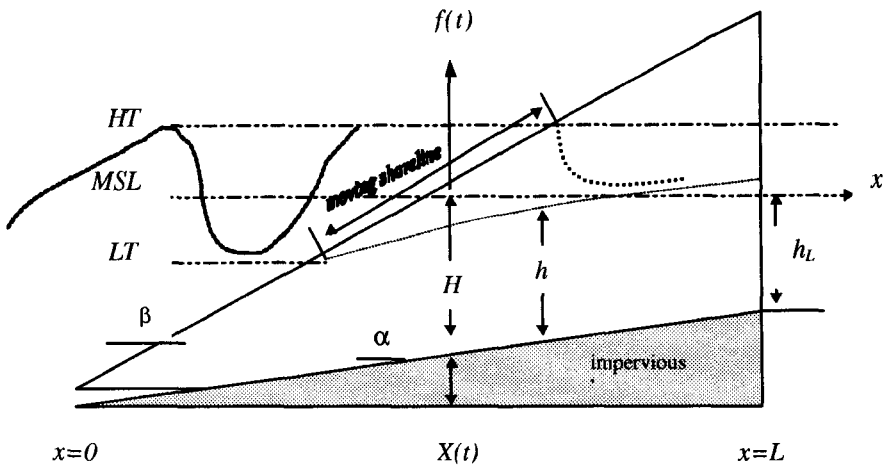


Figure 1. A schematic diagram of an unconfined coastal aquifer-tide system.

where

h is the height of watertable;
S is the specific yield;
x is the space coordinate;
K is hydraulic conductivity;
Θ is the slope of the impervious base, $\Theta = tg\alpha$
a is the angle between the horizontal datum and the impervious base of the aquifer;
$R_{(t)}$ is the time-dependent recharge rate; and
t is time.

In Figure 1, we choose the intersection between the mean sea level (MSL) or mid-tidal sea level (Li et al., 2001a) and the beach face as the origin of the x coordinate, and consider the problem defined in Figure 1 and (1) subject to the following conditions in a semi-indefinite domain:

$$h(x,0) = h_0, \qquad x > 0, \quad t = 0, \tag{2}$$

where

$$h\,(X(t),t) = f(t), \qquad x = X(t), \quad t > 0, \tag{3}$$

where $X(t)$ the x-coordinate of the moving boundary given by

$$X(t) = \cot\,(\beta) f(t), \tag{4}$$

where

β the angle between the beach face and the horizontal base;
h_0 the initial watertable;
$f(t)$ the periodic vertical perturbation relative to MSL due to tidal waves.
H the equilibrium depth of water on the beach taken as the mean sea level (MSL);
h_L the height of watertable at inland boundary L.

The groundwater discharge is given by

$$Q = Kh\left(\Theta - \frac{\partial h}{\partial x}\right). \tag{5}$$

As the term Kh denotes transmissivity, T, and if h deviates only by a small amount from the weighted depth, i.e., $\widehat{H} \approx h$, then transmissivity, T, could be regarded as independent of h (Werner, 1957). In such a case, (1) becomes the linear advection-diffusion equation,

$$\frac{\partial h}{\partial t} = \frac{\partial}{\partial x}\left(D\frac{\partial h}{\partial x}\right) - v\frac{\partial h}{\partial x}, \tag{6}$$

where

$$D = \frac{T}{S} \tag{7}$$

and

$$\nu = \frac{K\Theta}{S}. \tag{8}$$

We are only interested in a medium with an average hydraulic conductivity (Bear, 1972)

$$\widehat{K} = \frac{1}{h}\int_0^h K(x, y, z')dz' \tag{9}$$

and thereafter in this paper, K will be replaced by $\widehat{K}$.

3 The general solutions subject to periodic boundary conditions

As shown by Su et al. (2001), the solution of (6) is given by

$$h_n = \exp\left[\left(\frac{\nu}{2D} - \mu_n\right)x\right]\exp\left[i\left(\omega_n t - \rho_n x\right)\right], n \geq 0, \tag{10}$$

where

$$\omega_n = 2n\pi/T, \tag{11}$$

$$\mu_n = \sqrt{\frac{\sqrt{\phi^2+\psi^2}+\psi}{2}}, \tag{12}$$

$$\rho_n = \sqrt{\frac{\sqrt{\phi^2+\psi^2}-\psi}{2}}, \tag{13}$$

$$\phi = \frac{\omega}{D}, \tag{14}$$

$$\psi = \frac{v^2}{4D^2}, \tag{15}$$

with n being a counter $[0, \infty]$ of different waves

Now, consider a T -periodic boundary condition, $f(t)$, and define the pair of Fourier transformation

$$f(t) = \sum_{-\infty}^{\infty} a_n \exp(i\omega_n t), \tag{16}$$

$$a_n = \frac{1}{T}\int_0^T f(t)\exp(-i\omega_n t)\,dt. \tag{17}$$

A solution of (6) is now given in a complex conjugate form as

$$h = \sum_{n=0}^{\infty} a_n h_n(x,t) + \sum_{n=1}^{\infty} a_{-n}\widehat{h}_n(x,t) \tag{18}$$

where an overbar over a complex number means complex conjugate.

The phase-modulated tidal wave solutions

A more realistic model for semi-diurnal tidal waves is the compound cosine function which is essentially a solution of the one-dimensional wave equation for phase-modulated waves (Elmore and Heald, 1969). The solution is

$$f(x,t) = A_0 \cos[k_0 x - \omega_0 t + \Phi(x-ct)], \tag{19}$$

where k_0 and $\omega_0 = k_0 c$ pertain to the unmodulated waves with

c the wave velocity;
ω_0 the unmodulated frequency;
A_0 is amplitude, and
A_1 is a constant accounting for the displacement of the amplitude.

The function $\Phi(x - ct)$ in (19) is for phase modulation.

A solution of the one-dimensional wave equation is given by (19). Here we examine it as a function of space and time (Elmore and Heald, 1969), with

$$\Phi(x - ct) = m \sin(\omega_0 t), \tag{20}$$

where m is the modulation index. Using(20), (19) is written as

$$f(x,t) = A_0 \cos[k_0 x - \omega_0 t + m \sin(\omega_m t)]. \tag{21}$$

Phase shifts and a constant to account for amplitude shift can be introduced in (21),

$$f(x,t) = A_1 + A_0 \cos\{k_0 x - \omega_0 (t - t_1) + m \sin[\omega_m (t - t_2)]\} \tag{22}$$

where

A_1 is a constant to account for vertical amplitude shift, and t_1 and t_2 are phase shifts.

When (22) is fitted to the observed tidal data collected at Bundaberg Port, Queensland, Australia with the parameters $A_1 = 0.2$, $A_0 = 1.7$, $\omega_0 = 0.2529$, $\omega_m = 0.2529$, $m = 2.0$, $t_1 = 12.42$, $t_2 = 12.0$, the fit shown in Figure 2 is satisfactory.

4 Freshwater-seawater interface subject to phase-modulated tides

In the following section, we consider the transport of salts in the aquifer due to seawater intrusion subject to phase-modulated tidal waves. The basic equation governing the transport of salts is the Fokker-Planck equation which is given as

$$\phi \frac{\partial C}{\partial t} = \nabla \cdot (D \nabla C) - \nabla \cdot (Cv), \tag{23}$$

where

ϕ is porosity;
C is salt concentration;

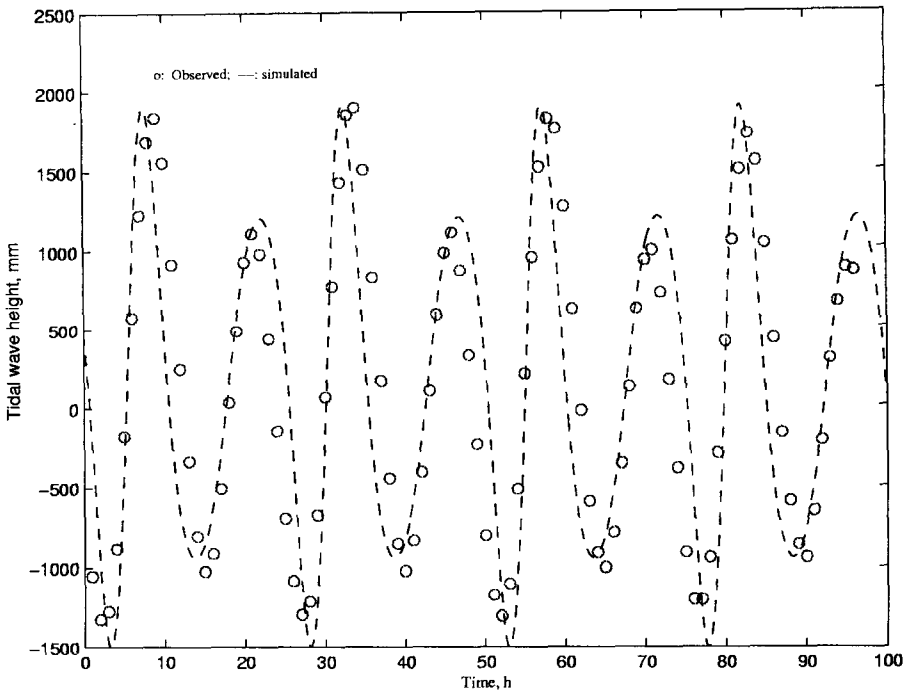


Figure 2. The observed and simulated semi-diurnal tidal wavesat Bundaberg Port, Queensland, Australia, starting from 00:00, AM, 01 Jan. 1999

$\boldsymbol{D}$ is the dispersion tensor which, in two dimensions, is given as

$$D_{xx} = \left(\alpha_L \frac{\nu_x^2}{|v|} + \alpha_T \frac{\nu_z^2}{|v|} + D_d T_{xx}\right) t^{-\lambda},$$
$$D_{zz} = \left(\alpha_T \frac{\nu_x^2}{|v|} + \alpha_L \frac{\nu_z^2}{|v|} + D_d T_{zz}\right) t^{-\lambda},$$
$$D_{xz} = D_{zx} = \left((\alpha_L - \alpha_T)\frac{\nu_x \nu_z}{|v|}\right) t^{-\lambda}, \tag{24}$$

where T_{xx} and T_{zz} are the principal components of the tortuosity tensor; and α_L and α_T are the longitudinal and transverse dispersivities respectively, and are defined as

$$\alpha_L = \frac{D_{0x}}{2}\sigma^2 x^m, \quad \alpha_T = \frac{D_{0z}}{2}\sigma^2 z^m, \tag{25}$$

where D_{0x} and D_{0z} are constants dependent on media properties; σ^2 is the variance of the lower fractal cutoff limit subject to a fractal dimension d $(1.0 < d < 2.0)$, with $m = 2d - 1$. (25) implies that in a heterogeneous porous medium there are two sources of variation in the flow travel path. The variation at a microscale is due to the heterogeneity of the media which is characterized by the variance of the fractal cutoff limit, while the variation at a geological scale is characterized by the fractal dimension. Incorporating these two sources of variation in the traditional Fokker-Planck equation is physically meaningful.

It is interesting to note that for a mean value of m, i.e., $m = 3/2$, the fractal dispersivity has been shown by Wheatcraft and Tylor (1988) to give the best fit of the field data from extensive tracer studies carried out under different conditions (Gelhar et al., 1992). When α_L and α_T are constants i.e., $m = 0$, and $\lambda = 0$, (25) reduces to the same form given in Bear (1999). Throughout this paper, the case where $\alpha_L = 0, \alpha_T = 0$ and $D_d \neq 0$ will be referred to as the constant coefficient dispersion; the case when $\alpha_L \neq 0, \alpha_T \neq 0, D_d = 0$ the variable coefficient dispersion, and the case with $m \neq 0$ the time- and scale- dependent dispersion.

The flow equation may be written as (Frind, 1982; Huyakorn et al., 1987)

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot [\mathbf{K}(\nabla h + \gamma C \nabla z)], \tag{26}$$

where

h is referred to as the freshwater reference hydraulic head;
$\mathbf{K}$ is the hydraulic conductivity tensor;
γ is the density coupling coefficient;
C is the solute concentration;
S_s is the specific storage;
t is time; and
z is elevation.

The reference head and the density coupling coefficient in (26) are defined as

$$h = \frac{p}{\rho_0 g} + z, \tag{27}$$

and

$$\gamma = \frac{\epsilon}{C_{\max}}, \tag{28}$$

where

p is the fluid pressure;
g is the gravitational acceleration;
$C_{\max}$ is the concentration that corresponds to the maximum density;
ρ_0 is the reference (freshwater) density;
ϵ is the density difference ratio defined as

$$\epsilon = \frac{\rho_{\max}}{\rho_0} - 1. \tag{29}$$

The relationship between fluid density, ρ, and concentration, C, under isothermal conditions can be expressed in the form

$$\rho = \rho_0(1 + \gamma C) = \rho_0(1 + \rho_r), \tag{30}$$

where ρ_r is the relative density.

With the coupled flow and transport equations, numerical solutions have been developed to simulate seawater intrusion (Liu et al., 2002). The simplest case of seawater intrusion is the so-called Henry problem.

For the Henry problem, two sets of parameters are used to simulate both hydraulic pressure and salt distributions across the model aquifer. The results are shown in Figure 3.

It is seen from Figure 3 that in the upper portion of the aquifer the simulated salts using the fractal model tend to move inland and mix better with freshwater. This result is different from the classic Henry isochlors (Reilly and Goodman, 1985). The explanations for this difference are that due to enhanced convection and mixing in the upper portion of the aquifer the salt concentration in the freshwater water increases. Further experiments are required to verify this phenomena with field data.

5 Conclusion

In the preceding analysis, some important characteristics of tides on beaches have been analysed in detail.

1. As semi-diurnal tides exhibit uneven twin peaks, it has been shown that the phase-modulated periodic solution of the wave equation satisfactorily simulate the uneven twin peaks of semi-diurnal tides observed in the field at Bundaberg on the east coast of Queensland, Australia.

2. Numerical analyses has shown that the Fourier series solutions of the Boussinesq equation subject to periodic BCs capture important features

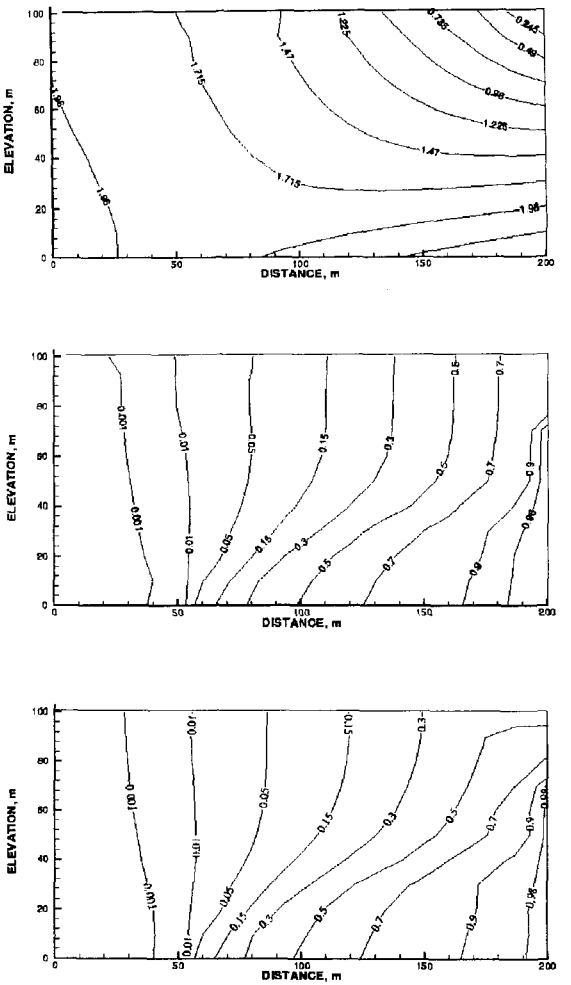


Figure 3. Distribution of hydraulic pressure (top, with seawater on the right hand side), and isochlors (salt contents) with different fractal dispersion parameters: $\lambda = 1/2; d = 5/4; D_{0x} = D_{0y}$ (middle), and $\lambda = 1/2; d = 5/4; D_{0x} = D_{0y}/3$ (bottom).

of tidal waves. First, tidal waves damp towards land, and second, the half amplitude of the tide above the mean sea level is greater than that below it. These methods have been confirmed by field data collected at Bundaberg Port, Queensland, Australia.

3. With a functional dispersivity, freshwater-seawater interface has been simulated which shows more realistic results. Further work and data collection are in progress to verify the findings.

Acknowledgments

This research was supported by the ARC SPIRT grant C10024101. Thanks are due to G.J. Broadbent, Maritime Division, Queensland Department of Transport who kindly provided data used in the figures.

REFERENCES

Bear, J. (1972). *Dynamics of Fluids in Porous Media.* New York: Dover Publications.

Bear, J., Cheng, A.H., Sorek, S., Ouazar, D. and Herrera, I. (1999). *Seawater Intrusion in Coastal Aquifers- Concepts, Methods and Practices.* Dordrecht, The Netherlands: Kluwer Academy.

Elmore, W. and Heald, M.A. (1969). *Physics of Waves.* New York: Dover.

Frind, E.O. (1982). Simulation of long-term trasient density-dependent transport in groudwater. *Adv. Water Resour* 5, 73-88.

Gelhar, L.W., Welty, C. and Rehfeldt, K.R. (1992). A critical review of data on field-scale dispersion in aquifers. *Water Resour. Res.* 28, 1955-1974.

Huyakorn, P.S., Andersen, P.F., Mercer, J.W. and White Jr., H.O. (1987). Saltwater intrusion in aquifers: develpoment and testing of a three-dimensional finite element model. *Water Resour. Res.* 23, 293-312.

Li, L. and Barry, D.A. (2000). Wave-induced beach groundwater flow. *Adv. Water Resour.* 23, 325-337.

Li, L., Barry, D.A., Stagnitti, F. and Parlange, J.Y (2000a). Groundwater waves in a costal aquifer: a new governing equation including vertical effects and capillarity. *Water Resour. Res.* 36, 411-420.

Li, L., Barry, D.A., Stagnitti, F., Parlange, J.Y. and Jeng, D.S. (2000b). Beach water table fluctuations due to spring-neap tides: moving boundary effects. *Adv. Water Resour.* 23, 817-824.

Liu, F., Turner, I., Anh, V. and Su, N. (2002). Finite volume methods for transient simulation of time-, scale- and density-dependent transport in heterogeneous aquifer systems (preprint).

Mariño, M.A. and Luthin, J.N. (1982). *Seepage and Groundwater.* Amsterdam: Elsevier.

Nielsen, P., Fenton, J.D., Aseervathan, R.A. and Perrochet, P. (1997). Watertable waves in aquifers of intermediate depths. *Adv. Water Resour.* 20, 37-43.

Parlange, J.Y. and Brutsaert, W. (1987). A capillary correction for free surface flow of groundwater. *Water Resour. Res.* 23, 805-808.

Reilly, T.E. and Goodman, A.S. (1985). Quantitative analysis of saltwater-freshwater relationships in groundwater systems - a historical review. *J. Hydrol.* 80, 125-160.

Su, N. and Liu, F. and Anh, V. (2001). Tides as phase-modulated waves inducing periodic groundwater flow in coastal aquifers overlaying a sloping impervious base. *Journal of Hydrology* (submitted).

Werner, P.W. (1957). Some problems in non-artesian ground-water flow. *Trans. Amer. Geophys. Union* 38, 511-58.

Wheatecraft, S.W. and Tyler, S.W. (1988). An explanation of scale-dependent dispersivity in heterogeneous aquifers using concepts of fractal geometry. *Water Resour Res.* 24, 566-578.

MODELLING HEALTH INEQUALITIES USING DYNAMIC MICROSIMULATION: STATISTICAL ISSUES AND FUTURE SCOPE

AGNES. E. WALKER
National Centre for Epidemiology and Population Health
Australian National University; and
National Centre for Social and Economic Modelling
University of Canberra, Australia
E-mail: agnesw@natsem.canberra.edu.au

The paper describes stage 1 of the development of a Health Module, within a dynamic microsimulation model simulating individuals over the lifecourse. The enhanced model accounts for Australians' health and mortality by socio-economic status. In developing the Module we found that household-based health surveys were unsatisfactory, as they excluded the very sick who resided in institutions. We thus chose the disability surveys which cover institutions as well as households. Individuals' disability status was used as a proxy for their health status, and disability and life expectancy were linked mathematically. Amongst the available indicators of socio-economic status, the geographically based index of socio-economic disadvantage was found to underestimate health inequalities, while an individual-level index based on family cash incomes was found to overestimate these. To obtain more accurate estimates, we propose indexes accounting for both cash income and assets. The paper also describes a range of analyses that could be attempted once that Module is completed. Such studies could for example compare the demographic, labour force, financial and distributional impacts of policies that aimed to lower health inequalities - and do that more accurately and at a much greater level of detail than was possible previously.

Keywords: health inequalities, socio-economic studies, probabilistic modelling, microsimulation.

1 Introduction

In recent years health inequalities became a much-researched subject, with a number of developed countries having already announced policies that aim to reduce inequalities in health - eg the UK and Canada (Acheson, 1998; King's Fund, 1999; McCain and Mustard, 1999). Much of the research published in Australia and overseas concerns differences in mortality patterns across socio-economic groupings. The virtually universal finding is that people with low socio-economic status (SES) - usually the 'poor' - die younger than people with high SES - usually the 'rich' (World Health Organisation, 2000; Glover

et al.,1999). More recently the focus of research has shifted to the reasons for premature death of the socially disadvantaged, and to the study of individuals' health and lifestyles as they affect mortality (Marmot, 1998). Also, there has been increasing interest internationally in the study of health in a lifecourse context (Barney et al., 2001). This project aims to simulate the life paths of individuals over a 20 to 30 year period using a dynamic microsimulation model (Antcliff *et al.*, 1996). In its first stage it involves the modelling of health as a function of socio-economic status - the subject of this paper - with mortality and disability being the indicators of health. In later stages we intend to add additional health indicators (such as the severity of disability and the main disease causing that disability), and to model health related decisions regarding issues such as remaining in (or exiting) the workforce, becoming a carer or, in old age, leaving home to enter an institution (eg hostel, nursing home). The aim of the full project is to make a unique contribution to the broadening of the way in which health inequalities are analysed.

2 Linkages between health and socio-economic status

In the Base population of the dynamic microsimulation model, DYNAMOD (Appendix A1), disability status is being allocated to individuals, by socio-economic status, in line with historically observed patterns. Disability is defined by the Australian Bureau of Statistics (ABS) as a limitation, restriction or impairment, which has lasted, or is likely to last, for at least six months and which restricts every day activities. Demographic, social and economic events occurring throughout the lives of people in this Base population - ie births, deaths, family structures, migration, education, labour force participation, earned income, government transfers, taxes and wealth - will then be simulated over a 20 to 30 year period. Figure 1 illustrates the way the link between health (ie mortality, disability) and SES is handled in the enhanced model.

Briefly, when a child is born, he/she is being allocated a date of death based on observed probabilities and on his/her parents' socio-economic status. A disabled new born is allocated an earlier date than an able-bodied new born. When considering probabilities of death, allowance is made for increases in future life expectancies, based on estimates by the ABS. It is only if a change occurs in a person's disability or socio-economic status that the initially allocated date of death is re-estimated. Appendices A2 and A3 describe the equations, for the model's input data, that mathematically specify the linkages between disability and mortality.

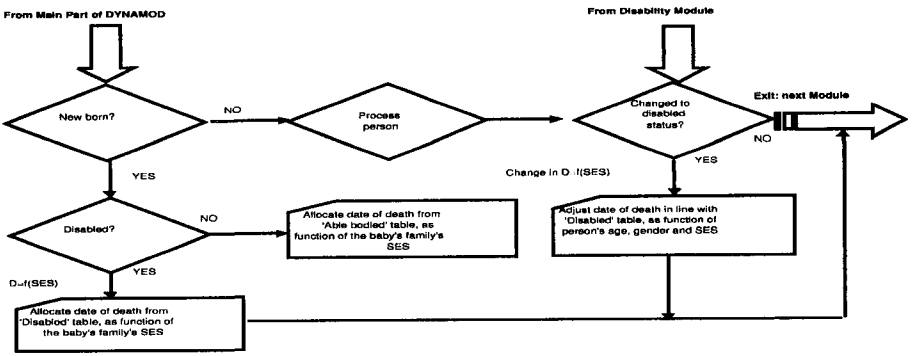


Figure 1. The links being modelled between mortality and disability, by socio-economic status

3 Choice of data sources

Amongst the available data sources we examined two large nationwide unit record surveys:. the ABS's latest (1995) National Health Survey (NHS) - a household survey (ABS, 1996) - and its 1993 and 1998 Disability surveys (ABS, 1999) - which include both households and institutions. Our earlier work suggested that data from household surveys were unlikely to be satisfactory when studying health status over the lifecourse, because such surveys excluded persons in institutions - such as hospitals, prisons, nursing homes (Walker, 2000). For example, we found that respondents to the 1995 NHS aged 70 years or more appeared to have better health on average than did younger age groups (Walker, 2000). This was thought to be because seriously ill or disabled Australians aged 70 or over who had moved into a hospital, hostel or nursing home were excluded from the survey. A similar pattern emerged from two other household surveys: the 1993-94 and 1998-99 Household Expenditure Surveys (HES) conducted by the ABS. As seen in Figure 2, in both years persons aged 75 or over appeared to have spent less on prescribed pharmaceuticals than did younger age groups. Another interesting finding was that in most age groups a higher proportion of persons spent on prescribed drugs in 1998-99 than in 1993-94 (Figure 2). More aggressive screening and treatment options available in recent years may have contributed to this pattern.

In addition, this pattern is consistent with the generally upward trend in the number of prescriptions subsidised under Australia's Pharmaceutical

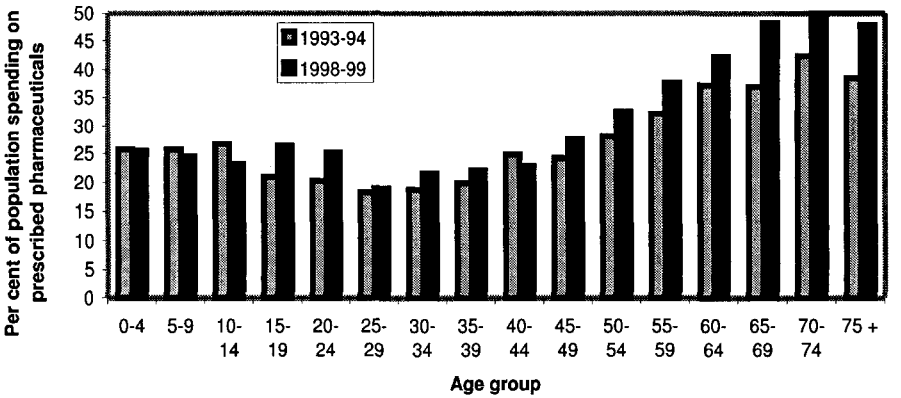


Figure 2. Spenders on prescribed drugs, per cent of the population by age group, 1993-94 and 1998-99

Benefits Scheme (a growth rate of close to 3 per cent a year). While it could be argued that those who survived to age 75 may have had a stronger constitution than did the 'younger old', this was not borne out by the statistics reported in the 1998 Disability Survey - a survey which included persons in institutions. That survey showed the proportion of Australians with a disability and/or long term illness as increasing steadily well beyond age 70 (Figure 3).

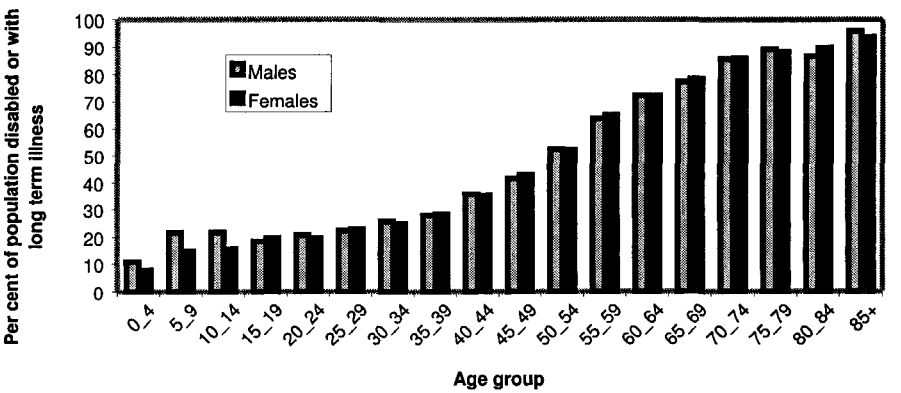


Figure 3. Proportion of population disabled and/or with long-term-illness, 1998

Comparisons across Figures 2 and 3 suggest that the HESs significantly

underestimate pharmaceutical spending in older age groups - probably from age 70 onwards. This pattern is consistent with the cut-off age of 70 that needed to be adopted in an earlier study analysing the 1995 NHS (Walker and Abello, 2000). The problem of excluding people aged over 70 or 75 years is that then the excluded group accounts for the majority of persons who are a year or two away from death (Walker, 1998) - a group of particular importance to studies of life time health expenditures and of disability-illness-mortality linkages. In view of the above, the Disability survey was chosen for this project as it clearly provides better lifecourse estimates of health status than the household-based surveys. The Disability survey has linkages to illnesses, through the 'main disabling condition' variable, as well as information on the severity of disability. Use of these variables is planned for later stages of the project and, if necessary, additional disease-related information may be imputed from National Health Surveys, but only to age 70 or 75.

4 Indicators of health status

Mortality

When studying health at the population level it is common to use various mortality rates and related life expectancies as proxies for health status - eg infant mortality, life expectancy at birth (World Health Organisation, 2000; Walker, 2000). However, health as indicated by mortality rates says nothing about how health problems developed over an individual's life; how they impacted on a person's family or employment prospects; what factors contributed to a particular mortality outcome; and how that outcome could have been altered through adoption of various interventions by individuals or government. In effect, we know very little from standard statistical collections about the characteristics of people who just died. In this project, by linking mortality to disability, we expect to fill in some aspects of this knowledge gap. For this project we used unpublished mortality statistics extracted by the Australian Institute of Health and Welfare (AIHW) from its 1995-97 Mortality Database. Apart from the variables often used in earlier mortality studies - age, gender and an index of socio-economic status based on area of residence - we also obtained data split into two groups based on cause of death. These were 'External' and 'Non-external' causes, which were separately identified so as to be able to distinguish between deaths due to accidents (eg drowning, car crashes) and deaths due to illnesses (eg heart attack). This distinction is important because people dying from external causes are generally much younger than people dying from diseases or old age - hence the former group,

although much smaller, will be important when estimating the effects of policy or other changes in terms of 'years of life lost'.

Disability

For this project we chose the ABS's Disability survey as a key data source (Section 3). The choices were to stay with the 1993-94 survey (used in the original model), or to up-date the model's Base data using the1998 survey. Because in our project we intended to model the SES-health link, and because the income variable was a potential candidate for indicating SES, we examined the income related statistics in these two surveys. We chose the 1998 Disability survey because we found that the non-response rate to the 'income' question in that survey was much lower than in the 1993-94 survey. Although the 1998 survey - covering some 43,000 persons - is cross sectional in nature, we were able to construct lifecourse patterns by assuming, for the model's default settings, that the 1998 patterns for groups with a particular set of demographic and socio-economic characteristics remained unchanged throughout the simulation period.

5 Indicators of socio-economic status

Based on income or area of residence

Two common indicators of socio-economic status used in Australian studies are the ABS's socio-economic indexes for areas (SEIFA) - (Dunn *et al.*, 2002) - and equivalent (or per person) family income (Saunders, 1996). In the literature family income was found to be a good single 'proxy' for socio-economic disadvantage (Walker and ABello, 2000) . With mortality statistics, only the SEIFA-based indicator is available. However, using the 1998 Disability survey we had the choice of either the SEIFA, or an income-based indicator of SES. The question was which of these two was preferable for our purposes. Studying both mortality and disability as a function the SEIFA index has the advantage of consistency across different aspects of the project. This however needed to be weighed against the advantages of the income-based indicator, in that the latter can be applied to individuals while the SEIFA applies to all people residing in a particular geographical area. Thus households in a particular area - in our case a Collector District covering on average around 200 households - will all have the same SEIFA quintile attached to them, and thus that data will not reflect the variability in socio-economic status that in reality does exist within each CD. In addition, since a 'per person' indicator cannot be obtained from the SEIFA, with that index it is not possible to distinguish between large and small families. For example, a two person couple family

with annual income of $50,000 a year residing in a particular area will have the same SEIFA quintile as a five person family, such as a couple with three dependent children, living in the same area. By comparison, with the income index it is common practice to use the 'equivalent family income' measure - which can be thought of as being a 'per person' indicator of family income. A priori, one would expect equivalent income to be a more appropriate and precise indicator of SES than the geographically based SEIFA index. This is because family income reflects the purchasing power of the particular family being studied, and not the collective wealth of the geographic area in which that family happens to reside.

However, a key disadvantage of the income-based indicator is that it takes no account of assets. In a lifecourse context many people start out by being 'cash poor and asset poor', then progress to a 'cash rich - asset poor' status, ending their lives as 'cash poor and asset rich'. This means that the income measure will tend to class older Australians into the lower SES groups which, for a significant proportion, may not be in line with what their life styles suggest (eg in terms large houses, expensive cars, overseas holidays). By comparison the SEIFA index has the advantage of taking account - albeit indirectly - of assets.

Using the 1998 Disability survey, we were able to study the way the allocation of the same population to SES quintiles differed across the two measures (SEIFA quintiles and income quintiles). For this exercise we used the ABS's index of 'Relative Socio-economic Disadvantage' as the SEIFA Indicator. For income we computed 'equivalent family income' (Walker and Abello, 2000). Because the SEIFA represents an average for a geographic area - and thus is unable to take account of variability - one would expect its use to result in a 'flattening' of the inequality gradient, compared with the individual family-based income indicator. This is supported by our findings when plotting the proportion of disabled by the two SES measures, with Figure 4 suggesting considerably greater inequalities in the proportion of disabled across income quintiles than across SEIFA quintiles. With the SEIFA measure 26 per cent of persons in the most disadvantaged SES group (quintile 1) were disabled in 1998, compared with 14 per cent in the least disadvantaged group (quintile 5). When using the 'equivalent income' measure, a much higher proportion of the quintile 1 persons were disabled (35 per cent), with a much lower proportion in quintile 5 (10 per cent). However, with the income measure we found that virtually no persons in the older 65+ group fell into the more affluent top three quintiles - probably due to the 'cash poor but asset rich' status of many in that age group (Walker, 2001). Overall, neither of the measures was found to be clearly superior. So for reasons of comparability, we chose the

ABS's index of 'Relative Socio-economic Disadvantage' (SEIFA) for both mortality and disability analyses, when preparing the input data to the dynamic microsimulation model.

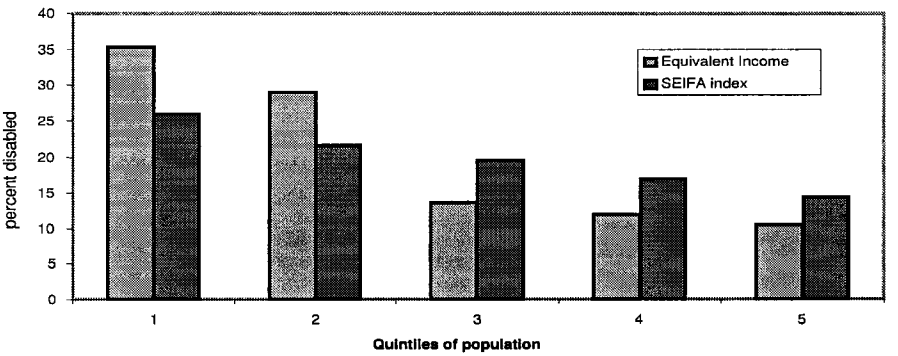


Figure 4. Proportion disabled by equivalent income and SEIFA quintiles, 1998

Income-wealth measure as indicator of SES in projection years

Although the projection years of the model are outside the scope of this paper, the proposed SES measure for that part of the project is briefly mentioned here. As seen in Section 5.1, an index which can be attached to individuals, and which accounts for both cash income and assets, seems preferable to the alternatives available. Because dynamic microsimulation is ideally suited to take account of wealth accumulation, and because assets are already computed within the model, we are able to propose that a combined income-wealth indicator be used during the model's projection years. It is expected that such an index would be a function of per year 'equivalent' family cash income - using model variables such as earnings and social security payments - and an annualised value of 'equivalent' family assets.

6 Modelling the health-SES link

Preparing the Base year data

The input data needed for the model - by gender, SES and single years of age (0 to 104) - comprises (Appendixes A2 and A3): Probability of death - for the able-bodied (who die from external causes), the disabled (who die from non-external causes) and the population in general (ie able-bodied plus

disabled). The probability of death is a function of disability - ie the disabled have lower life expectancies than the able bodied; Mortality improvement rates covering three periods: 1987 to 1994, 1995 to 2004 and 2005 to 2050 - based on ABS predictions. These account for likely future increases in life expectancies; Disability prevalence (computed by as the number disabled divided by the total number of persons in that class); Disablement rate (function of the number becoming disabled, the number of recoveries, the able-bodied population, and the number of deaths in the able bodied population); Recovery rate (with values chosen so that the number of persons becoming disabled remains positive for all ages within the SES quintile being studied). The data sources used were an extract from the AIHW's Mortality database (1995-97), the ABS's SEIFA estimates from the 1996 Census, and the ABS's 1998 Disability survey. For survey-based analyses we used the SAS programming language.

Mortality rates

(a) Rates by SES quintiles

To compute separate mortality rates for the able-bodied and disabled populations by SEIFA quintiles, we used the equations detailed in Appendix A2. It was only possible to obtain these separate mortality rates by using data from two sources: the AIHW extract for mortality by SES, and the 1998 Disability Survey for disability prevalence by SES. This was because - as the equations indicate - the disability prevalence rate mathematically links the able-bodied population to the disabled population. Mortality rates were initially computed by gender, SEIFA quintile and 5-year age groups and then converted - using the GAM spline program of SAS - into single years of age. The 'fit' in all cases was statistically significant (for smoothing: p-value less than 0.0001; for parameter estimates: p-value less than 0.0001). As an example of the results obtained by SES, Figure 5 presents for males mortality rates from external causes (such as accidents or suicides). As expected, the Figure shows that, when compared with most older age groups, males aged 20-34 years had exceptionally high such mortality rates. Indeed, the rates for that young group were similar to those observed for 70-74 year old men! The Figure also shows that quintile 1 people consistently had the highest mortality rates, and quintile 5 the lowest - with quintiles 2, 3 and 4 falling in between. We found that this ranking remained consistent not only when analysing mortality, but also across all the other aspects of the input data required by the model. This is remarkable, given that the various rates were computed using data from several unrelated sources and were linked through complex equations (Appendix A2). Although not reproduced in this paper,

for all cause mortality we found the pattern to age 39 to be similar to that in Figure 5. However, for the over 40 age groups all cause mortality rates were found to be well above those of Figure 5 (Walker, 2001) - illustrating the dominance of illness-related deaths after age 40, and the much greater importance of 'external cause' related deaths for the younger age groups. Hence the decision to model mortality from external and non-external causes separately.

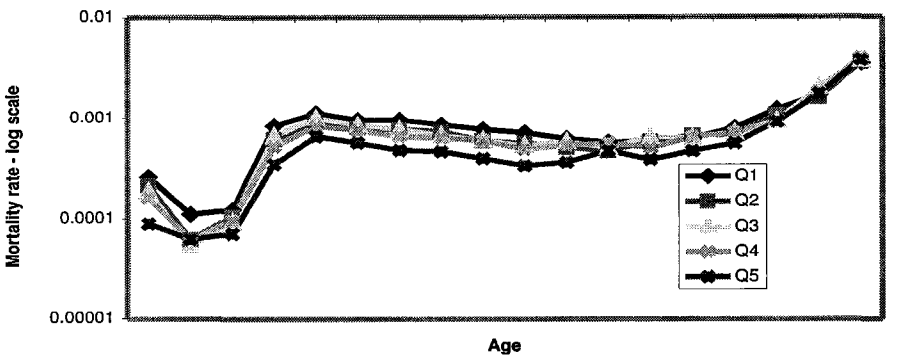


Figure 5. Mortality rates, external causes males age, SEIFA quintiles, 1995-7

(b) Rates over time

By comparing mortality statistics from the model's original 1990-92 based input data with the new 1995-97 data (computed by SES quintile), it was possible to examine how mortality rates changed over time. Figure 6 illustrates this for males, for non-external causes and - for sake of readability - for all SES quintiles.

We chose non-external causes, because for the able-bodied (ie people dying from 'external' causes) mortality rates changed very little over time. Figure 6 also shows that even in this relatively short period - five years between the early and late 1990s - mortality rates from non-external causes declined significantly. This is consistent with the general trend towards increasing life expectancies.

Disability prevalence

Unlike the trend for mortality (Figure 6), we found that disability rates generally increased between the early and late 1990s - as shown in Figure 7 which plots the proportion of the disabled by single years of age for 1993

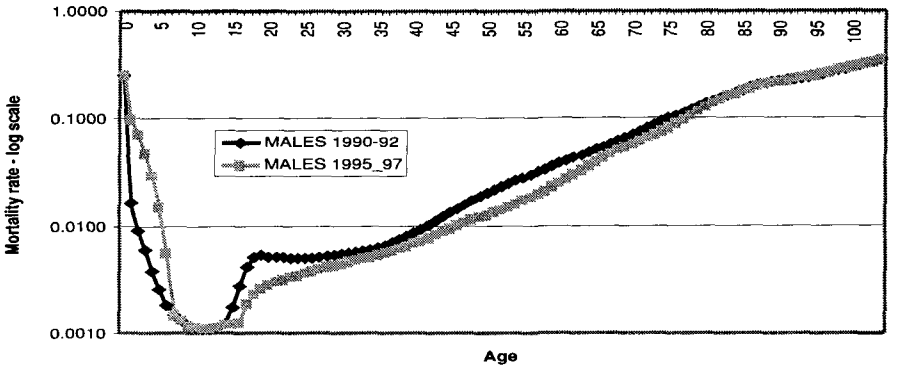


Figure 6. Mortality rates, non-external causes (ie the disabled population), by single years of age and gender, 1990-92 and 1995-97

(the year of original model's input dataset) and 1998 (the year for which we computed prevalence statistics by SEIFA quintile). While this finding may seem counter-intuitive - since many people seem to be of the view that Australians are living longer because they are healthier - it is in line with the findings of earlier studies. Considering the 1983 to 1995 period, one such study found that there were steady increases in the number of Australians reporting serious illnesses, as well as in the number of doctor visits per person, per year (Walker and Abello, 2000). Similar findings were reported in the literature for other developed countries. Amongst likely explanations are the possibilities that modern technology is now able to keep disabled people alive longer and/or that, due to more frequent doctor visits, people are now aware of a greater number of illness-related disabilities than previously (Walker and Abello, 2000). As for mortality rates, the disability data by SEIFA quintile indicated a consistently higher prevalence rate for quintile 1 people (the most disadvantaged) than for the quintile 5 group.

Disability decrement rates

Disability decrement rates were derived using standard multiple decrement techniques (Antcliff *et al.*, 1996). Briefly, the able bodied population is subject to exits due to death and disability and entrants due to recovery, while the disabled population is subject to exits due to death and recovery and entrants due to onset of disability. We assumed that, for all SEIFA quintiles, the recovery patterns specified in the original input dataset remained unchanged. For the new input data we then only needed to compute, by

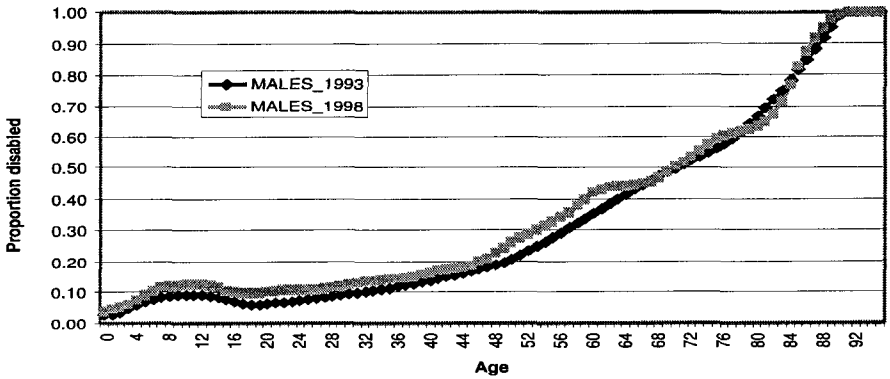


Figure 7. Disability prevalence rates by age and gender, 1993 and 1998

SEIFA quintile, the number - and rate - of people becoming disabled. We did this using equations A.5, A.6 and A.7 in Appendix A3. Our computed disablement rates for all quintiles were similar to the rates initially specified for the model.

7 Concluding remarks

Once the embedding of the input statistics in the simulation part of the model is complete, a limited number of applications will become possible. For example: what would be the impact on Australians' health (indicated initially by mortality/disability) of policies through which the health of the most disadvantaged was lifted to the level of, say, the national average, in terms of lower mortality rates. In later stages of the project the effects on a range of other factors, such as welfare payments, workforce exits due to health - based on family level decisions - and nursing home entries/costs could also be analysed. This initial illustrative simulation could be thought of as being similar to, but considerably more complex than earlier analyses (Mitchell *et al.*, 2000). Examples of relevant findings from (Mitchell *et al.*, 2000) are that: 7597 lives could be saved (7% of all deaths under age 65) if wealth redistribution patterns in the UK were reduced to those of the early 1980s; or 92% of avoidable child deaths could be prevented in areas where the death rates were higher than the national average if child poverty were eradicated; or 2504 lives could be saved through achievement of full employment. Another example, from (Turell and Mathers, 2001) is that: " If it were possible to reduce

death rates .. to a level equivalent to that of the least disadvantaged area, premature all-cause mortality for males in each age group would be lower by 22%, 28% and 26% respectively, and for females, 35%, 70% and 56%." Once the full project is completed it will be possible to study the impact of future health improvements on people remaining in the workforce after age 65. Such an impact could be measured, for example, in terms of changes in government expenditures on aged care and on the age pension. Overall, the ability to model the SES-health link over people's lives - as is attempted in this project - would significantly improve the tools currently available to support complex lifecourse related policy decisions.

A1 DYNAMOD

The dynamic micro-simulation model, DYNAMOD, to which the health module will be added, was developed at the National Centre for Social and Economic Modelling, University of Canberra. It simulates future events occurring in Australians' lives: such as couple formation, birth of a child, education, leaving home, migration, divorce, being employed, income from work and government, wealth accumulation, disability and death (Arntcliff *et al.*, 1996). The model is programmed in C and its Base year dataset is drawn from a 1 per cent sample of the Australian population (150,000 persons), extracted by the Australian Bureau of Statistics from its 1986 Census.

A2 Equations for computing mortality rates

For the disabled, mortality rates, at age x and with SES y, have been approximated by $q^d_{x,y}$ which is defined as

$$\frac{\text{no. of deaths for disabled aged } x \text{ in quintile } y \text{ due to non-external causes}}{\text{disabled populations aged } x \text{ in quintile } y}$$

Similarly, the mortality rates covering deaths from all causes, $q_{x,y}$, were calculated as:

$$q_{x,y} = \frac{\text{no. of deaths for disabled aged } x \text{ in quintile } y}{\text{disabled populations aged } x \text{ in quintile } y}$$

Thus,

$$q^d_{x,y} = q_{x,y} \text{x} \frac{\text{no. of deaths due to non-external causes}}{\text{total no. of deaths of those aged (x,y)}} \text{x} \frac{\text{total population(x,y)}}{\text{disabled populations aged (x,y)}}$$

If $p_{x,y}$ is the prevalence of disability at age x then

$$\frac{\text{total population(x,y)}}{\text{disabled populations aged (x,y)}} = \frac{1}{p_{x,y}}$$

Since the prevalence $p_{x,y}$ is known, we only need to compute the proportion of deaths due to non-external causes to obtain an estimate for $q^d_{x,y}$.

The equation for the able bodied population is similar, except that the mortality statistics are for external causes and the prevalence is for the able bodied:

$$\frac{\text{total population(x,y)}}{\text{able bodied population aged (x,y)}} = \frac{1}{1 - p_{x,y}}$$

A3 Equations linking disability and mortality

When computing the input data, disability decrements rates were derived using standard multiple decrement techniques (Arntcliff *et al.*, 1996). Briefly, the able bodied population is subject to exits due to death and entrants due to recovery, while the disabled population is subject to exits due to death and recovery and entrants due to disability onsets. Let:

$l^a_{x,y}$ be the able-bodied population aged x, with family SES quintile y

$l^d_{x,y}$ be the disabled population aged x, with family SES quintile y

$R_{x,y}$ be the number of recoveries aged x, with family SES quintile y

$D_{x,y}$ be the number of people becoming disabled, aged x, SES quintile y.

Then, in calculating independent death rates for the able-bodied population, those initially exposed to risk are given by:

$$E^a_{x,y} = l^a_{x,y} + \frac{1}{2}R_{x,y} - \frac{1}{2}D_{x,y}$$

and similarly for the disabled population

$$E^d_{x,y} = l^d_{x,y} + \frac{1}{2}R_{x,y} - \frac{1}{2}D_{x,y}$$

Then if $q^2_{x,y}$ is the mortality rate for the able-bodied population and $q^d_{x,y}$ the mortality rate for the disabled population, we have

$$E^a_{x,y}q^a_{x,y} + E^d_{x,y}q^d_{x,y} = E_{x,y}q_{x,y}$$

where $E_{x,y}$ is the initial exposed risk for all persons wthin that particular SES quintile. Thus,

$$q^d_{x,y} = \frac{l^a_{x,y} + l^d_{x,y}q_{x,y} - (l^a_{x,y} + \frac{1}{2}D_{x,y})q^a_{x,y}}{l^d_{x,y} - \frac{1}{2}R_{x,y} + \frac{1}{2}D_{x,y}} \tag{1}$$

Furthermore,

$$l^a_{(x+1),y} = l^a_{x,y} - D_{x,y} + R_{x,y} - (l^a_{x,y} + \frac{1}{2}R_{x,y} - \frac{1}{2}D_{x,y})q^a_{x,y} \tag{2}$$

and

$$l^d_{(x+1),y} = l^d_{x,y} - D_{x,y} + R_{x,y} - (l^d_{x,y} + \frac{1}{2}R_{x,y} - \frac{1}{2}D_{x,y})q^d_{x,y} \qquad (3)$$

The able bodied and disabled populations can be related at any point in time using the prevalence rate for the appropriate age within the SES quintile being studied. For example:

$$l^d_{(x+1),y} = l^a_{(x+1),y}\frac{p_{(x+1),y}}{1 - p_{(x+1),y}}$$

which, upon substituting from expression (1), (2) and (3), gives for each SES quintile

$$l^d_x - R_x + D_x - (l^d_x - \frac{1}{2}R_x + \frac{1}{2}D_x)\frac{(l^a_x + l^d_x)q_x - (l^a_x + \frac{1}{2}R_x - \frac{1}{2}D_x)q^a_x}{l^d_x - \frac{1}{2}R_x + \frac{1}{2}D_x}$$
$$= \frac{p_{x+1}}{1 - p_{x+1}}(l^a_x - D_x + R_x - (l^a_x + \frac{1}{2}D_x)q^a_x)$$

This equation is linear in $D_{x,y}$, and therefore can be solved as follows:

$$D_{x,y} = \frac{p_{(x+1),y}l^a_{x,y} + R_{x,y} - q^a_{x,y}(l^a_{x,y} + \frac{1}{2}R_{x,y})}{1 - \frac{1}{2}q^a_{x,y}} -$$
$$\frac{(1 - p_{x+1}, y)\{I^d_{x,y}(1 - q_{x,y}) - 1^a_{x,y}q_{x,y}\}}{1 - \frac{1}{2}q^a_{x,y}}$$

For any given values of $R_{x,y}$ this formula can be used to build up a double decrement table for death and disability. Values for $R_{x,y}$ have been chosen to ensure that $D_{x,y}$ remains positive for all ages within the SES quintile being studied and starts a 4% of the disabled population at age 0, increases to 15% of the disabled population over the age range 11 to 40 and then gradually declines to zero by age 94. Independent rates of increments due to disability and recovery, $r_{x,y}$ and $d_{x,y}$ can be calculated from $R_{x,y}$ and $D_{x,y}$ using the relationships:

$$r_{x,y} = \frac{R_{x,y}}{l^d_{x,y} + \frac{1}{2}D_{x,y} - \frac{1}{2}\theta^d_{x,y}}$$
$$d_{x,y} = \frac{D_{x,y}}{l^a_{x,y} + \frac{1}{2}R_{x,y} - \frac{1}{2}\theta^a_{x,y}}$$

During the simulation part of the model, survival functions are established which allow estimation of the time of death if/when a person becomes disabled (Arntcliff *et al.*, 1996).

REFERENCES

ABS (Australian Bureau of Statistics) (1996). *1995 National Health Survey: Users' Guide.* Cat. No. 4363.0, ABS, Canberra.

ABS (Australian Bureau of Statistics) (1999). *Disability Ageing and Carers, 1998, User Guide.* Cat No. 4431.0, ABS, Canberra.

Acheson, D. (1998). *Independent Inquiry Into Inequalities in Health Report.* London: The Stationary Office.

Antcliff, S., Bracher, M., Gruskin, A., Hardin, A. and Kapuscinski, C, (1996). Development of DYNAMOD: 1993 and 1994, National Centre for Social and Economic Modelling, Dynamic Modelling Working Paper No. 1, University of Canberra, and later DYNAMOD Technical Papers.

Berney, L., Blane, D., Davey Smith, G. and Holland, P. (2001). Lifecourse influences on health in old age. In *Understanding Health Inequalities*, Graham, H. (editor). Philadelphia: Open University Press.

Dunn, C., Sadkowski, K. and McPherson, M. (2002). Mortality Surveillance, Australia, 1987-1998. *Australian Institute of Health and Welfare*, forthcoming.

Glover, J., Harris, K. and Tennant, S. (1999). *A Social Health Atlas of Australia.* Commonwealth of Australia. Second Edition

King's Fund (1999). Local Inequalities Targets. UK: King's Fund Publications.

Marmot, M.G. (1998). Contribution of psychosocial factors to socio-economic differences in health. *Millibank Quarterly* 76, 403-48.

McCain, M. and Mustard, J. (1999). *Early Years Study.* Canada: Government of Ontario.

Mitchell, R., Shaw, M. and Dorling, D. (2000). *Inequalities in Life and Death: What if Britain Were More Equal?.* Bristol: Policy Press.

Saunders, P. (1996). Poverty, Income Distribution and Health: An Australian Study. *Social Policy Research Centre University of New South Wales.*

Turell, G and Mathers, C. 2001. Socioeconomic Inequalities in All Cause and Specific Cause Mortality in Australia: 1985-1987 and 1995-1997. *International Journal of Epidemiology* 30, 231-239.

Walker, A. (1998). Australia's Ageing Population: What Are The Key Issues and The Available Methods of Analysis. *Discussion Paper 27*, National Centre for Social and Economic Modelling University of Canberra.

Walker, A. (2000). Measuring the Health Gap Between Low Income and Other Australians, 1977 to 1995: Methodological Issues, *Discussion Paper 50.* National Centre for Social and Economic Modelling, University of Canberra.

Walker, A. (2001). Assessing health inequalities using a dynamic microsimulation model. Presented at the *Health Services & Policy Research Conference*, Victoria University, Wellington, New Zealand, 2-4 December, Section 5.

Walker, A. and Abello, A. (2000). Changes in the Health Status of Low Income Groups in Australia, 1977-78 to 1995. *Discussion Paper 52*, National Centre for Social and Economic Modelling, University of Canberra.

World Health Organization (2000). *The World Health Report 2000 - Health Systems: Improving Performance*, Geneva.

HYPOTHESIS TESTING OF MULTIVARIATE ABUNDANCE DATA

DAVID WARTON AND MALCOLM HUDSON
Department of Statistics
Division of Economic and Financial Studies
Macquarie University, NSW 2109, Australia
E-mail: DWarton@rna.bio.mq.edu.au, Malcolm.Hudson@mq.edu.au

Multivariate abundances are commonly collected in ecology and related disciplines. The main difficulties in modelling such data are that the number of variables is usually larger than the number of samples, and that the distribution of measurement variables is typically strongly right-skewed with a positive weight at zero. We consider methods of hypothesis testing for multivariate abundances, where permutation tests (permuting rows of the data matrix) are applicable. A number of such tests have been proposed in the ecological literature. We suggest alternative methods, and compare the different approaches using P-values from 20 datasets extracted from the ecological literature. All P-values were calculated using permutation tests. Statistics based on generalised linear models performed poorly in unbalanced designs, but other approaches provided comparable or more encouraging results than the test statistics presently used in ecology.

Keywords: multivariate analysis, non-normal data, zero-modified distributions, permutation tests.

1 Introduction

Multivariate abundances are quite commonly sampled in ecological research. In fact, textbooks have been written about the analysis of such data (Gauch 1982, Jongman *et al.* 1987). Typical measures of abundance are counts of organisms per trap and % projected cover in the sample area. Abundance measures are separately recorded for different taxonomic groups. Hence a multivariate abundance dataset $\mathbf{y}$ will have dimensions $N \times p$, corresponding to N multivariate samples of the abundance of p taxa.

There are two properties of multivariate abundance data which demand attention.

- Data depart markedly from a multivariate normal distribution. In particular, the distribution of abundances is strongly right-skewed with a positive weight at zero.

- The number of variables is large relative to the number of replicate samples.

Both of these features are apparent in the example dataset (Table 1).

Table 1. Example multivariate abundance dataset (Moulton 1982). Each row represents a single multivariate observation, classified into one of three treatments, each column represents one variable (abundances in one taxon). Twenty-one variables are not shown, so $p = 36$, whereas $N = 14$. Note also that most variables are strongly right-skewed, usually with a high probability of observing a zero.

														...		
Treatment 1	0	13	3	0	1	0	1	0	0	1	0	0	0	...	0	1
	0	4	6	0	0	0	0	0	0	0	0	0	0	...	1	1
	0	12	0	2	7	3	0	0	1	0	0	1	0	...	3	0
	1	23	1	1	0	0	1	0	0	0	1	0	0	...	0	0
	0	44	6	0	1	3	0	1	1	18	0	4	0	...	0	4
	0	6	1	1	0	0	0	0	0	0	2	8	0	...	0	1
Treatment 2	0	22	1	0	0	6	1	0	0	0	0	1	0	...	10	0
	0	12	8	0	0	2	0	4	1	7	1	2	0	...	29	0
	0	7	4	0	0	1	0	0	1	0	0	2	0	...	0	0
	0	24	5	0	0	0	0	0	1	1	0	2	1	...	0	0
Treatment 3	0	12	1	0	1	6	0	2	2	18	0	5	0	...	0	0
	1	22	3	0	2	4	0	2	1	11	2	38	1	...	0	0
	1	35	3	0	0	7	0	13	0	16	6	2	4	...	0	0
	1	13	0	0	0	2	0	4	4	4	4	1	1	...	0	0

This paper considers the case where the N multivariate samples have been classified into groups *a priori*, and there is interest in testing for differences in abundance across sets of grouping variables. Differences in abundance are assumed to be expressed through location parameters, and tested simultaneously for all taxa, using a single overall test. We consider tests of a single factor and tests of a nested factor in a multi-way layout.

An application of particular interest is assessing environmental impacts, when the N multivariate observations are classified as "impacted" or "control", for example, where there may be a blocking variable (*e.g.* topology - ridge/hill/gully). The example in Table 1, however, contains three treatments (corresponding to three experimental fires) and no blocking variable. Tabulated values are counts of invertebrates in different taxa, and there is interest in testing for location differences in taxon abundance across the fires.

Note that exact significance levels can be found for tests of the form specified above using restricted permutation testing (Edgington 1995).

For tests of the type described above, several procedures have been suggested in the ecological literature, as briefly reviewed in Section 2. Alternative approaches are then reviewed in Sections 3 and 4. There is no existing work comparing different test statistics for this type of data, the subject of Section 5.

2 Methods of analysis presently used

Practising ecologists presently use one of two types of test statistics - those found in a package known as CANOCO (ter Braak and Smilauer 1998), and those based on a matrix of pair-wise dissimilarities between observations. In all cases, the significance of a test statistic is found by resampling, and all variables are assumed to be independent in deriving the test statistics.

CANOCO methods Test statistics in the CANOCO package are explained in its documentation (ter Braak and Smilauer 1998), and are either referred to as "redundancy analysis" (RDA) or "canonical correspondence analysis" (CCA). Both RDA and CCA are scale dependent statistics, although in CCA data are transformed using sums of abundances in such a way that if abundances follow a Poisson distribution, they are standardised to have equal variance (ter Braak and Smilauer 1998).

Distance-based approaches In distance-based approaches, test statistics are functions of a matrix of pair-wise distances between multivariate observations. In ecology, the most common example is the Bray-Curtis distance, defined as

$$d_{ij} = \sum_{k=1}^{p} \frac{|y_{ik} - y_{jk}|}{y_{ik} + y_{jk}}. \tag{1}$$

The various test statistics which are used can all be expressed in the form of the statistic suggested by Mielke *et al.* (1976): for each of the g groups, we define the average within-group distance

$$\delta_k = \sum_{i=1}^{n_k} \sum_{j=i+1}^{n_k} \frac{d_{ij}^r}{n_k(n_k-1)} \tag{2}$$

and the test statistic is a weighted average of these,

$$\sum_{k=1}^{g} w_k \delta_k. \tag{3}$$

Here n_k is the number of observations in the kth group, a typical choice of r is 1 or 2, and we take w_k to be $(n_k - 1)/(N - g)$, although n_k/N is another common choice (Mielke and Berry 2001). Under H_1, this statistic will be unusually small compared to other permutations of the data.

Other statistics have been suggested in the ecological literature, which are monotonic functions of the Mielke *et al.* (1976) statistic. For example,

choosing $w_k = (n_k - 1)/(N - g)$ and $r = 2$ leads to a test statistic recently proposed (Anderson 2001). If in addition the Euclidean distance is used, this statistic is equivalent to RDA, and to statistics suggested by others (Romesburg 1985, Edgington 1995).

Clarke (1993) suggested a test statistic based on ranks of distances. It is a monotonic function of the Mielke *et al.* (1976) statistic if ranks of distances are used, with $w_k = n_k(n_k - 1)$ and $r = 1$.

A number of alternative approaches will also be considered in this study, and these are reviewed in the following sections. These approaches have been organised into sections to separately address the two important properties of multivariate abundances that were previously identified - nonnormality, and the large number of variables.

3 Approaches for non-normal response variables

A range of approaches are considered here to address the non-normality of abundances.

Transformation Data are transformed to approximate normality, and least-squares theory used. The transformations considered are $\sqrt[4]{\mathbf{y}}$ and $\log(\mathbf{y} + a) - \log(a)$, where a is the minimum possible non-zero abundance. The most common transformations used in practice for analysis of univariate abundances are $\sqrt[4]{\mathbf{y}}$ and $\log(\mathbf{y} + 1)$, but a modification of the log-transformation is used here so that the transformation is scale independent. Because data are typically overdispersed relative to the Poisson distribution, $\sqrt{\mathbf{y}}$ (variance-stabilising transform for the Poisson) is not considered.

Log-linear models Data are often obtained as counts, and so log-linear models (McCullagh and Nelder 1989) can be applied. We consider the Poisson (with overdispersion parameter) and negative binomial model, because counts are usually strongly right-skewed. Parameters of multivariate log-linear models are estimated using generalised estimating equations (Liang and Zeger 1986), with nuisance parameters estimated by moments. As a test statistic, Rao's score statistic (Rao 1948) is used. Some datasets did not contain counted data - in such cases data were transformed to $\frac{\mathbf{y}}{a}$, a previously defined, which would be expected to behave like counts.

Robust methods Rank-based analyses (Hollander and Wolfe 1999) and other robust approaches (Birkes and Dodge 1993) are obviously appro-

priate. The Kruskal-Wallis test and least-deviations or L_1 regression are considered here. For L_1 regression, a univariate statistic suggested by Birkes and Dodge (Birkes and Dodge 1993) is used.

Mixture modelling We also consider modelling the abundance in each taxon using a lognormal distribution, conditional on the presence of the taxon in a sample, which is modelled using the Bernoulli distribution. This approach has recently been used to model willingness-to-pay (Reiser and Shechter 1999). In the univariate case, a maximum likelihood test statistic consists of two additive components - a logistic regression term (for presence/absence) and a sums of squares term (for non-zero abundances). This decomposition of the test statistic is particularly attractive from the practitioner's viewpoint.

4 Approaches for multivariate data with a large number of variables

Relative to the number of samples taken, the number of variables is frequently large. When p is close to N, sample correlation and covariance matrices are near singular, and test statistics give spurious results. When $p > N$, as is often the case, test statistics requiring the inverse of a sample variance or correlation matrix can not be found (due to singularity).

Data reduction The number of variables could be reduced to a more manageable number prior to analysis, whether using additional information (functional groups, taxonomy) or data-driven approaches (*e.g.* principal components analysis). This is not considered further in this paper.

Subset selection Rather than construct a test statistic using all available variables, a subset could be selected to optimise some criterion, *e.g.* include the k variables which maximise the value of a given test statistic. This approach is presently being evaluated.

Simplified correlation A simplified correlation structure could be assumed in constructing test statistics. Two correlation structures are considered here:

- All variables are independent, so $\mathbf{R}$, the correlation matrix, is the identity.
- Groups of variables are independent of others, $\mathbf{R}$ is block diagonal. The grouping of variables will be arbitrary, the group size taken to

be $(N-g)/3$ (simulations suggest larger block sizes can give spurious results).

When all variables are assumed independent, each test statistic is a sum of univariate statistics. This general approach to multivariate analysis dates back to 1958 (Chung and Fraser 1958). The least-squares statistics, with nuisance parameters estimated under H_1, are then a sum of univariate F statistics, an approach which has previously been suggested (Edgington 1995). For robust methods and mixture models, the identity matrix is the only assumed correlation structure considered in this paper.

For all of the test statistics considered here, resampling is used for inference, as it can not be assumed that the simplified correlation structure is true.

5 Main Results

There are two distinct questions of interest, relating to power considerations. Questions are:

Under what conditions will one test be more powerful than others?

This is an area that has not been investigated at all for multivariate abundance data. An outline is presented of how the statistics of Section 1 may differ in power, given the significance level of each is found through permutation.

Lack of standardisation of data Remarkably, it is not presently routine in ecology to standardise abundances (or transformed abundances) according to the variability of different taxa. Ensuing test statistics would be more sensitive to effects amongst more variable taxa, and insensitive to effects amongst taxa with less variable abundances.

Assumptions of correlation matrix As verified by simulation (Mielke and Berry 2001), rotation invariant statistics that fail to account for correlation between variables will be more sensitive to location shifts along the first few principal component axes that to location shifts along other axes (where the principal components axes have been calculated from the sample correlation matrix of $\mathbf{y}$). Although this result will not strictly be true of rotation variant statistics, it is indicative of what one might expect in general.

Robustness to outliers If abundances contain outliers, then least squares methods will lack power, compared to more robust methods (Birkes and

Dodge 1993). Methods based on the Bray-Curtis distance would be expected to be robust in this sense.

Choice of abundance model If a particular model for abundances is more appropriate than others, statistics based on this model would be expected to generally be more powerful.

Closer consideration of the above issues is required.

In practice, how often is one statistic more powerful than others?

Different test statistics will be sensitive to different alternatives, and it is of interest to see if a particular type of alternative is frequently encountered in practice.

We addressed this question using 20 datasets taken from applications (Table 2). The use of 20 datasets from the literature ensured data were representative of multivariate abundance datasets encountered in practice. In preliminary work to compare the power of different statistics, P-values were calculated for each dataset, and compared across methods. A statistic that is

Table 2. Dataset source, and properties. "Sub-sampled" refers to whether the dataset was reduced in size by selecting a subset of all treatments, variables, observations or by analysing only one of several sampling times. For unpublished data, the name of the provider of the data is given.

Source	Abundance	Organisms	Sub-sampled:	Nested factor?
N. Andrews	count	invertebrates	time	yes
A. Pik	count	invertebrates		yes
Moulton (1982)	count	invertebrates	obs., time	yes
Moulton (1982)	count	invertebrates	time	
I. Lunt	count	plants	time	
B. Rice	% cover	plants	treatments	
J. Overton	% cover	plants		
B. Rice	% cover	plants	treatments, time	
B. Rice	% cover	plants	excluded	
Clements (1980)	% cover	plants	treat., var., time	
A. Pik	count	invertebrates		yes
Warwick *et al.* (1990b)	count	invertebrates	treatments	
Warwick (1971)	biomass	invertebrates		
Gray *et al.* (1990)	count	invertebrates	treatments	
Warwick *et al.* (1988)	count	invertebrates	treatments	
Gee *et al.* (1985)	count	invertebrates		
Warwick *et al.* (1990a)	count	invertebrates		
van Dobben *et al.* (1999)	count	invertebrates		
van den Brink *et al.* (1996)	count	invertebrates	time	
van der Aart and Smeenk-Enserink (1970)	count	invertebrates		

generally more powerful than others in practice will provide generally lower P-values. Subsequent power simulations are required to strengthen conclusions.

A hypothesis test was chosen such that most statistics suggest there is some evidence (in the range $0.001 < P < 0.1$) against the null. This ensured that in most cases, H_1 was true but there was not such overwhelming evidence for it that comparison of test statistics was trivial. To satisfy these conditions, subsets of some datasets were used, as indicated in Table 2. One dataset was excluded due to lack of evidence for H_1 ($0.1 < P < 0.6$ for most methods).

Variables with only one (or no) non-zero elements were excluded.

For all test statistics, P-values were determined by restricted permutation testing (Edgington 1995). This involved permuting observations (rows of $\mathbf{y}$) within each of the groups defined under H_0, except for datasets which included nested factors (indicated in Table 2) such that it was appropriate to permute blocks of $\mathbf{y}$ rather than observations. P-values were estimated from ten thousand random permutations, except when exact P-values could be found from less than ten thousand systematic permutations.

P-values are usually interpreted on a proportional scale, so instead of arithmetic mean P-values we used geometric means (Table 3) and P-values were plotted on a logarithmic scale (Figure 1).

Table 3. Geometric mean of P-values from 19 datasets. P-values estimated by permutation testing with 10,000 iterations. For all "test in literature", data were transformed to $\sqrt[4]{\mathbf{y}}$, Bray-Curtis distances used (where relevant), and "Mielke *et al.* (1976)" used $w_k = (n_k - 1)/(N - g)$. **R** indicates correlation matrix assumed to be block diagonal, **I** indicates correlation matrix assumed to be the identity matrix, H_0 indicates nuisance parameters were estimated under H_0. Values in table are geometric mean P-value×1000.

		R	**R**, H_0	**I**	**I**, H_0
transformation	$\mathbf{y}$ (none)	63	41	36	33
	$\log(\mathbf{y}+1)$	28	24	16	17
	$\sqrt[4]{\mathbf{y}}$	38	26	28	17
loglinear model	Poisson	89	37	34	26
	Negative binomial	60	31	21	27
robust methods	Kruskal-Wallis				15
	L_1 regression				38
mixture model					15
tests in literature:	RDA			14	
	CCA			19	
	Mielke *et al.* (1976), $r = 1$			14	
	Mielke *et al.* (1976), $r = 2$			15	
	Clarke (1993)			19	

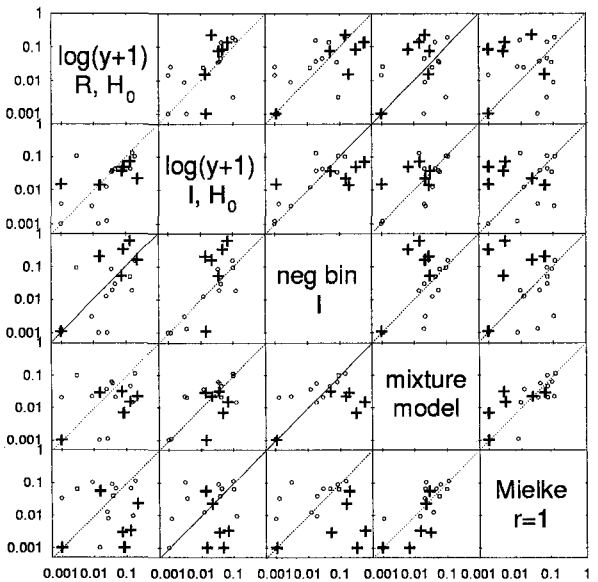


Figure 1. Scatterplot matrix of P-values from 19 datasets, calculated by five different methods that are labelled as in Table 3. + indicates a severely unbalanced design, with sample sizes in different groups varying over a factor of two.

6 Discussion

Results (Table 3) suggest the following conclusions for multivariate abundances.

- Least-squares has generally higher power following transformation of data, and loglinear models generally have low power. This is particularly true for unbalanced designs.
- L_1 regression has low power. It models the median, and so only detects effects expressed at the median, whereas often effects were observed in variables where more than half the values were zero.
- For these data, there is little obvious advantage in trying to model the dependence of variables (block diagonal $\mathbf{R}$), rather than using the independence model.
- Estimating nuisance parameters under H_0 rather than H_1 can be ad-

vantageous. This effect was only present in 3 datasets, for which many variables were rarely non-zero.

Otherwise, Table 3 suggests similar performance between methods reviewed in Section 1. Comparing individual P-values, however, it is clear that different test statistics often lead to different interpretations of a given dataset (Figure 1), most obviously for unbalanced data.

7 Conclusion

The above analysis was sufficient to identify methods that are clearly inadequate for multivariate abundance data (loglinear models, L_1 regression, untransformed least-squares).

In the place of methods presently used in the literature, results provide some justification for assuming independence, estimating nuisance parameters under H_0, and using either a transformed least-squares approach or a simple mixture model. These methods lead to test statistics which have particularly simple forms, and can be easily decomposed into the contributions of different variables. However, these results are only preliminary, and require support from simulations of data from known distributions, as presented in forthcoming publications.

REFERENCES

Anderson, M.J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology* 26, 32-46.

Birkes, D. and Dodge, Y. (1993). *Alternative Methods of Regression.* New York: John Wiley and Sons.

Chung, J.H. and Fraser, D.A.S. (1958). Randomization tests for a multivariate two-sample problem. *Journal of the American Statistical Association* 53, 729-735.

Clarke, K.R. (1993). Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 18, 117-143.

Clements, A. (1980). The Vegetation of Bushland in the Northern Sydney area. *M.Sc. Thesis.* Macquarie University.

Edgington, E.S. (1995). *Randomization Tests*, Third edition. New York: Marcel Dekker.

Gauch, H.G. Jnr. (1982). *Multivariate Analysis in Community Ecology.* Cambridge: Cambridge University Press.

Gee, J.M., Warwick, R.M., Schaanning, M., Berge, J.A. and Ambrose, W.G. Jr. (1985). Effects of organic enrichment on meiofaunal abundance and

community structure in sublittoral soft sediments. *Journal of Experimental Marine Biology and Ecology* 91, 247-262.

Gray, J.S., Clarke, K.R., Warwick, R.M., Hobbs, G. (1990). Detection of initial effects of pollution on marine benthos: an example from the Ekofisk and Eldfisk oilfields, North Sea. *Marine Ecology Progress Series* 66, 285-299.

Hollander, M. and Wolfe, D.A. (1999). *Nonparametric Statistical Methods.* New York: John Wiley and Sons.

Jongman, R.H.G., ter Braak, C.J.F. and van Tongeren, O.F.R. (1987). *Data Analysis in Community and Landscape Ecology.* Wageningen: Pudoc.

Liang, K.Y. and Zeger, S.L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* 73, 13-22.

McCullagh, P. and Nelder, J.A. (1989). *Generalised Linear Models*, Second edition. Boca Raton: Chapman and Hall.

Mielke, P.W. and Berry, K.J. (2001). *Permutation Methods - a Distance Function Approach.* New York: Springer-Verlag.

Mielke, P.W., Berry, K.J. and Johnson, E.S. (1976). Multi-response permutation procedures for *a priori* classifications. *Communications in Statistics - Theory and Methods* 5, 1409-1424.

Moulton, T.P. (1982). The Effect of Prescribed Burning and Simulated Burning on Soil and Litter Arthropods in Open Forest at Cordeaux, N.S.W, Australia. *Ph.D. Thesis*, Macquarie University.

Rao, C.R. (1948). Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society* 44, 50-57.

Reiser, B. and Shechter, M. (1999). Incorporating zero values in the economic valuation of environmental program benefits. *Environmetrics* 10, 87-101.

Romesburg, H.C. (1985). Exploring, confirming, and randomization tests. *Computers and Geosciences* 11, 19-37.

ter Braak, C.J.F. and Smilauer, P. (1998). CANOCO Reference manual and User's Guide to CANOCO for Windows: Software for Canonical Community Ordination (version 4). New York: Microcomputer Power.

van den Brink, P.J., van Wijngaarden, R.P.A., Lucassen, W.G.H., Brock, T.C.M., Leeuwangh, P. (1996). Effects of the Insecticide Dursban 4E (active ingredient chlorpyrifos) in outdoor experimental ditches: II. Invertebrate community responses and recovery, *Environmental Toxicology and Chemistry* 15, 1143-1153.

van der Aart, P.J.M. and Smeenk-Enserink, N. (1970). Correlations between distribution of hunting spiders (Lycosidae, Ctenidae) and environmental characteristics in a dune area. *Netherlands Journal of Zoology* 25, 1-45.

van Dobben, H.F., ter Braak, C.J.F. and Dirkse, G.M. (1999). Undergrowth as a biomonitor for deposition of nitrogen and acidity in pine forest. *Forest Ecology and Management* 114, 83-95.

Warwick, R.M. (1971). Nematode associations in the Exe estuary. *Journal of the Marine Biology Association of the United Kingdom* 51, 439-454.

Warwick, R.M., Carr, M.R., Clarke, K.R., Gee, J.M. and Green, R.H. (1988). A mesocosm experiment on the effects of hydrocarbon and copper pollution on a sublittoral soft-sediment meiobenthic community. *Marine Ecology Progress Series* 46, 181-191.

Warwick, R.M., Clarke, K.R. and Gee, J.M. (1990a). The effect of disturbance by soldier crabs, Mictyris platycheles H. Milne Edwards, on meiobenthic community structure. *Journal of Experimental Marine Biology and Ecology* 135, 19-33.

Warwick, R.M., Platt, H.M., Clarke, K.R., Agard, J. and Gobin, J. (1990b). Analysis of macrobenthic and meiobenthic community structure in relation to pollution and disturbance in Hamilton Harbour, Bermuda. *Journal of Experimental Marine Biology and Ecology* 138, 119-142.

ESTIMATION OF PARAMETERS IN PYROLYSIS KINETICS

JASON M. WHYTE
Department of Applied Mathematics
Adelaide University
Adelaide, SA 5005, Australia
E-mail: jwhyte@maths.adelaide.edu.au

MARK A. SUGDEN AND GEOFFREY D. ABBOTT
Department of Fossil Fuels and Environmental Geochemistry
University of Newcastle upon Tyne
Newcastle upon Tyne, NE1 7RU, England

ANDREW V. METCALFE AND CHARLES E.M. PEARCE
Department of Applied Mathematics
Adelaide University
Adelaide, SA 5005, Australia

A linear time invariant state space model is proposed for the production and decay of two epimers (R) and (S) of a hopane released from oil bearing rock during laboratory pyrolysis. Concentrations of R and S are measured over time. The parameters to be estimated are: the initial amounts of precursors X, for R, and Y, for S; the rate constants for the production of R and S; the rate constants for the decay of R and S; and the rate constants for the two-way epimerization between both X and Y and R and S. It is shown that the model is locally identifiable. The parameters are estimated by numerical integration of the rate equations, alternated with a derivative free least squares constrained, optimisation routine. Asymptotic standard errors and covariances of parameters are given and compared with those obtained from a resampling approach (parametric bootstrap). An alternative fitting procedure, based on estimating derivatives of the concentrations of R and S by fitting splines, is implemented and compared with that based on integration of the rate equations. The rate constants are important for elucidation of the reaction pathways, and the estimates of initial concentrations of X and Y have potential for inferring the yields of oil bearing rock.

Keywords: compartmental model, identifiability, resampling, splines, state-space model.

1 Introduction

Biological markers (biomarkers) are compounds detected in petroleum or rock extracts which indicate an unambiguous link with a natural product. Chemical changes which occur in the sediment may alter functional groups and

bonds in the natural product, but the carbon skeleton of the compound remains basically the same (Mackenzie, 1984). Hopanes are a class of biomarkers routinely found in buried sedimentary organic matter and are derived from the defunctionalisation of precursors with the biologically inherited (R)- configuration at C-22. The side chain $22R$-configuration dominates the isomeric distributions in the organic solvent extract of shallow, relatively cool sediments (Philp, 1985).

With increasing burial depth and the associated temperature rise the side chain R-configuration preferred in immature sediments converts gradually to an all-isomer mixture of R and S epimers. This process is gauged by measuring a biomarker maturity parameter from gas chromatography - mass spectrometry (GC-MS) data - the $17\alpha(H), 21\beta(H)22S/(22S + 22R)C_{31}$ homohopane parameter abbreviated to $(22S/(22S + 22R))$, is given by the peak area of the $22R$ epimer relative to the sum of itself and the peak area of its $22S$ counterpart in the m/z =191 mass chromatogram. These parameters are of great interest to the oil industry because they indicate the extent of the thermally-driven reactions which convert sedimentary organic matter into petroleum. The biomarker maturity parameter ranges from 0 to an end point of about 0.60, and high values suggest the rocks in a sedimentary basin have been buried to the depths necessary, about 3 km, for significant and economically exploitable petroleum generation to have occurred.

Kerogen is defined as the organic constituent of sedimentary rocks that is neither soluble in aqueous alkaline solvents, nor in the common organic solvents. Biomarkers may be trapped in the kerogen network or, alternatively, bound to kerogen by covalent bonds and then are released as a result of increasing burial within the Earth's crust. Generally biomarker maturity parameters increase with time and temperature during the pyrolysis of sedimentary organic matter (Lewan et al, 1986). In the 1980s it was held that the inherited biological (R)-configuration at C-22 converts to an epimeric mixture during burial and the associated thermal maturation, see Mackenzie (1984) and Mackenzie and McKenzie (1983). However, direct chiral isomerization at C-22 for hopanes did not appear to be the sole control on observed changes in the $22S/(22S + 22R)$ maturity parameter, during the hydrous pyrolysis of some kerogens (see Abbott et al., 1990)). Further controlling factors were cited which included sequential release of (R)- and possibly (S)-epimers from macromolecular and/or functionalised components, combined with degradation of both (R)- and (S)-epimers to mainly unknown products. Laboratory pyrolysis of kerogens or organic-rich rocks at temperatures in excess of 250 ^oC causes the release of bound biomarkers.

We analyse data from a laboratory experiment on a sample of Messel

shale. The free biomarker, R and S, was washed out with solvent over a two day period, and the sample was then subjected to pyrolysis, in 23 small glass tubes, in an oven which was maintained at $350°C$. These tubes were removed at different times throughout the pyrolysis and analysed for R and S content. Therefore, the laboratory experiment provided a set of n time (t) and estimated epimer concentrations, $[R]_t$ and $[S]_t$, triples. The most general chemical kinetic model considered here is:

$$\begin{array}{ccccc} X & \xrightarrow{k_3} & R & \xrightarrow{k_7} & D_R \\ k_2 \uparrow\downarrow k_1 & & k_6 \uparrow\downarrow k_5 & & \\ Y & \xrightarrow{k_4} & S & \xrightarrow{k_8} & D_S \end{array}$$

In this model, X and Y are precursors, referred to as bound biomarker, for R and S, and D_R and D_S are the degradation products. The dominant pathways in laboratory pyrolysis are X to R to D_R and Y to S to D_S. However, epimerization between X and Y, and between R and S occurs in the ground and this may be emulated in laboratory pyrolysis.

The washing before pyrolysis justifies an assumption that the initial concentrations of R and S are zero. The statistical problem is to estimate the rate constants $k_1, \ldots, k_8$, possibly with some constraints in addition to their being positive, and the initial concentrations of X and Y from the data which are subject to experimental errors. The initial concentrations of X and Y are of considerable practical importance. Their sum provides an estimate of amount of oil in the rock. The ratio $[Y]_0/([X]_0 + [Y]_0)$ is the bound maturity parameter. It has a similar range of values to the maturity parameter based on the free R and S found in the rock, and is potentially more reliable because the free R and S may have come from extraneous sources.

2 Results

2.1 State-space model

The kinetic model is a set of simultaneous first order linear differential equations. It can therefore be written in state space form:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} \\ \mathbf{y} &= \mathbf{C}\mathbf{x} \\ \mathbf{x}(0) &= \mathbf{x_0} \end{aligned} \tag{1}$$

Here $\mathbf{x}$ is the state vector

$$\mathbf{x}' = ([X], [Y], [R], [S])$$

$$\mathbf{x_0}' = ([X]_0, [Y]_0, 0, 0)$$

and $\mathbf{y}$ is the observation vector

$$\mathbf{y}' = ([R], [S])$$

In practice $\mathbf{y}$ will be sampled. The constant matrices $\mathbf{A}$ and $\mathbf{C}$ are defined as follows.

$$\mathbf{A} = \begin{pmatrix} -k_1 - k_3 & k_2 & 0 & 0 \\ k_1 & -k_2 - k_4 & 0 & 0 \\ k_3 & 0 & -k_5 - k_7 & k_6 \\ 0 & k_4 & k_5 & -k_6 - k_8 \end{pmatrix} \qquad \mathbf{C} = \begin{pmatrix} 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{pmatrix}$$

2.2 Identifiability

The model is deterministically identifiable if it is, in principle, possible to estimate all 10 parameters from noise free time series measurements $[R]_t$ and $[S]_t$. However, even if a model is identifiable it may not be possible to estimate the parameters with acceptable precision from finite length, noise corrupted, records. Also there may be different sets of parameter values for which the model's associated differential equations have identical solutions, $[R]_t$ and $[S]_t$. A model is described as locally identifiable provided a set of parameters, corresponding to any given solution of the differential equations, is unique within some neighbourhood of the point values of the parameters. A relevant example is the fast/slow ambiguity: there are two sets of parameters, $[X]_0, k_3$, and k_7, for which the system

$$X \xrightarrow{k_3} R \xrightarrow{k_7} D_R,$$

with $[R]_0$ equal to 0, has identical $[R]_t$, corresponding to the cases $k_3 > k_7$ and $k_3 < k_7$. However, it may be possible to distinguish between the two cases from knowledge of a likely range of values for $[X]_0$.

We should check that a model is identifiable before attempting to estimate its parameters. There are several methods for checking model identifiability; see for example Vajda and Rabitz (1988) or Eisenfeld (1985). Here we have chosen to use the Laplace transformation approach of Vajda and Rabitz (1988).

The Laplace transform of a first order time-invariant, linear differential equation system such as (1) is

$$\mathbf{Y}(s, \mathbf{k}) = \mathbf{C}(\mathbf{k})[s\mathbf{I} - \mathbf{A}(\mathbf{k})]^{-1}\,\mathbf{x}_0(\mathbf{k})$$

where $\mathbf{k}$ is the vector of unknown parameters. The vector $\mathbf{\Phi}$ of moment invariants is formed by all different coefficients of powers of s in the numerator and

denominator of the transfer functions $Y_i(s, \mathbf{k})$, where Y_i are the components of $\mathbf{Y}$. A Jacobian $\mathbf{J}$ is defined as,

$$\mathbf{J} = \frac{\partial \mathbf{\Phi}}{\partial \mathbf{k}}.$$

The system is identifiable, at least locally, if and only if the rank of $\mathbf{J}$ is equal to the number of unknown parameters, except, possibly, for a finite set of point values. Even with computer algebra, this criterion becomes impractical to implement with a large number of unknown parameters, and similar limitations apply to the other methods. In our case we were able to demonstrate that the system is identifiable if the initial conditions are known, from a Jacobian of size 10 by 8, but the symbolic algebra routine could not operate on a matrix of size 10 by 10 which would be needed for the case of unknown initial conditions. However, we claim that if a linear system is identifiable given initial conditions, observable and input connectable (see Davison, 1977), then it is also identifiable with unknown initial conditions. An argument for this follows.

Suppose a system is identifiable given initial conditions (I0). Then given a noise free signal $y(t)$ and I0 we can identify the unknown parameters and get an exact fit. But, we don't know I0. Therefore suppose we assume initial conditions I1. Then, either we do not get an exact fit and we try different initial conditions or we do get an exact fit with I1. The former could continue until we converge on I0, in which case the rate constants and the parameters are identifiable. So, we need only consider the second possibility. The second possibility would imply that non-zero initial conditions (I1-I0) give a zero response. This would contradict the assumption of input connectability. It should be noted that such initial conditions may contain negative components and are therefore not physically realisable. Nevertheless, the mathematical argument holds.

2.3 Fitting the model

It is possible to express the solution to the differential equation describing the system in algebraic form as in Marin et al. (1999). However, we have taken the more general approach of solving the equations numerically. This approach extends to $\mathbf{A}$ matrices of dimensions larger than 4×4. The solution, for a given set of parameter values, is compared with the data and the sum of squared errors is calculated. The parameter values are adjusted and the process is repeated until a least value of the sum of squared errors is obtained. We checked that this routine did identify the parameter values set equal to the estimates identified in Marin et. al (1999) as a test case, from sampled

solution time series R_t and S_t consisting of 400 points. We then fitted the following models to the 23 data triples $(t, [R]_t, [S]_t)$.

Model 0

This is the simplest model in which there is no epimerization between R and S or their precursors X and Y. Thus, the parameters k_1, k_2, k_5 and k_6 are set to 0. The results of fitting subject to the constraint that $[X]_0 > [Y]_0$ are given in Table 1. The constraint was imposed because the Messel shale is known to be thermally immature and the maturity parameter, $[Y]_0/([X]_0 + [Y]_0)$, was therefore expected to be well below 0.5. The residual sum of squares is 102.7. A plot of the fitted values is given in Figure 1.

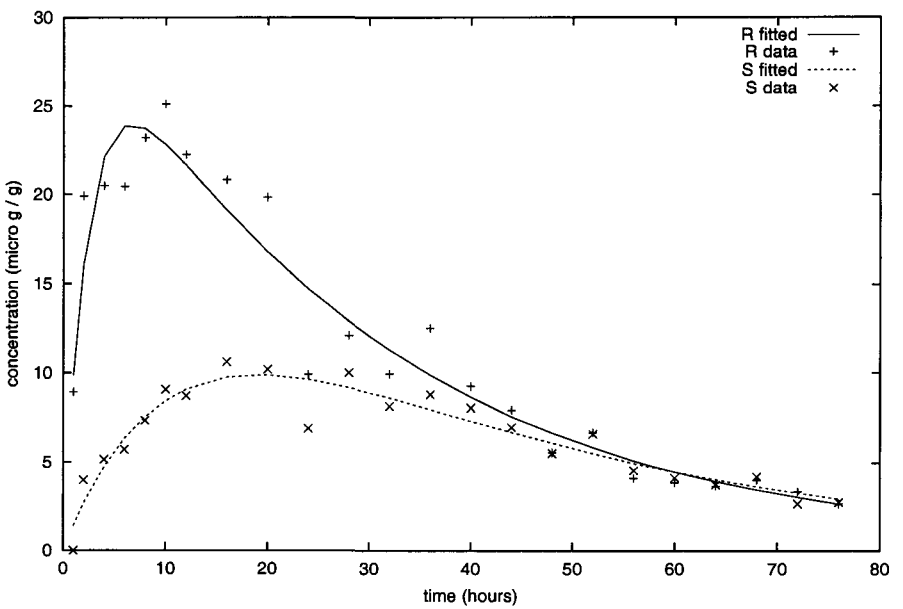


Figure 1. Data for R and S with fitted values for Model 0.

The parameter estimates were insensitive to small changes in the initial values taken for the algorithm, but using randomly generated initial values two other sets of estimates were found that gave the same residual sum of squares. However, these did not correspond to physically reasonable solutions:

for example, k_3 greater than k_7 yet k_4 less than k_8 is most unlikely given the chiral similarity of the molecules X and Y to the molecules R and S.

Model F

This is the full model which allows for epimerization between X and Y and between R and S. The constraints are that $[X]_0 > [Y]_0$, $k_1 \leq k_5 + 0.005$ and $k_2 \leq k_6 + 0.005$. The restrictions on the rate constants arise from a physical argument that the epimerization (which involves movement of a hydrogen atom) is impeded in the bound state.

The results depend on the starting values for the unknown parameters and are given in Table 1. The residual sum of squares for alternatives (i) and (ii) are 99.38 and 103.4 respectively. Even for the case of alternative (i), the reduction in the residual sum of squares relative to Model 0 is less than is expected with the loss of 4 degrees of freedom, if the additional parameters are zero (F=0.30 on 4 and 36 degrees of freedom). Furthermore, despite the slightly smaller residual sum of squares, alternative (i) is considered physically unrealistic because S is known to degrade directly.

Model C

This model allows for epimerization but requires specified values for the epimerization rate constants. For Model C(1) k_1 and k_2 were set at 0.005 and k_5 and k_6 were set at 0.01. For Model C(2), k_1 and k_2 were 0.02 and k_5 and k_6 were 0.025. These values are somewhat arbitrary but ensure at least a small amount of epimerization in the fitted model.

The estimation problem is therefore similar to Model 0.

The residual sum of squares for Model C(1) is 103.9. This is directly comparable with the residual sum of squares for Model 0, and is slighly higher. If more epimerization is assumed, Model C(2) for example, the residual sum of squares increases. For Model C(2) the residual sum of squares is 106.6. The estimate of $[Y]_0$ is quite sensitive to the amount of epimerization assumed, because epimerization provides an alternative pathway for the production of S.

2.4 Spline fits

The strategy was to use a spline fit to obtain estimates of first and second derivatives of $[R]$ and $[S]$, and hence to estimate the rate parameters directly by least squares without solving the differential equations. This proposed method was first tried on a rather simpler system consisting of a state space model for a mass on a spring with natural frequency 1 and damping parameter

Table 1. Parameter estimates with their accuracy and precision

	k_1	k_2	k_3	k_4	k_5
Model 0 estimates			.409	.088	
Bootstrap mean			.413	.090	
Bootstrap std dev.			.084	.040	
Asymptotic std dev.			.053	.036	
Model F					
estimates(i)	.019	.000	.036	.032	.066
estimates(ii)	.000	.000	0.400	.100	.005
Model C(1) estimates	(.005)	(.005)	.380	.110	(.01)
Bootstrap mean			.38	.11	
Bootstrap std dev			.079	.064	
Model C(2) estimates	(.02)	(.02)	.331	.170	(.025)
	k_6	k_7	k_8	$[X]_0$	$[Y]_0$
Model 0 estimates		.033	.028	29.9	16.7
Bootstrap mean		.034	.032	30.3	19.4
Bootstrap std dev.		.0036	.012	2.04	5.82
Asymptotic std dev.		.0023	.0105	1.35	4.65
Model F					
estimates(i)	.350	.430	.000	370	19.3
estimates(ii)	.005	.032	.026	30.5	14.4
Model C(1) estimates	(.01)	.031	.026	31.3	12.8
Bootstrap mean		.032	.030	31.9	15.6
Bootstrap std dev		.004	.012	2.35	6.35
Model C(2) estimates	(.025)	.027	.027	34.5	6.74

0.2. The method gave precise estimates of parameters from 250 noise free observations at 0.1 second intervals, but deteriorated rapidly if noise was added. In a simulation, the observations were multiplied by $(1 + 0.01 \times Z_i)$ where Z_i is a random number from a standard normal distribution. A centred moving average of order 7 was then applied to the noisy observations before fitting splines. The means and standard deviations shown in brackets, of 100 such estimates of the natural frequency and damping were -0.9186 (0.2519) and -0.4233 (0.0792) respectively. Fitting methods using the discrete time series were slightly better, but, so far, the method of Section 2.3 appears far superior. However, use of a Kalman filter with the discrete time series is likely to lead to more precise estimates and we will investigate this in future work.

3 Model Validation

There is no statistical evidence of epimerization and the simplest model, Model 0, was by a small margin the best fit. We have therefore chosen to present an analysis of the residuals and the results of a cross-validation for Model 0. The residuals for R and S are defined by:

$$r_{Rt} = [R]_t - [\hat{R}]_t,$$
$$r_{St} = [S]_t - [\hat{S}]_t,$$

where the circumflex accent denotes fitted value. An R^2 statistic was defined as

$$R^2 = 1 - \frac{\Sigma(r_{R_t}^2 + r_{S_t}^2)}{\Sigma\left\{([R]_t - [\bar{R}])^2 + ([S]_t - [\bar{S}])^2\right\}}$$

and equals 0.930. If this is adjusted by dividing by the appropriate degrees of freedom it becomes 0.913. It would be higher if the overall mean, $([\bar{R}] + [\bar{S}])$ was used in the definitions. There is clear evidence that the residuals for $[R]$ have a higher standard deviation (1.96) than those for $[S]$ (0.90). However, transforming the residuals by dividing by the square root of either $[\hat{R}]_t$ or $[\hat{S}]_t$, as appropriate, results in transformed residuals which have an approximately constant standard deviation, although there are outlying observations at 1, 2 and 24 hours, (Figures 2, 3) and an approximately normal distribution (Figures 4-6).

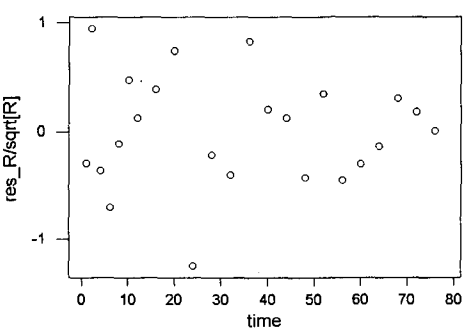


Figure 2. Transformed R residuals against time.

The sensitivity of the parameter estimates to the outlying observations, and all the other observations, was ascertained by repeating the analysis with each one of the 23 data triples omitted. These estimates are shown in Figures 7 and 8 and are reasonably stable. The PRESS residuals themselves are shown in Figure 9.

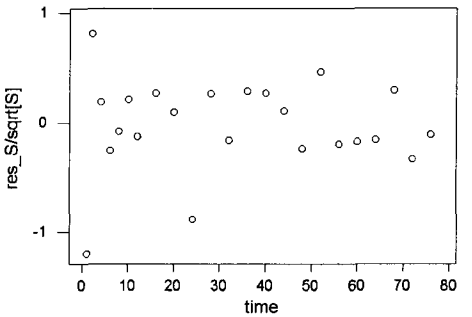


Figure 3. Transformed S residuals against time.

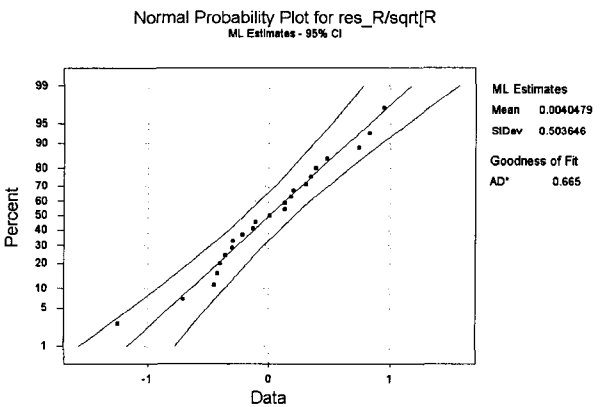


Figure 4. Normal plot for transformed R residuals.

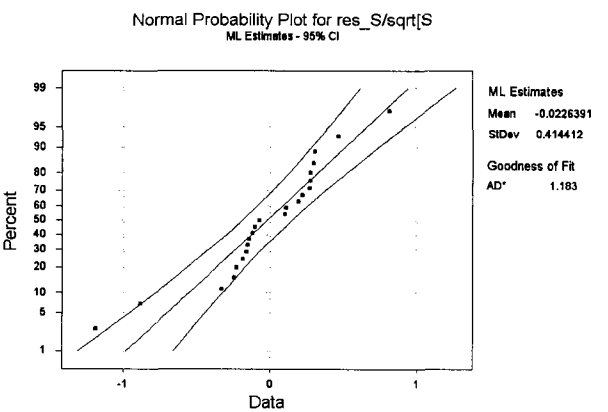


Figure 5. Normal plot for transformed S residuals.

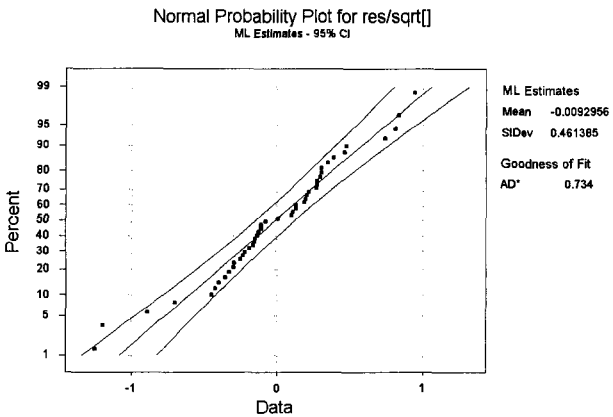


Figure 6. Normal plot for transformed residuals.

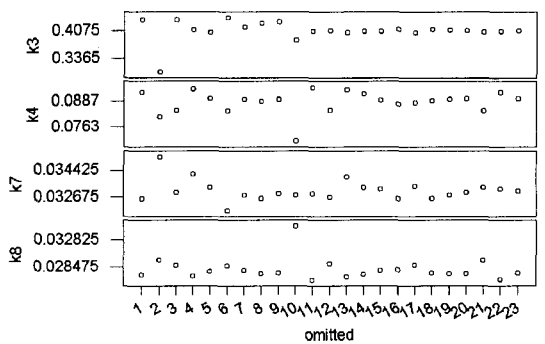


Figure 7. Sensitivity of rate constants to omitted data.

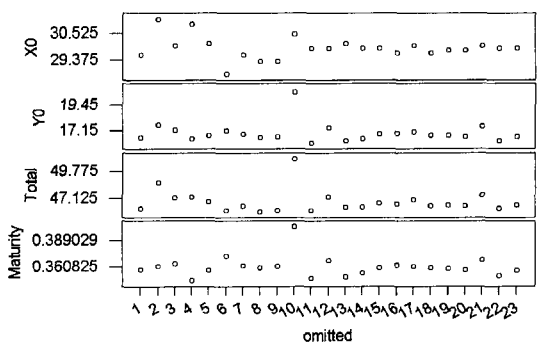


Figure 8. Sensitivity of concentration parameters to omitted data

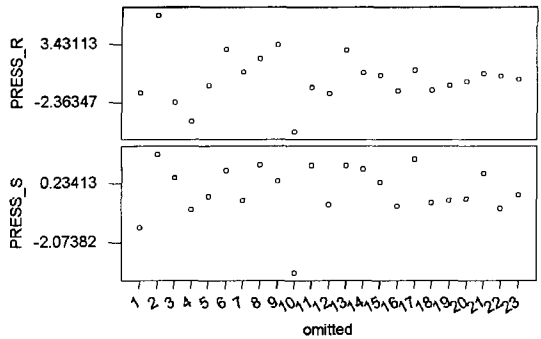


Figure 9. PRESS residuals

4 Discussion and Conclusions

In principle it is possible to make accurate estimates of all the rate constants and the initial amounts of bound biomarkers from the time series of concentrations of $[R]$ and $[S]$. However, in practice, the observations are not free of error and the number of sampling times is restricted (23 in this study). The data used in this study were the result of careful experimental work over several months. Nevertheless, they are not sufficient to make precise estimates of the unknown parameters because the least squares function of all 10 parameters has no clearly defined minima at physically realistic points and appears to have least values at boundaries.

If values are assumed for the epimerization rate constants (k_1, k_2, k_5 and k_6) the remaining parameters can be estimated with reasonable precision. However, a statistical test of any plausible hypothesis for the epimerization rate constants is unlikely to lead to rejection of the hypothesis at any useful level. In particular, there is no evidence against a hypothesis of no epimerization, and this simplest model, Model 0, was by a small margin the best fit.

The estimate of $[Y]_0$ and hence the estimate of the maturity parameter $[Y]_0/([X]_0 + [Y]_0)$ is quite sensitive to the assumed amount of epimerization, at least with the constraints $k_1 = k_2$, $k_5 = k_6$ and $k_1 + 0.005 < k_5$, although the estimate of the total amount $[X]_0 + [Y]_0$ is not. The case of Model C(1) is compared with Model 0 in Table 2 below.

If it is reasonable to assume the epimerization rate constants are small, for example when pyrolysis is at temperatures as low as $250°C$, Model 0 may be a reasonable approximation. Even if such an assumption is not justifiable, analysis with an assumption of no epimerization could be used to compare

Table 2. Bootstrap statistics for key parameter estimates

Model	mean sum	std dev (sum)	mean maturity	std dev (maturity)
0	49.75	6.04	0.380	0.074
C	47.48	6.57	0.320	0.086

different oil bearing rocks and to provide an estimate of an upper bound on the maturity parameter. However, it would be advisable to make independent estimates of the epimerization rate constants from either ab initio molecular orbital calculations or ancilliary laboratory experiments. This additional information could be combined with the method of this paper, for Model C, by a Bayesian analysis.

REFERENCES

Abbott, G.D., Wang, G.Y., Eglinton, T.I., Home, A.K. and Petch, G.S. (1990). The kinetics of sterane biological marker release and degradation processes during the hydrous pyrolysis of vitrinite kerogen. *Geochimica et Cosmochimica Acta* 54, 2451-2461.

Davison, E.J. (1977). Connectability and structural controllability of composite systems. *Automatica* 13, 109-123.

Eisenfeld, J. (1985). Remarks on Bellman's structural identifiability. *Mathematical Biosciences* 77, 229-243.

Lewan, M.D., Bjorøy, M. and Dolcater, D.L. (1986). Effects of thermal maturation on steroid hydrocarbons as determined by hydrous pyrolysis of Phosphoria Retort Shale. *Geochimica et Cosmochimica Acta* 50, 1977-1987.

Luisa Marin, M., Zaragoza, R.J. and Pérez Pla, F.F. (1999). Analyzing the uniqueness of the rate constants calculated from complex kinetic systems: a study of the hydrolysis of ciclohexanecarbonitribes. *International Journal of Chemical Kinetics* 31, 611-626.

Mackenzie, A.S. (1984). *Applications of biological markers in petroleum geochemistry.* In: Brooks, J. and Welte, D.H., (eds). Advances in Petroleum Geochemistry 1, 115-214.

Mackenzie, A.S. and McKenzie, D. (1983). Isomerization and aromatization of hydrocarbons in sedimentary basins formed by extension. *Geological Magazine* 120, 417-470.

Philp, R.P. (1985). Biological markers in fossil fuel production. *Mass Spectrometry Reviews* 4, 1-54.

Vajda, S. and Rabitz, H. (1988). Identifiability and distinguishability of first-order reaction systems. *Journal of Physical Chemistry* 92, 701-707.

NON-REPLICABILITY OF DISEASE GENE RESULTS: A MODELLING PERSPECTIVE

SUSAN R. WILSON
Centre for Mathematics and its Applications
The Mathematical Sciences Institute, and
Centre for Bioinformation Science
John Curtin School of Medical Research
Australian National University, Canberra ACT 0200, Australia
E-mail: Sue.Wilson@anu.edu.au

GAVIN A. HUTTLEY
Centre for Bioinformation Science
The Mathematical Sciences Institute and
John Curtin School of Medical Research
Australian National University, Canberra ACT 0200, Australia
E-mail: Gavin.Huttley@anu.edu.au

The prevalent paradigm for the analysis of common human disease assumes that a single gene is largely responsible for individual disease risk. The consequence of examining each gene as though it were solely responsible for conferring disease risk when in fact that risk is contingent upon interactions with other disease loci has not been considered. So, a general genetic model to analyse data for two marker loci has been developed. Based on this model it is shown that results can vary markedly depending on the parameters associated with the "unidentified" disease gene. In particular it is found that if parameters associated with the second gene were to vary between studies, then the conclusions from those studies may also vary. This is a theoretically broad conclusion with important implications for interpreting different results from individual genome studies and comparing results between studies.

Keywords: association studies, linkage, log-linear modelling, Simpson's paradox.

1 Introduction

With the completion of the human genome sequence, a major obstacle to identifying the genetic basis of human phenotypes, the positional cloning of genes, has been largely overcome. For each phenotype, a critical step that remains is to identify a list of possible genes as candidates. Regardless of how this list is determined, the next step in establishing causation is then to identify all the gene variants and show that their incidence correlates with phenotypic differ-

ences between individuals. Generally this is being done by considering each locus separately. Unfortunately, there is a disturbing trend in the outcome of such studies, specifically a difficulty in replicating positive results. There are many possible explanations for inconsistency between studies. Beyond the obvious interpretation that replication failure may originate in type I errors, is a prospect that these underlying models are too simplistic. If the aetiology of a common disease is due to interactions among two or more genes then the mode in which those interactions take place, such as additively or multiplicatively, may impact on our ability to map the responsible genes. Two of the most popular study designs, the transmission disequilibrium test (TDT) and the case control design have comparable statistical power to identify genes with variants that contribute a low relative risk for common disorders (Risch, 2000). The TDT uses triad data - consisting of an affected individual and their parents. For this test the untransmitted alleles serve as the control. This study design is widely considered robust to the effects of population stratification compared with case-control studies. Here, we consider the implications of ignoring epistatic interactions (interactions among genes) on the outcome of the TDT, and develop methods for examining such effects using the TDT.

2 Modelling

Consider affected (A) individuals whose parental marker haplotypes are i_1i_2, j_1j_2 for the male parent (MP) and k_1k_2, l_1l_2 for the female parent (FP) where subscript 1 refers to alleles at the first locus, and 2 at the second. Let $P_{wx|yz}^{i_1j_1i_2j_2|k_1l_1k_2l_2}$ be the conditional probability that the MP transmits haplotype wx and the FP transmits haplotype yz given that the MP is $i_1j_1i_2j_2$ and the FP is $k_1l_1k_2l_2$ and the child is affected, where w is one of the pair (i_1,j_1), x is one of (i_2,j_2), y is one of (k_1,l_1), and z is one of (k_2,l_2). Suppose there are two disease susceptibility ($\mathcal{DS}$) loci which lie between two marker loci. Let θ_1^m, θ_2^m and θ_3^m be the recombination fractions between the first marker and the first $\mathcal{DS}$ locus, between the two $\mathcal{DS}$ loci, and between the second $\mathcal{DS}$ locus and the second marker respectively for the MP, and with f replacing m for the FP. Then it can be shown, using Bayes' theorem as well as the assumptions that there is no selection and/or mutation in the production of gametes and there is random mating, that

$$P_{wx|yz}^{i_1j_1i_2j_2|k_1l_1k_2l_2} = p_{i_1i_2}p_{j_1j_2}p_{k_1k_2}p_{l_1l_2}E_{wx|yz}$$

where $E_{wx|yz}$ = P(Child A | MP transmits wx, FP transmits yz)/ P(Child A), and $p_{i_1i_2}$ represents the frequency of marker haplotype i_1i_2. Let $q_{s_1s_2}$

represent the frequency of the disease locus haplotype s_1s_2, and e the relevant linkage disequilibrium parameters for the marker-disease haplotypes. Let $E_{wx|yz} = E'_{wx|yz}/E_*$, where $E_* = \sum_{wx|yz} E'_{wx|yz}$. Based on a general Mendelian disease model for the two disease loci, it can be shown that

$$E'_{wx|yz} = \sum_{\mathbf{s},\mathbf{t},\mathbf{u},\mathbf{v}} e_{i_1i_2s_1s_2} e_{j_1j_2t_1t_2} e_{k_1k_2u_1u_2} e_{l_1l_2v_1v_2} q_{s_1s_2} q_{t_1t_2} q_{u_1u_2} q_{v_1v_2} \tau_{wx|yz}$$

where $\mathbf{s}$, $\mathbf{t}$, $\mathbf{u}$, $\mathbf{v}$ represent the disease locus haplotypes s_1s_2, t_1t_2, u_1u_2, and v_1v_2 respectively, $\tau_{wx|yz} = \theta_{wx}^{m^T} F \theta_{yz}^{f}$, where superscript T represents (here the vector) transpose, and

$$F = \begin{bmatrix} f_{s_1u_1s_2u_2} & f_{s_1v_1s_2u_2} & f_{s_1v_1s_2v_2} & f_{s_1u_1s_2v_2} \\ f_{t_1u_1s_2u_2} & f_{t_1v_1s_2u_2} & f_{t_1v_1s_2v_2} & f_{t_1u_1s_2v_2} \\ f_{t_1u_1t_2u_2} & f_{t_1v_1t_2u_2} & f_{t_1v_1t_2v_2} & f_{t_1u_1t_2v_2} \\ f_{s_1u_1t_2u_2} & f_{s_1v_1t_2u_2} & f_{s_1v_1t_2v_2} & f_{s_1u_1t_2v_2} \end{bmatrix}$$

where, for example, $f_{s_1u_1s_2u_2}$ is the penetrance of the disease locus genotype $s_1u_1s_2u_2$, that is Pr(Affected $| s_1u_1s_2u_2$), and analogously for the other terms, and $\theta_{i_1i_2}^{m^T}$, $\theta_{i_1j_2}^{m^T}$, $\theta_{j_1i_2}^{m^T}$, and $\theta_{j_1j_2}^{m^T}$ are respectively given by the following row vectors

$$((1-\theta_1^m)(1-\theta_2^m)(1-\theta_3^m), \theta_1^m\theta_2^m(1-\theta_3^m), \theta_1^m(1-\theta_2^m)\theta_3^m, (1-\theta_1^m)\theta_2^m\theta_3^m)$$

$$((1-\theta_1^m)(1-\theta_2^m)\theta_3^m, \theta_1^m\theta_2^m\theta_3^m, \theta_1^m(1-\theta_2^m)(1-\theta_3^m), (1-\theta_1^m)\theta_2^m(1-\theta_3^m))$$

$$(\theta_1^m(1-\theta_2^m)(1-\theta_3^m), (1-\theta_1^m)\theta_2^m(1-\theta_3^m), (1-\theta_1^m)(1-\theta_2^m)\theta_3^m, \theta_1^m\theta_2^m\theta_3^m)$$

$$(\theta_1^m(1-\theta_2^m)\theta_3^m, (1-\theta_1^m)\theta_2^m\theta_3^m, (1-\theta_1^m)(1-\theta_2^m)(1-\theta_3^m), \theta_1^m\theta_2^m(1-\theta_3^m))$$

The corresponding values for $\theta_{yz}^{f^T}$ are obtained by replacing m by f, and s_1 by u_1, s_2 by u_2, t_1 by v_1, and t_2 by v_2, in the above vectors. It is noted that the simplified formulae for a single $\mathcal{DS}$ locus can be obtained by ignoring all terms with the subscript 2 in the above, as well as θ_3, x and z, and reducing F to its appropriate 2×2 form. First, suppose each marker is very close to a $\mathcal{DS}$ locus, so it can be assumed that $\theta_1^m = \theta_1^f = \theta_3^m = \theta_3^f = 0$. It can be shown that

$$\frac{P_{i_1i_2|k_1k_2}^{i_1j_1i_2j_2|k_1l_1k_2l_2}}{P_{j_1j_2|l_1l_2}^{i_1j_1i_2j_2|k_1l_1k_2l_2}} = \frac{\sum_{\mathbf{s},\mathbf{u}} e_{i_1i_2s_1s_2} e_{k_1k_2u_1u_2} q_{s_1s_2} q_{u_1u_2} f_{s_1u_1s_2u_2}}{\sum_{\mathbf{t},\mathbf{v}} e_{j_1j_2t_1t_2} e_{l_1l_2v_1v_2} q_{t_1t_2} q_{v_1v_2} f_{t_1v_1t_2v_2}} \qquad (1)$$

(noting that the θ_2 terms cancel)

$$= \frac{\Pr(\mathrm{A} \mid \text{MP transmits } i_1 i_2 \text{ and FP transmits } k_1 k_2)}{\Pr(\mathrm{A} \mid \text{MP transmits } j_1 j_2 \text{ and FP transmits } l_1 l_2)},$$

$$= \frac{\Pr(\mathrm{A} \mid \text{MP transmits } i_1 i_2)\Pr(\mathrm{A} \mid \text{FP transmits } k_1 k_2)}{\Pr(\mathrm{A} \mid \text{MP transmits } j_1 j_2)\Pr(\mathrm{A} \mid \text{FP transmits } l_1 l_2)}$$

if $f_{s_1 s_2 u_1 u_2} = f_{s_1 s_2} f_{u_1 u_2}$, where

$$\frac{Pr(\mathrm{A} \mid \text{MP transmits } i_1 i_2)}{Pr(\mathrm{A} \mid \text{MP transmits } j_1 j_2)} = \frac{\sum_{\mathbf{s}} e_{i_1 s_1 i_2 s_2} q_{s_1 s_2} f_{s_1 s_2}}{\sum_{\mathbf{t}} e_{j_1 t_1 j_2 t_2} q_{t_1 t_2} f_{t_1 t_2}}$$

and analogously for FP.

The equality $f_{s_1 s_2 u_1 u_2} = f_{s_1 s_2} f_{u_1 u_2}$ above will occur for all s_1, s_2, u_1, u_2 if the penetrance model is fully multiplicative; namely that the penetrances obey not only the relation $(f_{a_i a_j})^2 = f_{a_i a_i} f_{a_j a_j} (a_i \neq a_j)$ for each $\mathcal{DS}$ locus considered separately, but the penetrances are also multiplicative across the loci. The result holds whether the $\mathcal{DS}$ loci are linked or unlinked. So before proceeding, it is useful to give a simplified matrix of penetrance values for a multiplicative model for two $\mathcal{DS}$ loci each having two alleles:

$$\begin{bmatrix} f & r_2 f & r_2^2 f \\ r_1 f & r_1 r_2 f & r_1 r_2^2 f \\ r_1^2 f & r_1^2 r_2 f & r_1^2 r_2^2 f \end{bmatrix}$$

The rows represent the three genotypes at the first locus, the columns the second locus, and f is the background penetrance rate with r_1 the relative risk of the disease for individuals having one copy of *allele* 2_1 at the first locus, and r_2 being the corresponding relative risk associated with having one copy of *allele* 2_2 at the second locus.

Now consider the general two allele, two $\mathcal{DS}$ loci case when parental origin of the alleles is ignored. The penetrance model will involve up to eight parameters, compared with only two parameters for the single $\mathcal{DS}$ locus two allele case. Let 1_1 *allele* and 2_1 *allele* denote the two alleles at the first $\mathcal{DS}$ locus, and 1_2 *allele* and 2_2 *allele* for the second $\mathcal{DS}$ locus. The general matrix of penetrance values can be represented as

$$\begin{bmatrix} f & f e^{\beta_{0,1}} & f e^{\beta_{0,2}} \\ f e^{\beta_{1,0}} & f e^{\beta_{1,1}} & f e^{\beta_{1,2}} \\ f e^{\beta_{2,0}} & f e^{\beta_{2,1}} & f e^{\beta_{2,2}} \end{bmatrix}$$

where f represents an arbitrarily chosen base-line risk, and $e^{\beta_{a_1,a_2}}$ represents the relative risk of being affected from having a_1 2_1 *alleles* at the first $\mathcal{DS}$ locus (where $a_1 = 0, 1, 2$), and a_2 2_2 *alleles* at the second (where $a_2 = 0, 1, 2$), compared with the risk of having no type *2 alleles* at either locus.

3 Analyzing triad data for epistasis using two marker loci

Consider marker allelic transmission data from both parents to an affected child, and suppose the marker loci are unlinked. From application of full likelihood methods, the log-linear approach to analysing such data, based on the above extended genetic model, is obtained. This section concentrates on two loci each having two alleles, otherwise the data could become very sparse indeed. However, the methodology is general, and so the results can be extended to more than two alleles at the loci, as well as to more than two loci. It is convenient to arrange the observations $Y_{M_1F_1C_1,M_2F_2C_2}$ of the parental transmissions to affected children, where M_1, F_1, C_1 represents the number of copies of the *allele 2*$_1$ carried by MP, FP and the affected child, respectively, at the first locus, and analogously for the second locus, as a two-way table with the rows representing the informative mating types, and the columns the affected offspring genotypes. For two loci each with two alleles, this matrix will have 35 rows of informative matings (i.e. matings for which more than a single genotype can occur for their (affected) children) and nine columns (with many zero entries for unobservable $M_1F_1C_1, M_2F_2C_2$ combinations); see Wilson (2001). Let $\mu_{M_1F_1C_1,M_2F_2C_2} = E(Y_{M_1F_1C_1,M_2F_2C_2})$. Then in the log-linear model formulation we can write $\ln \mu_{M_1F_1C_1,M_2F_2C_2}$ as

$$\begin{aligned}\phi_m + \beta_{c_1,c_2} I_{(C_1=c_1,C_2=c_2)} + \ln(2) I_{(M_1F_1C_1,M_2F_2C_2)=(x'y'z',111)} \\ + \ln(2) I_{(M_1F_1C_1,M_2F_2C_2)=(111,xyz)} + \ln(4) I_{(111,111)}\end{aligned} \tag{2}$$

where $\beta_{0,0} = 0$, and neither xyz nor $x'y'z'$ are equal to 111. Here ϕ_m are stratum parameters that constrain the model fit to be such that for each mating type m, the fitted number equals the observed number of matings of that type having affected children. Also, $\beta_{c_1,c_2} = lnR_{c_1,c_2}$ is the logarithm of the relative risk (relative penetrance parameter) associated with $C_1 = c_1, C_2 = c_2$ relative to $C_1 = 0, C_2 = 0$. As well, I is an indicator function taking the value 1 when the subscripted relationship holds and 0 otherwise. The constants ln(2) and ln(4) are called "offsets". These are required because of the Mendelian assumption that two parents who are heterozygous at the same locus are twice as likely to produce a heterozygous child at this locus as to produce either one with no copies of *allele 2* or two copies of this allele. Two doubly heterozygous parents are four times as likely to produce a doubly heterozygous child compared with producing one of the four doubly homozygous children, and it follows that they are twice as likely to produce a singly heterozygous child, i.e. a child who is heterozygous at one locus and homozygous at the other, compared with producing a doubly homozygous child.

Simplifications of the general model (2) and differences in the relevant deviances can be readily evaluated to determine which model gives an adequate representation of the available data (Wilson, 2001).

Suppose that there is complete linkage of each marker to one of the disease loci, including complete allelic association. In other words, in the table of haplotype frequencies written with the rows representing the marker loci and the columns the disease loci, the offdiagonal terms $h_{i_1 i_2 s_1 s_2}$ are zero. In this case the right-hand side of (1) is $\frac{f_{i_1 k_1 i_2 k_2}}{f_{j_1 l_1 j_2 l_2}}$. Next suppose data are only available for the first locus, yet the underlying genetic model for the complex disease involves two possibly epistatic loci, as well as many background causes and associations, each of individually small effect. Then the penetrances at this first locus for the three genotypes, that is for having zero, one or two copies of *allele* 2_1, would be $fq_2^2 + fe^{\beta_{0,1}} 2p_2q_2 + fe^{\beta_{0,2}} p_2^2$, $fe^{\beta_{1,0}} q_2^2 + fe^{\beta_{1,1}} 2p_2q_2 + fe^{\beta_{1,2}} p_2^2$, and $fe^{\beta_{2,0}} q_2^2 + fe^{\beta_{2,1}} 2p_2q_2 + fe^{\beta_{2,2}} p_2^2$ respectively, where p_2 is the frequency of *allele* 2_2 at the second locus, and $q_2 = 1 - p_2$. For example, for the multiplicative disease model, we have $e^{\beta_{0,1}} = r_2$, $e^{\beta_{0,2}} = r_2^2$, $e^{\beta_{1,0}} = r_1$, $e^{\beta_{1,1}} = r_1 r_2$, $e^{\beta_{1,2}} = r_1 r_2^2$, $e^{\beta_{2,0}} = r_1^2$, $e^{\beta_{2,1}} = r_1^2 r_2$ and $e^{\beta_{2,2}} = r_1^2 r_2^2$. Hence the baseline rate for those having no *allele* 2_1 at the first locus becomes $f(q_2 + p_2 r_2)^2$, and the relative risk associated with having *allele* 2_1 is r_1.

Let $P_{J_1}^{1_1 2_1}$, $J = 1, 2$ represent the conditional probability that at the first locus the parent transmits J_1 given that the parent is $1_1 2_1$ and the child is affected. Then tests like the TDT evaluate departures of the ratio $P_{1_1}^{1_1 2_1} / P_{2_1}^{1_1 2_1}$ from 1.0. It can be shown that here this ratio is

$$\frac{p_1(q_2^2 + e^{\beta_{0,1}} 2p_2q_2 + e^{\beta_{0,2}} p_2^2) + q_1(e^{\beta_{1,0}} q_2^2 + e^{\beta_{1,1}} 2p_2q_2 + e^{\beta_{1,2}} p_2^2)}{p_1(e^{\beta_{1,0}} q_2^2 + e^{\beta_{1,1}} 2p_2q_2 + e^{\beta_{1,2}} p_2^2) + q_1(e^{\beta_{2,0}} q_2^2 + e^{\beta_{2,1}} 2p_2q_2 + e^{\beta_{2,2}} p_2^2)}.$$

For the multiplicative model considered above, the ratio is r_1^{-1}.

4 Influence of a second epistatic locus on a single locus analysis

To simplify the presentation, assume again that the underlying disease mechanism involves two major unlinked disease loci, and that there is complete linkage by each marker to one of the disease loci. Further, a specific example is given, but the possibility of many generalisations to other situations is noted. Suppose that the first locus appears, on analysis say, to be common dominant. Then the penetrances could be represented by f, fe^W, fe^W dependent on whether there are zero, one or two copies of *allele* 2_1, respectively, and $W \geq 1$. Suppose a second locus is acting epistatically, and that the matrix of penetrance values is of the form

Table 1. A simple model parameterisation for disease susceptibility; f represents baseline risk

Genotype		bb	Bb	BB
	frequency	q_2^2	$2p_2q_2$	p_2^2
aa	q_1^2	f	αf	αf
Aa	$2p_1q_1$	βf	f	f
AA	p_1^2	βf	f	f

$$\begin{bmatrix} f & fe^X & fe^X \\ fe^W & fe^{W-Y} & fe^{W-Y} \\ fe^W & fe^{W-Y} & fe^{W-Y} \end{bmatrix}$$

where f represents an arbitrarily chosen base-line risk. This is a simplified version of the general table of penetrance values given above. Let fe^Z, fe^V, and fe^V represent the marginal penetrances associated with having zero, one or two copies of *1_2 alleles* at the first locus. Then it follows that $e^X = (e^Z - q_2^2)/(1 - q_2^2)$. Compared with the completely general parameterisation, it is easier to see in this case that different relative values of the epistatic parameters X and Y due to the second locus, as well as the second locus's allele fequencies p_2, q_2, could change the marginal appearance of the first locus from being common dominant to having no effect, or to being common recessive. Further, the ratio $P_{1_1}^{1_1 2_1}/P_{2_1}^{1_1 2_1}$ above simplifies to $1 + p_1(e^{Z-V} - 1)$. So, if $e^{Z-V} > 1$, *allele 1_1* will appear to be preferentially transmitted at the first locus, while on the other hand if $e^{Z-V} < 1$, *allele 1_2* will appear to be preferentially transmitted, and if $e^{Z-V} \approx 1$, the first locus will appear to have no effect. This is further illustrated in Figure 1. Under the simple disease model presented in Table 1, the plane represents the values of q_2 at which there is no transmission bias of a, and for a specific pair of penetrance parameter values (α and β), if the value of q_2 falls above (or below) the corresponding point on the plane, allele a (or allele A) will be perceived as increasing disease risk.

5 Discussion

Simpson's paradox, the reversal of the direction of the association by collapsing tables, is well known in applied statsitics. As we have shown here and elsewhere (Wilson, 2001) the effect of collapsing over other major loci when just a single locus is being investigated can be considerable. If the loci affect

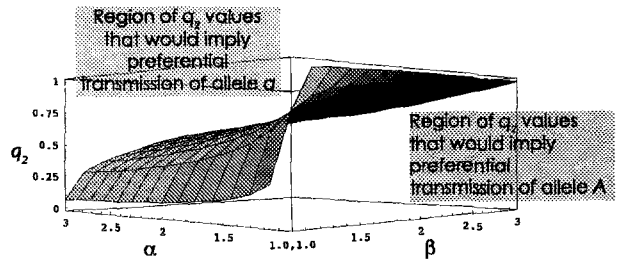


Figure 1. For a given α and β, either preferred transmission of a or A will be inferred, dependent on whether q_2 falls above or below the value on the plane, for parameterisation of Table 1.

disease outcome independently then a single locus approach can be used for finding the major genes that have an influence on disease development. If the loci do not act independently, however, then ignoring the other loci might have a major impact on one's result. In the above it was shown how this can occur for a quite simple background scenario, namely that of two loci each with two alleles that are the actual susceptibility alleles. It is worth emphasizing that a role for epistatic interactions in disease phenotypes is a real issue, not just a theoretical possibility. Examples of phenotypes where the causative loci act epistatically are known in variety of model organisms, including mice and humans.

Despite the perception that tests of association, such as TDT, overcome the well-known problems associated with the marker gene frequency varying from population to population, the above indicates this confidence can be misplaced. When there is epistasis between the genes affecting a phenotype, conclusions from family based tests of association such as TDT can also appear to vary from study to study if allele frequencies at the other loci that are not being simultaneously evaluated also vary between samples.

6 Conclusion

The thrust of the theoretical results presented here is to suggest that neither variations in empirical results from study to study, nor failure to demonstrate an association between a candidate gene locus and a phenotype in a single study, be necessarily judged as showing that the particular single locus is not a disease locus. Rather, consideration must now be given as to whether there may be one or more other major loci acting epistatically with this locus.

REFERENCES

Risch, N.J. (2000). Searching for genetic determinants in the new millennium. *Nature* 405, 847-56.

Wilson, S.R. (2001). Epistasis and its possible effects on transmission disequilibrium tests. *Ann. of Hum. Genet.* 65, 565-575.

STATISTICAL PROCESS MONITORING FOR AUTOCORRELATED DATA

NIEN FAN ZHANG
Statistical Engineering Division
National Institute of Standards and Technology
Gaithersburg MD 20899, USA
E-mail: zhang@nist.gov

In the past years statistical process control methodologies have been widely used in industry for process monitoring. However, the typical assumption that the process data are statistically independent is often invalid. This article discusses different approaches for dealing with process autocorrelation when using process control charts and process capability indices.

Keywords: moving average, process capability index, process control chart, stationary process, time series.

1 Introduction

In the past years, statistical process control (SPC) methodologies such as process control charts have been widely used in industry for process mean and variability monitoring. Most of SPC charts such as the Shewhart chart, the cumulative sum (CUSUM) chart and the exponentially weighted moving average (EWMA) chart were established based on the assumption that the data are statistically independent. However, this assumption is often invalid for many manufacturing processes such as those in continuous processes. In industry with continuous processes, on-line data acquisition systems are commonly used, and a huge volume of data on process variables is collected. Due to the high frequency sampling, most of the on-line data are autocorrelated. When the data are autocorrelated, although the traditional control charts still can be used, their use lacks of a solid scientific rationale for ascertaining whether the proess is in a state of statistical control and they are often not effective. Several papers discussed impact of autocorrelation on the performance of control charts. Johnson and Bagshaw (1974) and Bagshaw and Johnson (1975) found that the autocorrelation affects the performance of CUSUM chart. Harris and Ross (1991) discussed the impact of autocorrelation on CUSUM and EWMA charts.

To illustrate the impact of the autocorrelation on the performance of the

traditional control charts, a simulation study was done in Zhang (2000) for a (weakly) stationary first order autoregressive (AR(1)) process. An AR(1) process $\{X_t\}$ is defined by

$$X_t - \mu = \phi(X_{t-1} - \mu) + e_t \tag{1}$$

where $| \phi | < 1$ and μ is the process mean and e_t is white noise. Here we assume that e_t is also normally distributed with finite variance σ_e^2. Without loss of generality we assume that the process mean, μ, is zero. When ϕ=0, $\{X_t\}$ is an i.i.d. sequence. For each of the AR(1) processes with ϕ=0.25, 0.5, 0.75 and 0.9, at a certain point, a mean shift occurred with a magnitude of 0.5, 1, 2, and 3 in the unit of the process standard deviation, i.e., $\sigma_x^2 = \sigma_e^2/(1-\phi^2)$.

The average run length (ARL), the mean of the run length, is used to measure the performance of control charts. A desired control chart should have large in-control ARL and small out-of-control ARL. That is, for a desired chart when the process has no mean shift the ARL should be large, and when a mean shift occurs the ARL should be small to signal the mean shift quickly. For the AR(1) processes with ϕ specified as above, the ARL's were estimated for the X (individual) chart, CUSUM chart, and EWMA chart. For the X chart, 3σ control limits were used. For the CUSUM chart, the decision interval with h=5 and k=0.5 was used. For the EWMA chart, the parameter λ is 0.2 and 3σ control limits were used.

The simulation showed that when an AR(1) process is positively autocorrelated, even as weakly as with ϕ=0.25, the autocorrelation has a big impact on the CUSUM and EWMA charts and the in-control ARL's are smaller than those when ϕ=0. Namely, even when the process is weakly autocorrelated, the control charts will give frequent false alarms when the process is actually stable and under control. The details can be found in Zhang (2000). Autocorrelation also affects the use of process capability indices such as C_{pk} and the tolerance limits. For the effect of autocorrelation on tolerance limits, we refer it to Amin and Lee (1999).

In this article, we discuss various approaches for dealing with process autocorrelation. In Section 2, we discuss two approaches for handling autocorrelation in the use of control charts and present comparisons among various charts. In Section 3, we summarize other approaches for dealing with autocorrelation. In Section 4, we present a brief summary on the use of process capability indices when a process is autocorrelated.

2 Approaches for dealing with process autocorrelation

There have been two approaches for dealing with autocorrelation for control charts in the literature. We will discuss these approaches separately.

Residual charts

The first approach for dealing with autocorrelation is to use process residual charts proposed by Alwan and Roberts (1988). This approach requires one to model the process data and to obtain the process residuals. Assuming a true model, the residuals are statistically uncorrelated. Then the traditional SPC charts such as the X chart, CUSUM chart and EWMA charts can be applied to the residuals.

To illustrate the approach, consider an AR(1) process in (1). The residual at t is

$$e_t = X_t - [\hat{\mu} + \hat{\phi}(X_{t-1} - \hat{\mu})], \tag{2}$$

where $\hat{\mu}$ and $\hat{\phi}$ are the estimators of μ and ϕ. An X residual chart is a time-ordered plot of the residuals with $k\sigma$ limits. The center line is set to zero. Assuming a perfect model fitting, $\hat{\mu} = \mu$, $\hat{\phi} = \phi$ and the residuals are uncorrelated. For simplicity, the parameters and their estimators are not distinguished and denoted by μ and ϕ respectively, hereafter. Thus, the residual can be written as

$$e_t = X_t - [\mu + \phi(X_{t-1} - \mu)]. \tag{3}$$

The residual variance is the same as the white noise variance which is

$$\sigma_e^2 = Var(e_t) = (1 - \phi^2)\sigma_x^2. \tag{4}$$

Thus, the 3σ limits for the X residual chart is $3\sigma_e$. The CUSUM and EWMA residual charts are also constructed by applying the corresponding charts to the residuals.

Although the rationale for using residual charts is straightforward and it might be expected that the properties of the X residual chart are similar to those of the X chart, the X residual chart does not have the same properties as the X chart. Harris and Ross (1991) recognized that an X residual chart from an AR(1) process may have poor detection capability to signal a mean shift, for example when the process is positively autocorrelated. Longnecker and Ryan (1990) showed that for an AR(1) process with a positive ϕ, an X residual chart has a high probability of detecting a mean shift as soon as it occurs, but if the X residual chart fails to detect the shift immediately, then there is a low probability that the shift will be detected later.

Zhang (1997b) defined a measure of detection capability for the X residual chart for general stationary processes. Using this detection capability index, Zhang (1997b) compared the detection capability of the X residual chart with that of the X chart applied to an independent sequence. For general AR processes, Zhang (1997b) showed that the detection capability of the X residual chart is larger than that of the X chart immediately after the occurance of a mean shift. For stationary AR(2) processes, the relationship between the detection capability indices of the X residual chart and the X chart varies and depends on the process parameters. These results were extended to ARMA processes in Jiang and Wang (2001).

Jiang (2000) considered the detection capability index for EWMA residual charts. There have been no corresponding studies of the detection capability of the CUSUM residual charts. By simulation, Harris and Ross (1991) showed that the out-of-control ARL's for CUSUM residual charts are larger than those when the charts are applied to independent sequences . Runger, Willemnain, and Prabhu (1995) studied the ARL for CUSUM residual charts applied to AR(1) processes. Using a Markov chain approach, they provided an approximate expression for the ARL of a one-sided CUSUM of residuals from AR(P) models. A recent paper by Lu and Reynolds (2001) discussed the performance of the CUSUM charts for monitoring an autocorrelated process.

Lu and Reynolds (1999) discussed the performance of EWMA charts applied to the residuals of a process, which is an AR(1) process plus a random error. The model of the process is the same as the one Harris and Ross (1991) discussed, which is equivalent to an ARMA(1,1) process. Similar to Wardell, Moskowitz, and Plante (1994), they compared the EWMA chart and EWMA residual chart applied to the process mentioned in the above with various parameters. The conclusion was that when the process autocorrelation is weak or moderate, the two charts perform about the same, while the EWMA residual chart performs a little better than the EWMA chart.

Atienza, Tang, and Ang (1998) proposed a chart based on the statistics used for detecting outliers and level shifts in time series analysis. This approach requires time series modeling.

Control charts based on autocorrelation-adjusted limits

The second approach for dealing with autocorrelation, which is more direct, is to modify the existing SPC charts by adjusting the control limits to accommodate the autocorrelation. This approach has the advantage that it does not require time series modeling as the use of residual charts does. Vasilopoulos and Stamboulis (1978) proposed using the $\overline{X}$ chart with modified control limits to monitor autocorrelated data. Their studies, however,

were limited to some specific time series models such as AR(1) and AR(2) processes.

When a process model such as an AR(1) model is known, the control limits of the $\overline{X}$ chart can be calculated as a function of the model parameters such as ϕ for an AR(1) process and the process variance. For an $\overline{X}$ chart, we are charting the subgroup means of the process with subgroups of size n. When the process is an i.i.d. sequence, the 3σ limits are given by $\overline{\overline{X}} \pm 3\sigma_x/\sqrt{n}$, where $\overline{\overline{X}}$ is the grand mean. However, when the process is stationary and not white noise, the standard deviation of $\overline{X}$ is not $\sigma_x/\sqrt{n}$. The process autocorrelation has to be accounted for and the standard deviation of the sample mean is expressed in terms of process parameters and the process variance.

Modifications of CUSUM and EWMA charts have also been considered for autocorrelated data. Yashchin (1993) examined the use of CUSUM charts when the autocorrelation is moderate. The idea involves replacing the autocorrelated observations with an i.i.d. sequence for which the run length distribution is approximately the same. The article proposed a method to modify the control limits of the tabular form of the CUSUM chart by using the autocorrelation structure of the process.

Montgomery and Mastrangelo (1991) suggested monitoring a one-step ahead EWMA prediction error for autocorrelated data. The parameter of the EWMA was determined by minimizing the sum of the squares of the EWMA one-step-ahead prediction errors. The standard deviation of the one-step-ahead errors can be estimated using historical data. As shown by Box and Jenkins (1976, p.144-145), when $\{X_t\}$ is an integrated moving average (IMA) (0,1,1) process, the corresponding EWMA is the optimal one-step-ahead prediction of the process. This scheme is like a residual chart except that the IMA(0,1,1) model is assumed for all the processes and the prediction errors are used.

Schmid (1997) considered EWMA charts for time series. Zhang (1998a) proposed the EWMAST chart, which is an EWMA chart for stationary processes. The limits of the EWMA chart are adjusted to accommodate the autocorrelation and are determined by the process variance and autocorrelation. The EWMA of X_t is defined as

$$Z_t = (1 - \lambda)X_{t-1} + \lambda X_t, \tag{5}$$

where $z_0 = \mu$ and λ $(0 < \lambda \leq 1)$ is a constant. It is shown in Zhang (1998a)

that

$$\sigma_z^2 = Var[Z_t] = [\frac{\lambda}{2-\lambda}]\sigma_x^2\{1-(1-\lambda)^{2t}+2\sum_{k=1}^{t-1}\rho(k)(1-\lambda)^k[1-(1-\lambda)^{2(t-k)}]\} \tag{6}$$

where $\rho(k)$ is the autocorrelation function of X_t at lag k. When t is large, for a large integer M, an approximate variance of Z_t is given by

$$\sigma_z^2 \approx [\frac{\lambda}{2-\lambda}]\sigma_x^2\{1+2\sum_{k=1}^{M}\rho(k)(1-\lambda)^k[1-(1-\lambda)^{2(M-k)}]\}. \tag{7}$$

Assuming that X_t is normally distributed, the EWMAST chart is constructed by charting Z_t, which is also normally distributed. The centerline is at μ, and the $L\sigma$ limits are $\mu \pm L\sigma_z$.

Jiang, Tsui, and Woodall (2000) proposed to use ARMAST chart as an extension of the EWMAST chart. Essentially, for appropriate orders and parameters of an ARMA model, the chart is constructed by plotting the corresponding ARMA statistic of the process variable in the time order.

Comparisons among control charts for autocorrelated processes

In Zhang (1998a) and Zhang (2000), comparisons of ARL's were made among the EWMAST chart, X chart, X residual chart, CUSUM residual chart, EWMA residual chart, and the chart proposed by Montgomery and Mastrangelo (1991) (ϕ = 0.5, 0.75, and 0.95 for this chart) for AR(1) processes. It was concluded that when a process is weakly autocorrelated such as for an AR(1) process with ϕ=0.25, the EWMAST chart, EWMA residual chart, and the CUSUM residual chart perform equally well. When ϕ=0.5, the EWMAST chart performs better than the other charts. When ϕ=0.75, the EWMAST chart performs better than other charts for small to medium mean shifts. When ϕ=0.9, the X residual chart performs better than other charts for medium to large mean shifts while the EWMAST chart and the X chart perform better than other charts for small to medium mean shifts. The chart proposed by Montgomery and Mastrangelo (1991) performs very poorly when ϕ=0.5 and 0.75. Only when ϕ=0.95, does the chart performs relatively well. But even in this case, the X residual chart is slightly better than the chart proposed by Montgomery and Mastrangelo (1991).

The results for negative ϕ's show that for large mean shifts all charts perform well. For a small mean shift, the EWMAST, CUSUM residual, and EWMA residual charts perform best. The performance of the X chart is the worst for small to medium mean shifts. Overall, for negative ϕ's the EWMAST chart performs little better than all residual charts. An obvious

advantage of using the EWMAST chart is that there is no need to build a time series model for stationary process data.

The performance of the chart proposed by Atienza et al. (1998) is better than the residual charts and is comparable to the EWMAST chart for AR(1) processes. However, the approach requires time series modeling. The ARMAST chart proposed by Jiang, Tsui, and Woodall (2000) performs better than the EWMAST chart only when a process has strong and positive autocorrelations, i.e., when a process is near non-stationary with positive autocorrelation. A major disadvantage of using the ARMAST chart, however, is that the ad hoc approach proposed to choose the chart parameters would not be easy for many users.

3 Other approaches for dealing with autocorrelation

Another approach for dealing with autocorrelation is to reduce the process autocorrelation by some data treatment mechanism. Box and Jenkins (1976) and MacGregor (1976) discussed the effects of the choice of sampling interval on ARIMA process data. When a process is stationary and samples are taken less frequently in time, the autocorrelation of the sampled data will decrease. Thus, when the sampling interval is large enough, the data will appear to be uncorrelated. However, this approach discards the intermediate data, and therefore increases the possibility of missing important events in the process.

Instead of choosing a large sampling interval, moving averages of the process data with a fixed window size can be formed. Telser (1967) and Tiao (1972) discussed the autocovariances of moving sums of non-overlapping subgroups or temporal aggregation of AR and ARMA processes respectively. Brillinger (1973) showed that when $\{X_t\}$ is a stationary process and satisfies some regularity conditions, the non-overlapping means or batch means are asymptotically independent and normally distributed. Thus, in principle, when the batch size is large enough, the process formed by the batch means can be treated as white noise. For a stationary ARMA(1,1) process Kang and Schmeiser (1987) discussed the autocorrelation at lag 1 of the process formed by batch means as a function of the batch size and the parameters of the original process. Alwan and Radson (1992) discussed the time series behavior of the subsample means when the underlying process is an AR(1) process. Runger and Willemain (1995) and (1996) proposed batch means control charts when the underlying processes is an AR(1) processes.

In Zhang (2001), a general case of moving averages called generalized moving averages is discussed: the new data sequence is formed by the averages of n consecutive observations of the original process moving k step at a

time. Assuming that the original process is $\{X_t\}$, $\{Y_T(n,k)\}$ for T=1,2,.., the sequence formed by the generalized moving averages is defined as follows:

$$Y_1(n,k) = \frac{X_1 + X_2 + ... + X_n}{n}$$

$$\vdots \quad \vdots$$

$$Y_T(n,k) = \frac{X_{(T-1)k+1} + X_{(T-1)k+2} + ... + X_{(T-1)k+n}}{n}$$

When $n = k$, this is the case of averages of consecutive non-overlapping subgroups or batch means of the original process data, which was discussed in Brillinger (1973) and Runger and Willemain (1995). In Zhang (2001), it is shown that

(a) When a process is an MA(q) process, the process formed by the generalized moving averages is another MA(Q) process. When $k \geq n, Q$, the order of the new process is less than or equal to q.

(b) When k, the size of the moving step is not smaller than n, the size of averaging, and $k \leq n + q - 1$, the process formed by the generalized moving averages is an MA(1) process. In addition, if $k - n$ is a constant, the first lag autocorrelation of the new process decreases to zero when n approaches infinity. That is, the process becomes white noise. Because a stationary process can be approximated by an MA process by Wold decomposition Theorem, these results can be applied to any stationary process in practice. Thus, for a given set of stationary process data the size of averaging and the size of moving steps for which the process formed by the generalized moving averages to be treated as uncorrelated can be estimated. The estimation approach can be found in Zhang (2001). Once the process formed by the generalized moving averages are uncorrelated or approximately uncorrelated, the traditional SPC charts can be applied to it.

There are some other approaches for dealing with process autocorrelation. Consider a dynamic process system which consists of input streams and output streams as described in Zhang and Pollard (1994). When the input streams and output streams are autocorrelated, the mass imbalance, which is the difference between the input streams and output streams, is also autocorrelated in general. Zhang and Pollard (1994) and Zhang (1997a) showed that under some conditions, the autocorrelation of the mass imbalance is less than the autocorrelations of the input and output streams. Since the mass imbalance has weaker autocorrelation than those of the input and output streams, it is better to apply appropriate control charts to the mass imbalance process.

4 Process capability indices for autocorrelated data

Process capability indices are widely used in manufacturing industries to measure and monitor the performance of a process in meeting preset specification limits. As in the case of control charts, there is a concern about the assumption of the independence of the process observations. In this section we will briefly discuss the use of process capability indices when the process is autocorrelated.

The first proposed process capability index is C_p. Because of the shortcomings of C_p, other process capability indices including C_{pk} and C_{pm} were developed subsequently. The definitions and properties of these indices can be found in Kotz and Johnson (1993). When a process is autocorrelated, Yang and Hancock (1990) discussed the expectation and variance of $\overline{X}$ and S^2, which are used to form the process capability indices. Wallgren (1996a) and (1996b) modified the definitions of C_{pm} and C_{pk} by incorporating the autocorrelation when the process is an AR(1) process. However, these modified indices have limited usage in practice because they are only applied to AR(1) processes.

Zhang (1998b) discussed the use of C_p and C_{pk} for a stationary process. When a process is stationary, C_p and C_{pk} can still be used as when the process is an i.i.d. sequence. However, the corresponding variances for sample C_p and sample C_{pk} in this case will not be the same as given in Zhang, Stenback, and Wardrop (1990). When a process is stationary and Gaussian, Zhang (1998b) gave approximate variances of sample C_p and sample C_{pk}. Assume that the autocorrelation of the process at lag k is $\rho(k)$. Then

$$Var[\hat{C}_p] \approx C_p^2 \frac{F(n,\rho)}{2(n-1)^2 f^3(n,\rho)} \tag{8}$$

and

$$Var[\hat{C}_{pk}] \approx \frac{C_{pk}^2}{f^2(n,\rho)}\left[\frac{g(n,\rho)}{9nC_{pk}^2} + \frac{F(n,\rho)}{2(n-1)^2 f^2(n,\rho)}\right] \tag{9}$$

where $f(n,\rho)$, $F(n,\rho)$, and $g(n,\rho)$ are functions of $\rho(k)$ and the sample size, n:

$$f(n,\rho) = 1 - \frac{2}{n(n-1)} \sum_{i=1}^{n-1} (n-i)\rho(i)$$

$$F(n,\rho) = n + 2\sum_{i=1}^{n-1} (n-i)\rho^2(i) + \frac{[n + 2\sum_{i=1}^{n-1}(n-i)\rho(i)]^2}{n^2}$$

$$- \frac{2\sum_{i=0}^{n-1}\sum_{j=0}^{n-i}(n-i-j)\rho(i)\rho(j)}{n}$$

$$g(n,\rho) = 1 + \frac{2\sum_{i=1}^{n-1}(n-i)\rho(i)}{n}$$

The approximated variances in (8) and (9) are useful because they are in terms of true C_p or C_{pk} values as well as the autocorrelation and the sample size. In addition, when the process is an i.i.d. sequence, $f(n,\rho) = 1, g(n,\rho) = 1$ and $F(n,\rho) = n-1$. From (8) and (9), we have

$$Var[\hat{C}_p] \approx \frac{C_p^2}{2(n-1)} \tag{10}$$

$$Var[\hat{C}_{pk}] \approx C_{pk}^2[\frac{1}{9nC_{pk}^2} + \frac{1}{2(n-1)}] \tag{11}$$

which are the results in Bissel (1990). Based on the variances given in (8) and (9), interval estimators of C_p and C_{pk} can be obtained.

5 Conclusion

In this article, we discussed several approaches for dealing with process autocorrelation in the area of process monitoring. For process control charts, we discussed the use of process residual charts and modifications of existing SPC charts by adjusting the control limits to accommodate the process autocorrelation. We have made comparisons among the control charts when a process is autocorrelated. Other approaches for dealing with autocorrelation were also discussed. In addition, we also discussed process capability indices for autocorrelated data.

REFERENCES

Alwans, L.C. and Roberts, H.V. (1988). Time-series modeling for statistical process control. *Journal of Business and Economics Statistics* 21, 1025-1049.

Amin, R.W. and Lee, S.J. (1999). The effect of autocorrelation and occasional outliers on two-sided tolerance limits. *Journal of Quality Technology* 31, 286-300.

Atienza, O.O., Tang, L.C. and Ang, B.W. (1998). A SPC procedure for detecting level shift of autocorrelated processes. *Journal of Quality Technology* 30, 340-351.

Bagashaw, M. and Johnson, R.A. (1975). The effects of serial correlation on the performance of CUSUM tests II. *Technometrics* 17, 73-80.

Bissell, A.F. (1990). How reliable is your capability index. *Applied Statistics* 39, 331-340.

Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control.* San Francisco: Holden-Day.

Brillinger, D.R. (1973). Estimation of the mean of a stationary time series by sampling. *Journal of Applied Probability* 10, 419-431.

Harris, T.J. and Ross, W.H. (1991). Statistical process control procedures for correlated observations. *Canadian Journal of Chemical Engineering* 69, 48-57.

Jiang, Y. (2000). Detection capability of residual EWMA chart for stationary AR process data. *Communication on Applied Mathematics and Computation* 14, 24-32.

Jiang, Y. and Wang, J. (2001). Detection capability of X-chart for stationary invertible ARMA process data. *Chinese Journal of Applied Probability and Statistics* 17, 51-58.

Jiang, W., Tsui, K., and Woodall, W.H. (2000). A new SPC monitoring method: The ARMA chart, *Technometrics* 42, 399-410.

Johnson, R.A. and Bagshaw, M. (1974). The effect of serial correlation on the performance of CUSUM tests. *Technometrics* 16, 103-112.

Kang, K. and Schmeiser, B. (1987). Properties of batch means from stationary ARMA time series. *Operations Research Paper* 6, 19-24.

Kotz, S. and Johnson, N. (1993). *Process Capability Indices.* London: Chapman and Hall.

Longnecker, M.T. and Ryan, T.P. (1992). Charting correlated process data, *Technical Report 166*, Dept. of Statistics, Texas A&M University.

MacGregor, J.F. (1976). Optimal choice of the sampling interval for discrete process control. *Technometrics* 18, 151-160.

Montgomery, D.C. and Mastrangelo, C.M. (1991). Some statistical process control methods for autocorrelated processes. *Journal of Quality Technology* 31, 166-188.

Runger, G.C. and Willemain, T.R. (1995). Model-based and model-free control of autocorrelated processes. *Journal of Quality Technology*, 35, 283-292.

Runger, G.C. and Willemain, T.R. (1996). Batch-means control charts for autocorrelated data. *IIE Transactions*, 483-487.

Runger, G.C., Willemain, T.R. and Prahbu, S. (1995). Average run lengths for CUSUM control charts applied to residuals. *Communications in Statistics - Theory and Methods* 24, 273-282.

Schmid, W. (1997). On EWMA charts for time series. *Frontiers in Statistical Quality Control 5.* Heidelberg: Physica-Verlag, 115-137.

Telser, L.G. (1967). Discrete samples and moving sums in stationary stochastic processes. *Journal of American Statistical Association* 62, 484-499.

Tiao, G.C. (1972). Asymptotic behavior of time series aggregates. *Biometrika* 59, 525-531.

Vasilopoulos, A.V. and Stamboulis, A.P. (1978). Modification of control chart limits in presence of data correlation. *Journal of Quality Technology* 1, 20-30.

Wallgren, E. (1996a). *Properties of the Taguchi capability index for Markov dependent quality characteristics.* University of Örebro, Sweden.

Wallgren, E. (1996b). *Confidence limits for the process capability index Cpk for autocorrelated quality characteristics.* University of Örebro, Sweden.

Wardell, D.G., Moskowitz, H. and Plante, R.D. (1994). Run-length distributions of special-cause control charts for correlated process. *Technometrics* 36, 3-17.

Yashchin, E. (1993). Performance of CUSUM control schemes for serially correlated observations, *Technometrics* 35, 37-52.

Zhang, N.F. (1997a). Autocorrelation analysis of some linear transfer function models and its applications in the dynamic process systems. *Lectures in Applied Mathematics 33, 385-400, American Mathematical Society.*

Zhang, N.F. (1997b). Detection capability of residual chart for autocorrelated data. *Journal of Applied Statistics* 24, 475-492.

Zhang, N.F. (1998a). A Statistical control chart for stationary process data. *Technometrics* 40, 24-38.

Zhang, N.F. (1998b). Estimating process capability indexes for autocorrelated data. *Journal of Applied Statistics* 25, 559-574.

Zhang, N.F. (2000). Statistical control charts for monitoring the mean of a stationary process. *Journal of Statistical Computation and Simulation* 66, 249-258.

Zhang, N.F. (2001). What the generalized moving averages can do for process monitoring. *2001 Proceedings of Section of Physical and Engineering Sciences of American Statistical Society*, 730-735.

Zhang, N.F. and Pollard, J.F. (1994). Analysis of autocorrelation in dynamic processes. *Technometrics* 36, 354-368.

Zhang, N.F., Stenback, G.A., and Wardrop, D.M. (1990). Interval estimation of process capability index C_{pk}. *Communications in Statistics: Theory and Methods* 19, 4455-4470.