Book Reviews

Editor: Ananda Sen

Genomic Clinical Trials and Predictive Medicine
Richard M. Simon
Cambridge University Press, 2013, xiii + 144 pages, $44.99, paperback
ISBN: 978-1-1074-0135-8

Readership: Statisticians, clinical investigators and translational scientists.

This book is about how genomic data can be used in designing and analysing clinical trials. It is timely given the rapid increase in the availability of genomic data and interest in using it to personalise treatment. In addition to covering clinical trial design, this book covers the development of diagnostic tests that use biomarkers to identify patients that should receive a particular treatment. However, this book is not a general tutorial on how to analyse genomic data. It also does not cover the large statistical literature on identifying predictive markers from observational data. Rather, it is focused on the role of genomic data in designing clinical trials. A clear strength of the book is the presence of many practical examples, which serve to make concrete the ideas being presented. In terms of subject matter, the book is heavily oriented towards oncology, although the statistical concepts and approaches will certainly be more broadly applicable.

The first chapter contains a brief introduction to clinical trials. It covers the differences between phase I, II and III trials, power calculations for designing trials and interim analyses. This will be helpful for statisticians who have some familiarity with clinical trials but need a refresher. It will not be sufficient for someone new to clinical trials. For clinical investigators in need of a statistics refresher, there is an appendix that covers statistical basics such as hypothesis testing, regression modelling and Bayesian methods.

The second chapter covers 'Actionable Prognostic Biomarkers'. A somewhat common misconception, even amongst statisticians, is that only predictive markers are actionable. However, as the author points out, prognostic markers are actionable in certain cases. The highlighted case is the situation when standard of care treatment is intensive and the marker identifies patients who have good prognosis with lower intensity therapy. If their prognosis is good enough (e.g. survival close to 100%) with the low intensity therapy, then there may be little to gain by using the more intensive therapy. Of course, identification of such markers requires data on patients with low intensity treatment that is no longer standard of care.

Identifying the appropriate population of patients for treatment is often a secondary aim of phase II trials but is particularly important in the context of molecularly targeted agents. Chapter 3 covers phase II trials and the role of genomic markers in these trials. A number of phase II designs incorporating biomarkers have been proposed and several of these are discussed in this chapter. Much of the work in this area has involved Bayesian adaptive designs. Perhaps two of the best known such trials are the BATTLE (Zhou *et al.*, 2011) and I-SPY2 (Barker *et al.*, 2009) trials. The former is discussed in Chapter 3, although the reference given (Kim & Simon, 2011) appears to be the wrong reference. The I-SPY 2 trial is not mentioned.

For phase III trials of molecularly targeted agents, an important design choice is whether to include all patients or only patients who are 'marker positive'. The definition of marker positive can take on many forms but will typically be based on biomarkers of the targeted pathway. If the treatment effect is larger in marker-positive patients, then including only marker-positive patients allows a smaller trial. On the other hand, if the new treatment is beneficial in marker-negative patients, then only an all-comers design would allow this to be assessed. Chapters 4 and 5 discuss the relative merits of these various trial designs.

Chapter 7 deals with developing and validating predictive markers in a randomised trial. The goal in this setting is to use one or more biomarkers to jointly identify a group of patients in which the experimental treatment is significantly better than control. The author discusses the efficiency of various designs, the use of cross-validation or split samples and controlling overall type I error rates. Additional details and discussions regarding various statistical approaches to prognostic classification are given in Appendix B.

The final chapter (Chapter 8) of the book provides an overview of what has been termed the 'Prospective-Retrospective' design. This design combines features of a prospective clinical trial, including a prospectively written analysis plan, with archived samples from a previously conducted randomised trial.

This book will be a valuable resource to those involved in genomic clinical trials. The author touches on many of the important issues in this field and provides a useful selection of approaches to handling them. References are given in almost every section and will be helpful for those wanting more detail. The writing is very clear, and the issues are often illustrated with real life examples. Software to implement many of the proposed designs and analyses is available online at the provided link.

Matthew Schipper: *mjschipp@umich.edu*
Departments of Radiation Oncology and Biostatistics
University of Michigan, Ann Arbor, MI 48109, USA

References

Barker, A.D., Sigman, C.C., Kelloff, G.J., Hylton, N.M., Berry, D.A. & Esserman, L.J. (2009). I-SPY 2: An adaptive breast cancer trial design in the setting of neoadjuvant chemotherapy. *Clinical Pharmacology and Therapeutics*, **86**, 97–100.
Kim, K.I. & Simon, R. (2011). Probabilistic classifiers with high dimensional data. *Biostatistics*, **12**(3), 399–412.
Zhou, X., Liu, S., Kim, E.S., Herbst, R.S. & Lee, J.J. (2008). Bayesian adaptive design for targeted therapy development in lung cancer – a step toward personalized medicine. *Clinical Trials*, **5**, 181–193.

Statistical Learning with Sparsity: The Lasso and Generalizations
Trevor Hastie, Robert Tibshirani and Martin Wainwright
Chapman and Hall/CRC, 2015, 367 pages, £57.99, hardcover
ISBN: 978-1-4987-1216-3

Readership: Statistics graduate and advanced undergraduate students, as well as practitioners in the biological and physical sciences who have a basic knowledge of probability, statistics and numerical analysis.

In the current era of *Big Data*, vectorised computing and high-speed networking, it is refreshing to find a book that approaches statistical analysis and modelling from an entirely different perspective, namely, that of sparse statistical modelling where only a small number of non-zero

parameters (or weights) are envisaged. With such an approach, simpler models are contemplated for which parameter estimation is necessarily easier and for which model interpretation is generally simpler and often more intuitive.

To my mind, one of the major concerns and drawbacks with *Big Data* modelling is that with a large amount of data and reliance on multi-parameter models, identifiability of parameters quickly becomes an issue when models are fitted to data and further that such models describing past data sets can then have either little, or no, predictive ability. This is true in modelling of the marine environment and the modelling of biological processes of fish stocks (O'Brien & Little, 2006) – two areas where the sparsity modelling and techniques described by the authors in their text will be most welcome and have immediate applications.

The authors of this book provide a well-written and authoritative text on how best to deal with sparse statistical models in which only a relatively small number of parameters (or *predictors*) play an important role. This book is easily accessible to anyone with a basic knowledge of linear algebra, a familiarity with the method of least-squares, knowledge of the L1-norm and applied numerical analysis. The authors have adopted a simplified notational scheme to enhance readability. There is a section at the end of each chapter providing exercises that can help in pedagogic delivery in classroom and aid self-instruction for private study as well. The book's methods and techniques are demonstrated throughout by applications that help reinforce the theory presented – thus, making this a text accessible to both theoreticians and practitioners.

The field of sparse statistical modelling is rapidly expanding, and the Least Absolute Selection and Shrinkage Operator acronym (LASSO) of the book's title was introduced nearly 20 years ago in the mid-1990s. This is an area that I predict will continue to develop and expand as budget reductions and financial constraints necessitate a more parsimonious approach to data collection and routine monitoring – possibly, with a move away from *Big Data per se*.

I thoroughly recommend this book and look forward to reading about the future developments in the methods and their applications which will no doubt soon be coming.

Carl M. O'Brien: *carl.obrien@cefas.co.uk*
Centre for Environment, Fisheries & Aquaculture Science
Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 0HT, UK

Reference

O'Brien, C.M. & Little, A.S. (2006). Incorporation of Process Information into Stock-recruitment Models. ICES Cooperative Research Report, No. 282, 152 pp.

Bayesian Nonparametric Data Analysis
Peter Müller, Fernando Andres Quintana, Alejandro Jara and Tim Hanson
Springer, 2015, 301 pages, £76.50/$109.00, hardcover
ISBN: 978-3-319-18967-3

Readership: Individuals with an understanding of Bayesian principles and readers interested in non-parametric applications from a data analysis perspective.

The book *Bayesian Nonparametric Data Analysis* by Peter Müller, Fernando Andres Quintana, Alejandro Jara and Tim Hanson presents different non-parametric Bayesian methods and models useful in the context of data analysis. This book provides a good review of well-known non-parametric Bayesian methods. The book is organised from a data analytic perspective. In that vein, it discusses important statistical inference problems, starting from density estimation,

followed by regression and non-parametric inference. For each of the inference topics, authors review relevant non-parametric Bayesian models and approaches including Dirchlet process (DP) models and variations, Pólya trees, wavelet-based models, neural network models, spline regression, CART, dependent DP models and model validation with DP and Pólya tree extensions of parametric models.

Chapter 1 in the book starts with a general overview of the non-parametric and semiparametric inference and discusses the distinction between the Bayesian and the classical non-parametric inference. Chapters 2 and 3 are devoted to density estimation. The main emphasis of this chapter is on the Dirichlet process. Non-parametric regression is discussed in Chapter 4. This chapter includes non-parametric priors on residual distribution, non-parametric mean functions and fully non-parametric regression (also known as density regression).

Chapter 5 introduces methods for categorical data. In this chapter, some contingency tables for multivariate categorical data and some methods specifically for ordinal data can be found. Chapter 6 of the book discusses applications for survival analysis. In this chapter, some models for hazard function and random probability models for event times can be found. Chapter 7 considers the use of hierarchical models. Models for random clustering and for feature allocation problems are discussed in Chapter 8, along with some very useful methods in machine learning. Finally, Chapter 9 concludes with a brief discussion of some additional inference topics. The appendix is dedicated to the DPpackage, a public R package that implements inference for many of the models that are discussed in the book.

The discussed methods are illustrated with a wealth of examples, ranging from stylised applications to case studies from recent literature. The book is a good reference for statisticians interested in Bayesian non-parametric data analysis. It is well-written and structured. Readers can find the algorithms, examples and applications easy to follow and extremely useful. This book makes a good contribution to the literature in the area of Bayesian non-parametric statistics.

Diego Andres Perez Ruiz: *diego.perezruiz@manchester.ac.uk*
School of Mathematics
University of Manchester, Oxford Road, Manchester M13 9PL, UK

Current Trends in Bayesian Methodology with Applications
Satyanshu K. Upadhyay, Umesh Singh, Dipak K. Dey and Appaia Loganathan
Chapman and Hall/CRC, 2015, 680 pages, £54.39, hardcover
ISBN: 978-1-4822-3511-1

Readership: Graduate students, statistics researchers and practitioners.

This is an impressive and useful book. It covers a diversity of latest methodological and applied aspects of Bayesian statistics. It includes Bayesian methodologies such as Bayesian variable selection, clustering, multi-state models, factor analysis, spatial statistics and Bayesian computing. It also covers a wide range of applications in economics, astrophysics and biomedicine.

There is a surge in Bayesian publications in recent years. This book will provide the opportunity to gaze through the novel Bayesian methodologies that are scattered around in different journals.

This book consists of 30 chapters. Each chapter is independent and is motivated by real applications. The chapters are well-structured. Each chapter starts with an overview of a problem

area and provides specific background. This motivates the discourse quite nicely. For example, the background of neuroscience and imaging is very helpful for people who are interested in this field but do not have an exposure to it. The problem that needs to be solved is clearly stated through real applications, and a methodological solution is proposed. In the solution, complicated theoretical details are left out. The book includes an extensive list of references at the end as a useful aid for the researchers. Overall, it is an exciting book and would benefit researchers who are interested in applying advanced Bayesian methods to practical problems.

Lili Zhao: *zhaolili@umich.edu*
Department of Biostatistics
1415 Washington Heights, Ann Arbor, MI 48109, USA

Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction
Guido W. Imbens and Donald B. Rubin
Cambridge University Press, 2015, xix + 625 pages, $60.00, hardcover
ISBN: 978-0-5218-8588-1

Readership: Undergraduate and graduate students and practitioners in any field who have a knowledge of matrix algebra.

This is a large book with over 600 pages consisting of 26 chapters that are distributed over five topical themes. The book is well-written with a very comprehensive coverage of many issues associated with causal inference. As can be seen from its Table of Contents, the book uses multiple perspectives to discuss these issues including theoretical underpinnings, experimental design, randomization techniques and examples using real-world data. For each of the data sets used, the authors give some background and references. While the theoretical development of the topics depends heavily on matrix algebra, the authors have put as many proofs as possible as appendices. This helps the chapters flow better.

The authors list their decision to focus their discussion on binary treatments as a limitation. However, I see this as an advantage, because it allows the authors to focus more on the causal inference issues that occur in actual studies. As the authors point out in the Introduction, many of their results can be generalized easily to multi-valued treatments.

My only disappointment with the book is that the authors only lightly discuss some of the limitations and philosophical, ethical and practical issues associated with particular causal inference designs. Adding short sections within relevant chapters discussing limitations and these other issues would have made a great book even better.

Carol Joyce Blumberg: *cblumberg@gmail.com*
Professor Emerita, Winona State University and Retired, U.S. Energy Information Administration
2035 Lyttonsville Rd, Silver Spring, MD 20910-2264, USA

Bayesian and Frequentist Regression Methods
Jon Wakefield
Springer, 2013, 674 pages, €79.49, hardcover
ISBN: 978-1-4419-0924-4

Readership: Advanced undergraduate and graduate students in statistics as well as researchers and practitioners.

Jon Wakefield's *Bayesian and Frequentist Regression Methods* is an impressive book both in terms of its coverage and its contents. It is, by no means, an introductory book, rather, in my opinion, the book provides an overview with insightful discussions and some details (but not intricate details), on a selected list of topics in regression, and is more suited for readers who have some familiarity with inference and regression topics. In some sense, it is two books into one comprehensive volume; the first four chapters providing coverage of Frequentist and Bayesian inference and variable selection while the remaining chapters delving into increasingly complex regression models. An accompanying website at http://faculty.washington.edu/jonno/regression-methods.html includes R (and some WinBUGs) codes for all analyses and figures in the book.

The first four chapters of the book, covering all of the first 190 pages, are overview with some details on classical and modern Frequentist and Bayesian inference and variable selection topics. The Frequentist inference emphasises estimating functions but also includes conditional, marginal, profile, quasi likelihoods and sandwich estimators, and the Bayesian inference includes Bernstein-von Mises theorem as well as substantive priors, direct sampling and Markov chain sampling. The hypothesis testing and variable selection is a thoughtful review of p-value, Neyman–Pearson and Bayes factor approaches followed by multiple testing, FWER, FDR and similar notes. It is a thought-provoking read for someone who is familiar with the topics.

The regression part of the book is subdivided into independent data (Chapters 5, 6 and 7), dependent data (Chapters 8 and 9) and non-parametric modelling (Chapters 10, 11 and 12). Each chapter is accompanied by data sets such as Minnesota lung cancer, contraception, seizure, BPD and birthweight data sets. The linear regression in Chapter 5 includes review of least-square and likelihood methods, a short review of Bayesian inference, with analysis of variance, regression diagnostics and diagnostics. The general regression models in Chapter 6 mostly cover generalised linear models (with a brief mention of other non-linear models) and primarily focus on likelihood and quasi-likelihood methods with a subsection devoted to Bayesian inference. Chapter 7 is on binary data models, logistic regression, likelihood, quasi-likelihood, Bayesian and conditional likelihood inference with notes on overdispersion, collapsibility and case-control studies.

The dependent data sub-block of the book focuses on linear mixed and generalised linear mixed models (GLMM) as well as non-linear mixed models and overviews likelihood, Bayesian and generalised estimating equation approaches. It includes interesting subtopics such as GLMM with spatial dependence and Markov random field priors with example and accompanying code.

In the non-parametric modeling sub-block, the first chapter briefly reviews Mallows Cp, general cross-validation, AIC, Ridge regression and Lasso as well. The other two chapters include spline and kernel regression for linear, linear mixed and generalised linear mixed models, as well as regression trees followed by an overview of classification methods.

The book impressively covers both Bayesian and Frequentist approaches in each of linear to advanced non-parametric regression chapters, but I felt that while the book's title is Bayesian and Frequentist (following the alphabetical ordering) and the two have equal footings in the inference block of the first four chapters, Bayesian played a second fiddle in the remaining regression chapters. In terms of my other critiques and nitpicking, while concepts of variable selection and multiple testing are covered in inference block of the book, they are not used much in the regression part of the of the book. I would also have preferred to see AIC, cross-validation and perhaps Lasso introduced much earlier in the book, possibly before or within the linear regression chapter. Finally, while plenty of regression models are discussed in the book, one notable exception in the coverage of this book is regression models for time-to-event data.

Overall, Jon Wakefield's *Bayesian and Frequentist Regression Methods* provides an excellent parallel treatment of Frequentist followed by Bayesian approaches to linear, generalised linear, generalised linear mixed and non-parametric regression models. This book is impressive both in terms of its coverage and its contents and is an exceptional resource for students and researchers who have some familiarity with these topics.

Sanjib Basu: *basu@math.niu.edu*

Northern Illinois University, Division of Statistics, De Kalb, IL 60115, USA

John Napier: Life, Logarithms and Legacy
Julian Havil
Princeton University Press, 2014, xv + 279 pages, £24.95/$35.00, hardcover
ISBN: 978-0-691-15570-8

Readership: Everyone interested in the history of mathematics.

John Napier invented logarithms as an aid to calculation and in 1614 provided extensive tables for them in his *Mirifici logarithmorum canonis descriptio*. The current book was published on the 400th anniversary of this invention. Napier's tables of logarithms may seem odd to the modern reader (or the slightly less modern for those who experienced the use of logarithm tables in school prior to the emergence of pocket calculators). The tables are readily accessible on Early English Books Online (EEBO) or for other online sources. There are 90 pages of tables organized in pairs of facing pages, one pair of pages for each of angle 0º to 44º with each line in a pair of pages corresponding to the minutes 0´ through 60´. The columns in the tables are values of sines (or cosines since $\cos(\theta) = \sin(90 - \theta)$) at each angle, the logarithms of these values and the logarithms of the tangents at each angle. Clearly, Napier's original purpose in inventing logarithms was to ease the calculation of trigonometric ratios, problems related to astronomy, navigation and cartography.

As expected, the book begins with a biography of Napier with an acknowledgement that very little source material about his life survives. Havil very nicely places Napier in his historical context. If the reader has not read the Table of Contents in advance, he or she would be jolted by the second chapter. Rather than the expected logarithms, it is a treatment of Napier's analysis of the Book of Revelation, the last book of the Christian Bible and probably the most difficult to understand. Napier predicted the end of the world by 1786. Initially published in 1593, Napier's analysis went through 10 editions in English and another 13 in other languages (Dutch, French and German corresponding to the languages of the reformed church). Once again, Havil nicely puts Napier's analysis into the historical context of the Protestant Reformation. Readers who do receive a slight jolt from this chapter should quickly flip through the rest of the book to find that Havil devotes a chapter to the analysis of each of Napier's publications: *A plaine discovery of the whole Revelation of Saint John set downe in two treatises* (1593), *Mirifici logarithmorum canonis descriptio* (1614), *Rabdologiae, seu Numerationis per virgulas libri duo* (1617), *Mirifici logarithmorum canonis constructio* (1619) and *De arte logistica* (in manuscript form, published 1839). Chronologically, the tables of logarithms come two decades after Napier's theological work. Only a glance at the *Descriptio* will make you realize why 20 years – 180 pages of hand calculations to seven and eight digits of accuracy does take some time. The *Rabdologiae* is a description of Napier's physical calculation systems including his best known, Napier's bones or Napier's rods. While the *Descriptio* contains the tables and instructions for their use, the posthumously published *Constructio* is a treatise on how the logarithms in the table

are calculated. Finally, *De arte logistica* is Napier's treatment of calculation in arithmetic and algebra. Taken as a whole, the main theme of Napier's corpus is calculation. From the title, calculation in the study of the Books of Revelation is not obvious. Napier spilt much ink interpreting the treatment of dates and time in the Bible so that calculations related to the end of time could be made. Havil has given a thorough analysis of each of these publications.

Napier's logarithms are sometimes equated to natural logarithms. The relationship is close but not a straight equivalence. Suppose $L(x)$ is Napier's logarithm for a value x. Then the connection to natural logarithms is expressed as

$$L(x) = \frac{\ln\left(\frac{10^7}{x}\right)}{\ln\left(\frac{10^7}{10^7-1}\right)}$$

The relationship comes from Napier's method of construction. Napier constructed his logarithms using arithmetic and geometric progressions. Two points A and B on a line move such that A's speed is constant (set at 10^7 so that A travels a distance of 10^7 in one unit of time) and B's speed decreases geometrically at a rate such that the remaining distance to travel to the end point 10^7 at time t is $10^7\left[(10^7-1)/10^7\right]^t$. The time t is Napier's logarithm of the remaining distance.

What I have given for $L(x)$ leads to a minor quibble with Havil. He gives the relationship of $L(x)$ to natural logarithms as $L(x) = 10^7 \ln\left(\frac{10^7}{x}\right)$ (page 77). It is a very minor difference, but a difference all the same that does not follow from Napier's method of construction. There are some other minor quibbles. For example, Havil seems to imply (page 24) that King Phillip II of Spain was married to Mary, Queen of Scots rather than Mary I, Queen of England. In terms of the subject material covered, this is a minor mistake.

Why is this book being reviewed for the *International Statistical Review*? It contains no data or statistical theory. It does, however, touch on the origins of one of the most common functions used in statistical theory, the natural logarithm and its inverse function the natural base e. And it does so in a thorough and interesting way.

David Bellhouse: *bellhouse@stats.uwo.ca*
Department of Statistical and Actuarial Sciences
University of Western Ontario, London, Ontario N6A 5B7, Canada

Analysis of Categorical Data with R
Christopher R. Bilder and Thomas M. Loughin
Chapman & Hall/CRC, 2015, xiii + 533 pages, £49.99/$89.95, hardcover
ISBN: 978-1-4398-5567-6

Readership: Advanced undergraduate and graduate students in statistics and related areas, as well as researchers and practitioners.

This book presents an extensive introduction to analysis of categorical data with R. The context is relevant for a multitude of application areas such as biology, ecology, medicine and sports just to name a few. Recent model-building techniques are covered. Readers of the book are recommended to have a basic background in regression. Although a strong mathematics background is not required, a basic understanding of calculus and matrices would be helpful.

The first two chapters introduce binary response variables and the associated regression models, with fundamental ideas like odds ratio, Monte Carlo simulation and generalised linear

models. Chapter 3 focuses on a multi-category response and discusses in detail contingency tables, nominal, ordinal and other regression models. Chapter 4 details methods to analyse count response variables with a focus on Poisson regression models. Chapter 5 is devoted to model selection and evaluation issues including tools to assess model fit and overdispersion, which is relevant for all regression models discussed. Chapter 6 covers additional or advanced topics such as error testing, exact inference, mixed models and Bayesian methods for categorical data.

At the end of the book, there are two Appendices. Appendix A includes a basic and systematic introduction to R, for those unfamiliar with the program. Throughout the book, R is used not only as a data analysis tool but also as a learning tool. It also provides a basic regression example in R as a lead-in to the more complex topics presented in the book. Appendix B covers likelihood methods, providing the principal, mathematical examples, Newton–Raphson algorithm, Delta method and inference tests (including Wald, likelihood ratio and score tests).

The book takes an easy-to-understand approach by partnering practical explanations with numerous illustrative examples. These step-by-step examples are supportive and cover the underlying definitions, ideas and methods behind the practical data analysis using R code. A number of examples use the same data set, but each example focuses on a different aspect, giving students a broader understanding of the data set. To help students apply their knowledge, the book has also provided an extensive number of exercises. Solutions to all these exercises are available to instructors.

The book has a website http://www.chrisbilder.com/categorical/ with the data sets and R code used for the worked examples and also the authors' in-class recorded lectures, which supports students to be flexible in their learning. The textbook can also be a very useful reference.

Shuangzhe Liu: *shuangzhe.liu@canberra.edu.au*
Program of Mathematics and Statistics
University of Canberra, Canberra ACT 2601, Australia

Meta-Analysis: A Structural Equation Modeling Approach
Mike W.-L. Cheung
Wiley, 2015, xxii + 378 pages, €60.80/£45.00, hardcover
ISBN: 978-1-119-99343-8

Readership: Statistics graduate students, as well as researchers and practitioners in social, behavioural and medical sciences with basic knowledge of either structural equation modelling or meta-analysis.

This book is a comprehensive discourse on how to conduct meta-analyses within the structural equation modelling (SEM) framework. The book also contains a lot of pragmatic examples, mainly on how to use the R-package metaSEM by the same author. That package is a meta-analysis frontend for the OpenMx R-package for structural equation modelling.

The first chapter in the book very briefly reviews the history and selected applications of meta-analysis and SEM. The second chapter is a review of key concepts and topics of SEM that will be needed for the meta-analysis applications. Although the chapter is only about 30 pages, it provides a nice overview with just enough statistical details and extensions of the basic modelling required to follow the presentation in later chapters. Chapter 3 is an outlier in the book. It introduces certain basic techniques to estimate common effect sizes and related sampling variances in a meta-analysis, but this topic may require supplements from other sources for most readers not familiar with meta-analysis.

Chapters 4–8 have similar structure: they start with a description of the meta-analytic model(s), followed by SEM formulation of the same and conclude with illustrations on the application and estimation of the models using R. In the theoretical part of the chapter, the model is in focus and links to applications are secondary The applications part of the chapter provides illustrations with real data. The complexity of models is elevated in each chapter: Chapter 4 is on univariate meta-analysis, Chapter 5 discusses multivariate meta-analysis, Chapter 6 describes three-level meta-analysis, Chapter 7 talks about meta-analytic structural equation modelling and in Chapter 8, advanced topics in SEM-based meta-analysis are discussed.

As the formulation, analysis and estimation of the meta-analytic models within SEM framework require certain manipulations, it may be a bit challenging to utilise popular SEM software for meta-analysis. Chapter 9 addresses this challenge by providing pragmatic illustrations of analysis of many of the discussed models with M*plus*. This is naturally very handy for the users of that software who are not willing to use R, and may also help to write the code required to estimate the same models in other SEM software packages.

The book also contains an appendix that gives a brief introduction to R as well as to R-packages metaSEM and OpenMx. This appendix is already a bit out-dated while writing this review, but not in a bad way as the installation of required components is much easier now, because all of them were released to the CRAN library. Wiley's companion site for the book was not the easiest one to find, but it is also definitely worth checking as all the R scripts as well as M*plus* scripts used in the book were available there. If there had been one homepage containing updated links and information about the book, related packages and scripts, some extra work for hunting them down could have been avoided.

It may be noteworthy that the freely available vignettes and other material related to the metaSEM package contain material that are quite similar to those illustrated in the book. The added value of the book seems to be a more comprehensive theoretical background material for the models. The book clearly provides one of the most comprehensive reviews of the recent developments and open questions on the SEM approach for meta-analysis

As a conclusion, I would recommend this book for statisticians and researchers who are interested in presenting their models in the SEM framework or who need complicated meta-analysis that actually require the use of SEM models. On the other hand, the value of SEM formulations of the basic applications of simple meta-analytic models may not be obvious, so the researchers interested in well-established meta-analytic models alone will likely benefit more from material that focuses on traditional meta-analysis rather than the approach adopted in this book.

Reijo Sund: *reijo.sund@helsinki.fi*
Centre for Research Methods, Department of Social Research
University of Helsinki, P.O. Box 18, FI-00014, Finland