International Statistical Review (2009), 77, 3, 463–482 doi:10.1111/j.1751-5823.2009.00095.x

Short Book Reviews

Editor: Simo Puntanen

Survival and Event History Analysis: A Process Point of View
Odd O. Aalen, Ørnulf Borgan, Håkon K. Gjessing
Springer, 2008, xviii + 539 pages, € 64.95 / £ 58.99 / US$ 84.95, hardcover
ISBN: 978-0-387-20287-7

Table of contents

Readership: Practicing statisticians as well as theoreticians interested in survival analysis. Also suitable for a graduate course. Very well written.

A thesis written in 1975 at Berkeley changed the way we think of survival analysis. Point processes and martingales became fundamental parts of this area. It led to deep theoretical treatise by Andersen et al. (1993, Springer); it became part of even excellent applied texts like Klein and Moeschberger (2003, Springer) and the equally good text by Fleming and Harrington (1991, Wiley) which blended theory and application. The thesis was written by Odd O. Aalen.

Aalen, Borgan and Gjessing have written a new book which is also likely to have a profound influence on the subject, possibly both from the classical and Bayesian point of view, though Aalen et al. only discuss the (philosophically) classical approach.

The book is based on point processes but, as mentioned earlier, written for both practitioners and theoreticians. Deep facts about these processes as well as martingales and stochastic integrals are introduced and used throughout with clarity and intuitive insight. Most readers who are seeing these topics for the first time will not miss a more theoretical approach or lack of proofs, but will still need to work hard. A rigorous theory is too technical for the intended large readership and must be learnt separately by those who are interested. Fleming and Harrington (1991), mentioned above, would be a good source.

The book then covers all the standard things like Kaplan–Meier, Cox regression model and partial likelihood, frailty models etc., taking throughout a point process based approach to survival.

Some novelties in the first seven chapters include a deep and novel treatment of additive regression models for the hazard function. Additive models have not been as popular as the

multiplicative model of Cox. The multiplicative models go better with both the actuarial approach and the point process approach. Additive models may also lead to negative estimates of the hazard rates at different points. Aalen et al. provide a good discussion and explain why in some cases additive models are to be preferred. They introduce new models and a martingale based analysis. Point processes are also used.

The remaining chapters, namely, chapters 8 through 11, are either quite new or at the existing cutting edge. The chapter on Causality goes much beyond the usual formulation of causality for epidemiologists due to Bradford Hill. Connoisseurs will notice that even the famous Feynman–Katz formula makes an appearance in the appendices.

There is however something I miss . . . the outstanding work of Groenboom and Wellner on the nonparametric MLE of the survival function under interval and double censoring. Groenboom, alone or with Wellner, proved consistency, rates of convergence, limiting distribution, where one existed, and also presented much faster algorithms than those due to Turnbull, which are commonly used. Sources are the monograph by Wellner and Groenboom (1992, Birkhauser) and Groenboom's St Fleur lecture notes in the Springer lecture note series of the late eighties or early nineties. All this material is very technical and either unknown or inaccessible to most people interested in survival analysis. A nontechnical exposition either in a book or article would improve awareness of these neglected but both useful and beautiful results.

Jayanta K. Ghosh: *ghosh@stat.purdue.edu*
Department of Statistics, Purdue University
West Lafayette, IN 47909, USA

Creative Mathematics

H. S. Wall
Mathematical Association of America, 2009, xx+195 pages, £ 27.99 / US$ 52.95, hardcover
ISBN: 978-0-88385-750-2

Table of contents

Readership: Everybody interested in teaching mathematics.

The first edition of *Creative Mathematics* was published in 1963 by the University of Texas Press in Austin and reprinted in 2006 by the Educational Advancement Foundation. The current book is its revised and updated edition.

The preface begins as follows: "This book is intended to lead students to develop their mathematical ability, to learn the art of mathematics, and to create mathematical ideas. This is not a compendium of mathematical facts and inventions to be read over as a connoisseur of art looks over the paintings in a gallery. It is, instead, a sketchbook in which readers may try their hands at mathematical discovery."

The book applies "the Moore method" or "the Texas method", which can also be described as "do-it-yourself-method". The emphasis is strongly geometric. For example, "simple graphs" are considered instead of functions.

One problem of this method is that it takes more time than ordinary teaching. Another problem is that it may be too tough for average students, but the author is optimistic. At the end of the preface he writes as follows:

"I find that this method of teaching can be an inspiration both to students who discover that they have unusual mathematical ability and decide to specialize in mathematics or a mathematical science, and to those of less ability. Students who are unable to prove any but the simplest propositions still get training in language and logical thinking and also obtain the benefit of seeing the presentations by the students who prove their results in class. Moreover, they find out something about what mathematics really is."

Jorma Kaarlo Merikoski: *jorma.merikoski@uta.fi*
Department of Mathematics and Statistics
FI-33014 University of Tampere, Finland

R for SAS and SPSS Users

Robert A. Muenchen
Springer, 2009, xvii + 470 pages, € 59.95 / £ 53.99 / US$ 74.95, hardcover
ISBN: 978-0-387-09417-5

Table of contents (including number of pages for each chapter)

Readership: Users of SAS or SPSS who want to learn the basics of R.

The author is the manager of the Statistical Consulting Center at the University of Tennessee and has been consulting in statistics for over 25 years. The author's first sentence, "While SAS and SPSS have many things in common, R is very different" suggests the difficulty in writing this book, whose title conveys its goal. (This reviewer's familiarity is with SAS and R, so the review will be aimed accordingly – however, analogous statements will usually apply to SPSS users.) As the number of pages for each chapter indicates, the author has adopted a non-standard idea of "chapter." Chapter 2 is short but useful. Chapters 4 and 5 are correct technically – we learn downloading and updating packages, using the source and sink functions, and even running R

from SPSS and using the Rattle package for Data Mining, but there are very few opportunities for the reader to have a sense of what R does at a simple level. Chapters 8 and 9 provide such a sense, and the latter chapter introduces the author's comparison paradigm: show an example in R (e.g., reading a delimited text file) in detail, and later show the corresponding SAS and SPSS code. (In other cases, some basic SAS and SPSS ideas are reviewed before the R code is shown.). The connection among the examples should be clear, and this is often achieved in Chapters 10–12, and much of the important Chapter 14. However, in some examples in Chapter 14, the examples no longer coincide. For example, in SAS, we find the N, mean, sd, minimum, and maximum of a number of variables in essentially two PROC MEAN lines of code, but the much lengthier code in R covers only the mean and N, and the results appear in separate vectors. The author introduces a number of useful packages in his text, but oddly "Hadley Wickham's powerful reshape package" is only used to supply a rename function for a simple aggregation example, instead of using the melt/cast pair to show how to produce output like that of PROC MEANS. This type of omission can give the impression that R cannot be used efficiently for serious aggregation problems. Similar problems arise elsewhere, such as in the discussion of the first.var and last.var type of variables created in SAS. While the author shows a neat way of *extracting* such variables in R, the user will be left wondering how to actually *mimic* first.var and last.var (as I did when I tried to reference this book to solve this problem). The chapters on graphics are good introductions, with many examples, of two approaches to graphics in R. Here, the link to corresponding graphs in SAS is very slim, but this seems fine because of R's much richer capabilities. Chapter 23 closes with a good set of examples of basic statistical modeling, including correlation, linear regression, ANOVA, and some non-parametric tests.

The author has a website of errata and the files used in the text (but the reader will need to figure out which files correspond to which examples). The number of typos and name changes (e.g., "mydata.Rdata" in the download, but "myWorkspace.RData" in the book) in the first edition may be annoying.

It must be a challenge to write such a text – if the second edition can provide better parallels of more realistic (complex) data manipulation in R (and with more complex data sets), the book will become very useful to SAS and SPSS users who wish to use R. The author may also better serve users by showing only his recommended way for R code to mimic SAS and SPSS code – having 3 or 4 options for calculating means with BY processing, for example, may be both intimidating and confusing.

Joseph G. Voelkel: *jgvcqa@rit.edu*
Applied Statistics, Kate Gleason College of Engineering, Rochester Institute of Technology
98 Lomb Memorial Dr., Rochester, NY 14623-5604, USA

Self-Normalized Processes: Limit Theory and Statistical Applications
Victor H. de la Peña, Tze Leung Lai, Qi-Man Shao
Springer, 2009, xiii + 275 pages, € 59.95 / £ 53.99 / US$ 89.95, hardcover
ISBN: 978-3-540-85635-1

Table of contents

Readership: Research workers in applied probability.

The t-statistic for testing a hypothesised mean value μ, based on a random normal sample of size n, has form $T_n = (S_n - n\mu)/W_n$, where S_n is the sample sum and W_n is proportional to the sample standard deviation. This is the prototypical example of a self-normalized sum, W_n being based on the sample values along with S_n. This book describes extensive developments of T_n, regarded as a process with parameter n, and its completion marked the centenary of Gosset's (1908) *Biometrika* paper in which the distribution of the t-statistic was derived.

The stated purpose of the book is to present the main techniques involving the asymptotic (large n) distribution theory of a variety of generalisations of T_n. As such, it serves as a reference text for a special-topic course for PhD students; each chapter after the first ends with a collection of problems and the material is based on such a course taught by two of the authors at Stanford and Hong Kong.

The basic form of T_n is generalised early on, in Section 2.2.1, to sums of independent random variables of any common distribution, and with a slightly modified form for W_n. Detailed asymptotic results follow for T_n and its variants, including Laws of Large Numbers, Central Limit Theorems, Law of the Iterated Logarithm, and large and moderate deviations. Later on, sums of i.i.d. variables are generalised to empirical processes, U-statistics, martingales, etc. The style is largely Theorem–Proof–Commentary and there is much detail, e.g. the result stated in equation (5.57) is awarded two pages of outline proof followed by five pages of formal proof. In Part III (chapters 15–18) the asymptotic results are applied to some standard statistical methods.

This is not a book for the faint-hearted! It is a thorough, painstaking study of an area of applied probability that underlies important statistical methodology. The authors are respected

researchers in the topic and I am sure that the text will encourage others to join them in their work.

Martin Crowder: *m.crowder@imperial.ac.uk*
Mathematics Department, Imperial College
London SW7 2AZ, UK

Biscuits of Number Theory

Arthur T. Benjamin, Ezra Brown (Editors)
Mathematical Association of America, 2009, xiii+311 pages, US$ 62.50, hardcover
ISBN: 978-0-88385-340-5

Table of contents

Readership: Teachers and students of number theory, and anybody else interested in this field.

This book is a collection of interesting and well-written papers on basic topics of number theory. All papers except one have previously been published, most in *American Mathematical Monthly*. Among the authors there are Euler and other great masters. Nine papers have been awarded.

Jorma Kaarlo Merikoski: *jorma.merikoski@uta.fi*
Department of Mathematics and Statistics
FI-33014 University of Tampere, Finland

Bayesian Reliability

Michael S. Hamada, Alyson G. Wilson, Shane C. Reese, Harry F. Martz
Springer, 2008, xvi + 436 pages, € 69.95 / £ 62.99 / US$ 89.95, hardcover
ISBN: 978-0-387-77948-5

Table of contents

Readership: Reliability practitioners, Bayesian researchers in reliability. The book may also be used as a textbook for a course for advanced undergraduates or graduate students with at least one basic course in Statistics.

This is a very well written Bayesian book on reliability with almost encyclopedic coverage. The first chapter provides an overview of the reliability aspects of the whole book, while the next two chapters introduce readers to basics of Bayesian analysis right up to MCMC and various relevant software. Chapters 4 and 5 cover component and system reliability. A good first course on reliability can be based on these four chapters and parts of Chapter 1. The remaining chapters deal with more advanced and relatively new topics, namely, repairable systems, regression models in reliability, modeling reliability via degradation data, planning and choice of optimal design using a genetic algorithm, and assurance testing. This reviewer hasn't seen the last three topics so well analyzed elsewhere. All these topics are of great practical importance and point to the continuing vitality of the subject.

The strength and pioneering contributions of the book include Bayesian goodness of fit tests and model selection, Bayesian analysis of reliability based on fault trees and Bayesian networks, modeling with covariates, and use of degradation data to anticipate reliability breakdowns. The illustrative examples are taken from real life, and for each example several plausible models are tried out and compared, providing a flavor of robustness. My only criticism, if it can be called a criticism, is that it would have been good to compare the new tests of Valen Johnson with the more standard ones based on different kinds of Bayesian p-values due to Rubin, Gelman, Berger, and Bayarri. A similar remark holds for model selection.

Given the strengths of the book in both coverage and detailed modeling of reliability based on many different kinds of data, real life examples, and a thorough, rigorous, and

insightful Bayesian analysis, the book makes a major contribution to the literature on reliability.

Jayanta K. Ghosh: *ghosh@stat.purdue.edu*
Department of Statistics, Purdue University
West Lafayette, IN 47909, USA

Design and Analysis of Experiments, Volume 1: Introduction to Experimental Design, Second Edition

Klaus Hinkelmann, Oscar Kempthorne
Wiley, 2008, xxiv + 631 pages, € 112.50 / £ 90.50 / US$ 135.00, hardcover
ISBN: 978-0-471-72756-9

Table of contents

Readership: Students in Statistics, graduate students, consulting statisticians and researchers interested in Theory of Experiments, both theoretical and applied.

The book of Hinkelmann and Kempthorne is a very good source of knowledge about designing and analysis of experiments. It makes the theory of experiments easier to understand and it may be useful as a textbook as well as a reference book.

This volume is a great extension of its first edition, especially by emphasizing the practical aspects of designing and analyzing experiments. Chapter 2 is expanded by two new subsections, in which particular steps of planning experiments are emphasized. The book allows practitioners and statisticians to understand each other better as well as to understand the relationship between the quality of experimental design and the validity of conclusions. The most commonly used experimental designs such as, e.g., Randomized Block Designs, Latin Squares, Factorial Experiments, Split-Plot Designs, or Repeated Measurements Designs are described with extended – in relation to the first edition – details on repeated measures.

The advantage of Hinkelmann and Kempthorn's book is the huge amount of numerical examples at the end of most chapters. Authors use the SAS® package, but the way of presenting the results may be easily applied to other statistical software.

The layout of the book, large and rich collection of examples and exercises, and very clear writing make this book highly recommended for students as well as for researchers in wide areas of science including agriculture, engineering, medicine, or economy.

Katarzyna Filipiak: *kasfil@up.poznan.pl*
Department of Mathematical and Statistical Methods
Poznań University of Life Sciences
Wojska Polskiego 28, 60-637 Poznań, Poland

Introduction to Scientific Programming and Simulation Using R
Owen Jones, Robert Maillardet, Andrew Robinson
Chapman & Hall/CRC, 2009, xix + 453 pages, £ 48.99 / US$ 79.95, hardcover
ISBN: 978-1-4200-6872-6

Table of contents

Readership: Undergraduate students of statistics, computer science and engineering; postgraduate students in applied disciplines.

This book is based on a course given by the first two authors at the University of Melbourne, Australia. After many years of developing their course, the book has two principal aims – to teach scientific programming and to introduce stochastic modelling. As a minimum, the authors assume that the reader has completed a first year undergraduate level course on calculus and from such a basis the book succeeds with its aims.

The authors do not describe statistical techniques as implemented in R but show how to convert algorithms into computer-based codes. The chapters within the book are structured to provide a core programme covering basics and discrete random variables; together with chapters covering additional material on continuous random variables and numerical optimisation. Personally, I think that the techniques of scientific programming presented will soon enable the novice to apply statistical models to real-world problems. The writing style is easy to read and the book is suitable for private study. If you have never read a book on scientific programming and simulation then I recommend that you start with this one.

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

Introduction to Spatial Econometrics
James LeSage, Robert Kelley Pace
Chapman & Hall/CRC, 2009, xiii + 354 pages, £ 54.99 / US$ 89.95, hardcover
ISBN: 978-1-4200-6424-7

Table of contents

Readership: Students and researchers with a basic knowledge of linear regression methods.

The scope and topic of this book are in the title, and the authors provide the right balance of theoretical detail and applied illustrations of the methods discussed. The authors' treatment of the subject is challenging of the reader; providing a number of motivations for the phenomenon known as *simultaneous spatial dependence*. The field is rapidly evolving and much of the material reflects recent ideas that have not appeared elsewhere.

The text would be suitable for either an advanced undergraduate or a postgraduate course in spatial econometric modelling. For an undergraduate course, however, be prepared for many student questions and possible challenges to the ideas presented in the text. In some parts of the text, the breadth of coverage comes at the expense of detailed derivations and in the authors' defence, I believe that the text does provide extensive references which will facilitate further study and motivate further research for those wishing to learn more. For those interested in implementing the methods and with knowledge of MATLAB, there is publicly available code which the authors refer to.

The authors' topic is an exciting area of statistics which will clearly evolve further and develop in the coming years. For these reasons alone, the text deserves wider exposure and the effort that will need to be invested in order to fully understand the material presented should be well rewarded.

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

Nonlinear Regression with R
Christian Ritz, Jens Carl Streibig
Springer, 2008, xii + 148 pages, € 39.95 / £ 34.99 / US$ 54.95, softcover
ISBN: 978-0-387-09615-5

Table of contents

Readership: Under- and post-graduate students of statistics and of applied disciplines in biology, chemistry, engineering, fisheries science, medicine and toxicology.

All too often nonlinear regression can be a confined and narrow topic within statistics and thankfully, this is not the case here. The scope and topic of this book are in the title and the authors take as their starting point the function nls() and subsequently, related functions in R. However, do not expect a textbook on nonlinear regression – the authors have been far more adventurous than that! They have successfully demonstrated the use of nonlinear regression from its many and varied applications in applied sciences.

Much is written in the statistics literature about linear regression and many applied statisticians can work successfully for years within this field but there are many like me who rarely see a problem for which linear regression is appropriate.

Developing and devising methodologies and techniques to deal with non-convergence of nonlinear estimation routines; together with the associated diagnostics can be a long, laborious and frustrating process. I was delighted to read in this text that a number of the rules-of-thumb that I and others apply to real-life problems find their place in this book. I strongly recommend this book – if you are a young scientist with limited exposure to nonlinear estimation then this book will save you hours of wasted exploration and investigation to find the allusive solution to your nonlinear estimation problem.

The authors' topic is an exciting area of statistics which will clearly evolve further and develop in the coming years. For these reasons alone, the text deserves wider exposure and the effort that will need to be invested in order to fully understand the material presented should be well rewarded.

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

Statistical Detection and Surveillance of Geographic Clusters
Peter Rogerson, Ikuho Yamada
Chapman & Hall/CRC, 2009, xxiv + 322 pages, £ 57.99 / US$ 89.95, hardcover
ISBN: 978-1-58488-935-9

Table of contents

Readership: Under- and postgraduate students of statistics, computer science and geography; together with practitioners in the public health sector.

A significant portion of this book is based upon the work of the first author which he has carried out over the last decade. Naturally, material from previous papers and book chapters have been used in writing this book and the authors helpfully provide a list of the correspondence between sections of this book, and their previously published papers and books.

The authors have a simply defined purpose for their book – to provide statistical tools for deciding whether the data on a given map deviate significantly from expectations, and deciding quickly whether new map patterns are emerging over time. The latter problem of monitoring geographic data over time has seen a recent surge in interest and this developing field is referred to as *spatial surveillance*; the former is long-established and the authors merely review established methods.

In contrast to most edited books, the authors' text provides a comprehensive presentation that will provide a valuable reference source for years to come. I recommend that anyone working with geographic information systems (GIS) consults this volume, as there are many valuable nuggets of information – many unexpected!

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

Computational Intelligence in Medical Imaging: Techniques and Applications
Gerald Schaefer, Aboul Ella Hassanien, Jianmin Jiang (Editors)
Chapman & Hall/CRC, 2009, xiv + 487 pages, £ 60.99 / US$ 99.95, hardcover
ISBN: 978-1-4200-6059-1

Table of contents

Readership: Post-graduate students of computer science, statistics and medicine.

In choosing this book the reader will be exposed to the range of exciting research that is being conducted in the context of medical imaging. This research utilises computational intelligence techniques employing neural networks, fuzzy logic, Bayesian statistics and Monte Carlo techniques, to mention but a few. From almost 40 proposals, the editors and reviewers have selected 15 contributions (most of which are co-authored) to present. The latest trends and developments in the field of computational intelligence in medical imaging are presented and discussed.

Medical imaging has become a routine tool for many branches of medicine and facilitates the capture, transmission, and analysis of medical images; and so aids in medical diagnoses. This is a growing field of research with many challenges remaining to be resolved. New approaches are emerging and I am sure that this collection of the latest trends and developments will further stimulate discussion and development of new solutions. The book will be of interest and relevance to anyone involved in the computational analysis and interpretation of images – whether medical or not.

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

Applied Multivariate Statistics for the Social Sciences, Fifth Edition
James P. Stevens
Routledge, 2009, xii + 651 pages, £ 40.50 / US$ 72.00, softcover (also available in hardcover)
ISBN: 978-0-8058-5903-4

Table of contents

Readership: Students in psychology, social science, education and business attending courses on multivariate statistics as well as practicing researchers in these areas.

As should be evident from the title, the main audience of this book is those who use advanced statistical methods, rather than those who develop these methods. A prerequisite for students is to have at least two quarter courses in statistics covering factorial ANOVA and covariance, although a two-semester sequence of courses is preferable. No prior knowledge of matrix algebra is necessary, since a chapter covering the basics of matrix algebra is included. An accompanying website contains data sets and answers to even numbered exercises.

The book covers most of the main multivariate statistical methods, but I miss a chapter about cluster analysis. Also, logistic regression is only briefly discussed. Given the importance of this subject, an extended coverage of it had been appreciated.

A nice feature of this book is its pedagogical way of introducing the different multivariate statistical methods. When possible, the author tries to start with the equivalent univariate method and then extends this step-by-step to a fully multivariate method. E.g., MANOVA is presented by first discussing the univariate Student's *t*-test and showing how this test is generalized to the multivariate two-group case of Hotelling's T^2 test. It is then shown how the univariate *k*-group ANOVA can be extended to the multivariate *k*-group MANOVA.

Another feature of this book is its heavy reliance on numerical examples using SPSS and SAS, providing both program code to run the examples and detailed outputs. Both the outputs and the program code are annotated and relevant parts are commented on. This makes it easy to interpret the outputs and understand what the numbers mean, as well as what the program code does.

Andreas Rosenblad: *andreas.rosenblad@ltv.se*
Center for Clinical Research Västerås, Uppsala University
Central Hospital, Västerås, Sweden

Statistics Applied to Clinical Trials, Fourth Edition
Ton J. Cleophas, Aeilko H. Zwinderman, Toine F. Cleophas, Eugene P. Cleophas
Springer, 2009, xxii + 559 pages, € 74.95 / £ 67.99 / US$ 119.00, hardcover
ISBN: 978-1-4020-9522-1

Table of contents

Readership: Students, physicians and investigators interested in statistical methods for clinical trials.

This book was originally written for a course in medical statistics given in the EU sponsored program European Interuniversity Diploma of Pharmaceutical Medicine starting in the year 2000. Since then it has been expanded and updated in order to serve as a guide and reference-text to students, physicians and investigators.

Frankly, this is the worst book on statistics that I have ever seen, and it is unbelievable that it has been printed by such a well-reputed publisher as Springer. It contains countless factual and grammatical errors. This starts already in the first chapter, where the authors state, e.g., (p. 5) that the mean ± 2 SDs is called "the 95% confidence interval of the data, which means that 95% of the data of the sample are within" and that statistics "test the null hypotheses of finding no difference from zero in your sample" (p. 5–6).

Formulas are inconsistent and hard to follow. Multiplications may be written as a*x*b, a*x*b, a.b, a·b or a × b, and it is sometimes not clear if x is used as a variable or as a multiplication sign.

Sometimes the authors talk about chi-square tests, and sometimes about χ^2 tests. The graphs are a mess, small sized with bad resolution and texts that are hard to read. All of a sudden there may be half a page empty without any reason. Names are misspelled, e.g. "Kruskall-Wallis test", "Markow models" and "Mantel-Haenszl". References are sometimes non-existing. E.g., on p. 404 is a reference to "Young 1948", which does not appear in the reference list, and a footnote referring to a website with an incorrectly given address. When finally locating the website, the reference cannot be found. Additionally, the correct reference would have been "Young 1941".

Andreas Rosenblad: *andreas.rosenblad@ltv.se*
Center for Clinical Research Västerås, Uppsala University
Central Hospital, Västerås, Sweden

Introduction to Meta-Analysis

Michael Borenstein, Larry V. Hedges, Julian P.T. Higgins, Hannah R. Rothstein
Wiley, 2009, xxx + 421 pages, € 42.00 / £ 34.95 / US$ 70.00, hardcover
ISBN: 978-0-470-05724-7

Table of contents

Readership: Researchers, clinicians and statisticians from the fields of medicine, epidemiology, social science, business, ecology, and others interested in meta-analysis.

The purpose of this book is to provide the reader with a framework that explains the logic of meta-analysis and how to apply and interpret meta-analytic procedures in a proper way. It is also intended to be useful for a course in meta-analysis. Supplementary material, exercises and data files are said to be available on an accompanying website. Additionally, the authors have developed software called Comprehensive Meta-Analysis, with a free trial version available from the website. However, when trying to install this software the installation procedure fails, and no other supplementary material, exercises or data files are found on the website.

The book has a statistical approach, providing all the necessary formulas along with worked examples, but the authors state that the reader should be able to skip the formulas and still understand the difference between two analysis methods and the mechanism behind a certain statistical method. While this may be true, to fully benefit from this text the reader should have a decent understanding of basic statistical methods and be able to read and understand the given formulas. Further, the focus of the text is on meta-analyses of effect sizes, while other approaches are only discussed briefly. For example, meta-analyses of diagnostic tests are not discussed at all.

The layout of the text is pedagogical, with each separate subject discussed in a separate chapter. The standard layout of a chapter is to start with an introduction to the subject that will be discussed, followed by several sections discussing the various aspects of the subject, and ending with a summary of the most important points to be learned from the chapter. The book is also richly illustrated with useful graphs.

Andreas Rosenblad: *andreas.rosenblad@ltv.se*
Center for Clinical Research Västerås, Uppsala University
Central Hospital, Västerås, Sweden

Model Based Inference in the Life Sciences: A Primer on Evidence
David R. Anderson
Springer, 2008, xxiv + 184 pages, € 32.95 / £ 29.99 / US$ 39.95, softcover
ISBN: 978-0-387-74073-7

Table of contents

Readership: Researchers and graduate students in ecology and other life sciences.

This monograph expounds ideas that the author has developed over many years with Burnham. It is heavily example-based, and aimed at working scientists. Examples are predominately from ecological studies.

Chapters 1 and 2 set out the author's general philosophy of inference, leaning heavily on the ideas of the geologist Thomas C. Chamberlin. Rather than using a study to test a single hypothesis, Anderson argues for starting with several carefully considered hypotheses in mind. This protects against allowing prior prejudices and understandings undue influence in the interpretation of results. Furthermore, "the main purpose of an experiment is to estimate the size of the effect on a response variable of interest caused by the treatment". Experiments are contrasted with descriptive studies and confirmatory studies. Anderson argues that too much emphasis has been placed on descriptive studies, at the expense of experiments (where these are possible) and confirmatory studies.

Chapter 3 and 4 argue for the use of the "corrected" AICc version of Akaike's AIC measure for model inference. The use of such a statistic to choose between a limited number of scientifically plausible models is relatively uncontroversial. More controversial is the use of AICc to quantify the relative evidence for the competing models. The AICc for the "best" model is subtracted from each model's AICc, yielding differences Δ_i. The quantities $\exp(-1/2\Delta_i)$ are then divided by their sum, yielding weights that Anderson terms "likelihoods" that are used in a model averaging procedure. Model averaging approaches offer a mechanism for incorporating model uncertainty into standard errors of model parameters and predictions.

Notwithstanding the comment that this likelihood "is very different from the usual one used in parameter estimation", the terminology is confusing, Why not call it a "penalized likelihood"? The "usual" likelihood appears in at least two places elsewhere in the book.

Anderson argues that the relative weight has a Bayesian interpretation as a posterior probability, providing a "savvy" prior is used. With such a prior, the model averaging approach that is advocated in Chapter 5 becomes a form of Bayesian model averaging. Thus, what is a savvy prior? Anderson does not explain, and it is necessary to look elsewhere; see the reference below. Basically, it is chosen so that the relative weight, defined as above, becomes a Bayesian posterior probability! The savvy prior has the same precision as the likelihood.

Chapter 6 has a brief discussion of multilevel models. For models for counts, Anderson seems to prefer use of a correction for overdispersion to modeling of the error components. The claim that model choice and averaging ideas carry over to "random effects" models is made too lightly. The use of AIC-like statistics in a multi-level model context is not (or should not be) a straightforward extension of methods that are applicable for models with independent errors.

There is an emphasis on the importance of careful selection of the prior set of hypotheses. The choice should evolve, up to the time of conduct of the study. Appendix F, at the end of the book, lists "Common Misuses and Misinterpretations". Note in particular the warning against "rampant data dredging".

This is an interesting and challenging, albeit idiosyncratic, book.

John H. Maindonald: *john.maindonald@anu.edu.au*
Centre for Mathematics & Its Applications
Australian National University
Canberra ACT 0200, Australia

Applied Statistics for Engineers and Physical Scientists, Third Edition
Johannes Ledolter, Robert V. Hogg
Prentice Hall, 2009, xv + 591 pages, US$ 134.67, hardcover
ISBN: 978-0-13-601798-1

Table of contents

Readership: Undergraduate and graduate level engineers and scientists who want an introduction to applied statistics, applied statisticians.

This third edition of *Applied Statistics for Engineers and Physical Scientists*, written by two leading statisticians seventeen years after the second edition, brings the book up to date with modern engineering statistics. The first edition with the title *Engineering Statistics* was published by Macmillan in 1987. The book is intended to be practical and it emphasizes the actual application of statistical methodology to realistic engineering settings. The authors encourage students to try several available programmes in analysing their data. Much of the discussion in this book is tied to Minitab, although SAS, SPSS, and R can also be used. The R package is very general and somewhat more difficult to learn, because it requires the user to execute R-specific instructions. However, it is easy for engineering students who are used to writing programs in Matlab. Instructions to use Minitab and R are included in the instructions manual available online.

Chapters 2–4 cover discrete probability, continuous probability, and applications of sampling distribution theory. The last four chapters cover useful stuff for engineers and scientists: statistical process control, single and multifactor experiments, and response surface methods. This perfect engineering statistics textbook does not provide only the statistical tools but also insights to the underlying statistics. Such a text allows the instructor to make necessary compromises. The book can be used for a strong two-semester sequence in applied statistics, and it can also be used for various one-semester courses depending on the interest of the students and their instructor.

Erkki P. Liski: *erkki.liski@uta.fi*
Department of Mathematics and Statistics
FI-33014 University of Tampere, Finland

The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition

Trevor Hastie, Robert Tibshirani, Jerome Friedman
Springer, 2009, xxii + 746 pages, £ 62.99 / € 69.95 / US$ 89.95, hardcover
ISBN: 978-0-387-84857-0

Table of contents

Readership: Everyone interested in statistical learning.

The first edition of this book from 2001 might be considered already as a modern classic on statistical learning. This second edition pays tribute to the many developments in recent years in this field, and new material was added to several existing chapters as well as four new chapters (15–18) were included. Especially the chapter on high-dimensional problems ($N \ll p$) is a welcome and important addition. These additions make this book worthwhile to obtain also for readers who already have the first edition.

In general the book discusses modern methods of supervised and unsupervised learning where the focus lies definitely on the former one. The material usually motivates and explains the ideas of the different methods without going too much into mathematical detail and outlines algorithms for practical implementation. Then the methods are applied to interesting real world problems coming from various fields. Seeing how many of the methods work on the different examples makes one often wish to play around a bit with the methods or to apply them immediately to one's own data. Here is then nice that several of the data sets of the book are available as an R package, and in many places the book refers to R packages which could be used for the corresponding methods. The R package ProDenICA, supposedly written by two of the authors of the book, especially raised my interest, but seems, however, not to be on CRAN as claimed in the book.

In general this is a well written book which gives a good overview on statistical learning and can be recommended to everyone interested in this field. The book is so comprehensive that it offers material for several courses.

Klaus Nordhausen: *klaus.nordhausen@uta.fi*
Tampere School of Public Health
FI-33014 University of Tampere, Finland