Short Book Reviews

Editor: Simo Puntanen

Statistics in Psychology Using R and SPSS
Dieter Rasch, Klaus Kubinger, Takuya Yanagida
Wiley, 2011, xii + 552 pages, £55.00/€66.00/$90.00, hardcover
ISBN: 978-0-470-97124-6

Table of Contents

Readership: Students and teachers of behavioural sciences, especially psychology, on multiple study levels (Bachelor's, Master's, doctoral).

If you should choose one textbook of applied statistical methods that would include relevant content for multiple study levels, from Bachelor's studies to Master's and even doctoral studies, this book could be a very good candidate. It has a broad coverage with a special structure for multiple target groups. It also shows in detail how to perform the various statistical analyses by two different statistical software programs, R and SPSS.

The three target groups mentioned above are pointed out by special labels (Bachelor, Master, Doctor) throughout the text, in order to support the students' navigation across the book. Although this type of classification is necessarily approximate, I feel that it works well with this book. Indeed, quite often the labels Master and Doctor seem to appear together, as it is more essential to draw some lines between those two and the Bachelor group.

However, it is not meant that the Bachelor's students should focus only on the Bachelor content and so on. The book is much more versatile. At least for the more advanced students, it may be quite stimulating and motivating to be able to get a taste of the more demanding contents as well. This is an important point, because it encourages the students to continue their

studies further. At the same time, it passes on a clear message that the Bachelor level is usually not enough for understanding the more advanced methods.

Occasionally, the other direction may be useful, too. For a Master's student, the Bachelor labels offer 'checkpoints' of contents that should be mastered already. In addition to the three labels, there are didactic notes for the assumed lecturers. These notes offer background information, explanations and viewpoints, notes of traditions, notations, references, persons and so on. These contents may be useful, at least for inexperienced teachers.

Hence the book certainly has a multi-layered structure that makes it rather special in its genre. It covers the basic issues of measurement and study design, and then approaches to univariate and multivariate data analysis and statistical inference, introducing many statistical methods and models relevant to psychology.

And there is even one more layer: the software. Both R and SPSS are used and their input and output are explained in detail throughout the book. The software specific contents are marked clearly, which makes it easy to follow either one of these or make comparisons between them.

Kimmo Vehkalahti: *kimmo.vehkalahti@helsinki.fi*
Department of Social Research, Statistics
FI-00014 University of Helsinki, Finland

New Perspectives in Statistical Modeling and Data Analysis: Proceedings of the 7th Conference of the Classification and Data Analysis Group of the Italian Statistical Society, Catania, September 9–11, 2009
Salvatore Ingrassia, Roberto Rocci, Maurizio Vichi (Editors)
Springer, 2011, xxii + 587 pages, £90.00/€106.95/$129.00, softcover
ISBN: 978-3-642-11362-8

Table of Contents (parts)

Readership: Researchers interested in the methodological and applied aspects of statistical modelling, data analysis and classification.

This volume contains contributions from the seventh biannual meeting of the Classification and Data Analysis Group of the Italian Statistical Society, organised by the Faculty of Economics at the University of Catania in September 2009.

The volume is divided into 12 parts that include both theoretical and applied parts. The main topics of the contributions are in the area of data analysis (including categorical data) and modelling (with applications in evaluation, economics, environmental and medical sciences industry

and services); classification and statistical methodology (functional data analysis, kernel estimation, multivariate analysis). Moreover, some contributions focus on computational aspects of statistics (visualisation of relationships, computer intensive methods).

Most of the contributions provide a clear overview of current areas of active research in statistics. As such, they can help the reader to have an idea about the current state-of-art in different topics.

The volume can be used primarily as reference material.

Fabrizio Durante: *fabrizio.durante@unibz.it*
School of Economics and Management
Free University of Bozen-Bolzano
Piazza Università 1, 39100 Bolzano, Italy

Regression Models as a Tool in Medical Research
Werner Vach
Chapman & Hall/CRC, 2013, xxi + 473 pages, £57.99/$89.95, hardcover
ISBN: 978-1-46-651748-6

Table of Contents

Readership: The book can be a very helpful contribution especially for researchers in medical sciences when performing their statistical analyses and trying to interpret the results obtained.

Such models as linear regression, logistic regression and Cox proportional hazards model are standard tools in medical research. Unfortunately, the application of these tools rises a number

of issues that can be quite difficult especially for inexperienced users. This book provides a plenty of practical knowledge about these basic models and also some of their extensions, that is often not easy to find from statistical textbooks or from software manuals. The basic methods are well explained and illustrated by numerous practical examples, mainly using simulated datasets. Although the situation is often more complicated in practice, the solution is pedagogically founded. Example data sets are analysed using Stata software. Although Stata is certainly a very good choice for these analyses, the use of non-commercial software (such as R) might have been more useful.

Tapio Nummi: *tapio.nummi@uta.fi*
School of Health Sciences
FI-33014 University of Tampere, Finland

Principles of Business Forecasting
Keith Ord, Robert Fildes
Cengage Learning, 2013, xxii + 506 pages, $142.49, hardcover
ISBN: 978-0-324-31127-3

Table of Contents

1. Forecasting, the why and how
2. Basic tools for forecasting
3. Forecasting trends: exponential smoothing
4. Seasonal series: forecasting and decomposition
5. State-space models for time series
6. Autoregressive integrated moving average (ARIMA) models
7. Simple linear regression for forecasting
8. Multiple regression for time series
9. Model building
10. Advanced methods of forecasting
11. Judgment-based forecasts
12. Putting forecasting methods to work
13. Forecasting in practice

Readership: The target audience is business people who have time constraints but are interested in getting a new overview of existing methods.

How would you write a compendium book about forecasting for people who are not directly interested in statistics but want to know what technologies are currently available? I recommend the style of Ord and Fildes. It took me many years to appreciate the notion that deriving statistics formulas for models is one thing and making the models and techniques understandable for a large audience is another thing.

The title of the book, Principles of Business Forecasting, really hits the two goals of the book: Principles are explained in first place - and not formulas or newest statistical models that led to the forecasting procedures - and the second target is the audience, business people who have notorious time constraints but are interested in getting a new overview of existing methods. I think that both goals are really nicely reached.

Let me point out some highlights:

- An emphasis on incorporating the latest research in forecasting methodology helps students prepare for a managerial or analytic career and to help current practitioners make more effective forecasts.
- The authors base their approach to forecasting on a series of stages: purpose, horizon, information, value, and evaluation.

- The organisation of chapters is flexible and designed for students or current practitioners: introduction (Chapters 1–2), extrapolative methods (Chapters 3–6), statistical model building (Chapters 7–9), and advanced methods and forecasting practice (Chapters 10–13).
- The authors illustrate forecasting techniques in a variety of software environments, including Excel®, Minitab®, and SPSS, so that users can apply the methodology that best fits their available resources.
- Discussion questions are placed throughout each chapter for students to think beyond technical issues.
- End-of-chapter principles reinforce a consistent foundation of good forecasting practices.
- End-of-chapter exercises help students consider the role of the practising forecaster.
- Students and instructors can conveniently access the text data sets on their companion sites.

There is a quite common saying among people who work in the area of statistical modelling: The student does not see the forest for the trees because of the many trees. The business of forecasting can be a nice stroll through the woods, and if you got to meet the right ranger of this national park it might be really enjoyable.

The first book that tried to do this type of guided ranger tour through the forest of econometrics was *A Guide to Econometrics* by Peter Kennedy (MIT Press, 2003). The students love this book and it has seen many editions since it came out the first time. It addresses the basic questions: What can be measured in empirical economics? What makes economics so difficult to be a simple exercise of statistical modelling? What is the econometrics on top of simple statistical applications? And finally: Is it worthwhile to go this long way and have we understood more of the empirical and statistical relationships after we have studied the involved models in detail and tried to understand the mathematical technicalities?

The traditional approach is the hard one: First study mathematics and statistics and at the end of the course I show you how to put together simple models. Later making the remark: For the real serious models you have to take an advanced course that will show you what powerful models can really do. But when you are there you are probably not interested anymore. So, what was it good for?

Who should read or use the book? Clearly people studying business administration. They are also used to these high prices of text books in this field. There is the possibility to use the Cengage software along with the book, but I could not even open a learning example, unfortunately. Finally, researchers in new methods might get some ideas where to look for interesting challenges. Given the prices involved, I recommend to read the book in the library.

Wolfgang Polasek: *polasek@ihs.ac.at*
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56, 1060 Vienna, Austria

Multivariate Analysis for the Biobehavioral and Social Sciences: A Graphical Approach
Bruce L. Brown, Suzanne B. Hendrix, Dawson W. Hedges, Timothy B. Smith
Wiley, 2012, xiv + 475 pages, £66.95/€80.40/$104.00, hardcover
ISBN: 978-0-470-53756-5

Table of Contents

Readership: This book is oriented to researchers in natural and biological sciences and work in the area of multivariate statistics. But most of the content is certainly not suitable for beginners in these fields.

The title of the book is very promising and who would not like to learn about graphical approach in bio-behavioural and social sciences. But to my disappointment, I found all known graphical techniques described that can be found in modern multivariate statistics books.

Chapter 5 gives some insight to the philosophy of this 'newly proposed' statistical approach to statistical graphics. On page 230 the authors come up with 'Desirable characteristics in multivariate graphics (MG)'. For readability I have shortened all points.

1. MG should not be in any way misleading.
2. MG should be aesthetically pleasing.
3. MG should take a Gestalt approach to variables.
4. MG should take a Gestalt approach to individual observations.
5. . . . .
8. MG should be realistic, let us understand the world better (like the Minard graph).

I have created this list to show how difficult it is to come up with a theory of graphing. As in any axiom system (see e.g. Peano axioms for integer numbers) the explicandum 'MG' (multivariate graphics) will be explained by a set of properties that are essentially undefined. In our case: misleading, pleasing, Gestalt approach, and realistic. As one can easily imagine that it might be very time consuming to come up with any sort of agreement.

All books on statistical graphics cite the Minard graph and many attempts have been made to come up with a similar graph. But this is a 'statistical dream', because here are only a few high-dimensional relationships that can be transformed into an interesting graph. Thus, the Minard graph is rather harmful since it distracts people to create their own graphs. It also shows that some interesting features of the graph can only be obtained from extra information that is not revealed in the graph: The death of French soldiers was due to the rising (!) temperatures that melted the frozen rivers and not because of falling winter temperatures as many people think.

The authors have intended to write a book on statistics that gives some idea on nice graphical displays of the used methods. Unfortunately this approach requires the knowledge of a lot of basic statistical ideas, so the book follows almost entirely other textbooks on multivariate statistics. Only a few examples or hints try to follow own ideas. This might not always be very illuminating, like the comparison of male names in the 19th and 20th century (Chapter 5, p. 233 ff.)

The book is rather silent on the software that should produce the graphs. The examples of the book are based on the SAS program, which is unfortunately a software package that is expensive to buy. A library in R would have been more useful to distribute some plots or results.

I would say that this is not a text book for studying statistics by nice pictures but a book that introduces you into the spirit of statistics with a minimum of formulas. This approach was partially successful but the future will show if this road leads to more lovers of statistics. By the way, this will certainly become a challenge in statistics for the future: Is it possible to convey statistical ideas without going all the way through the whole way of boring linear algebra and linear models? If applied statistics will not solve this challenge soon enough, then the next step of developing more attractive applied methods will come from computer science. As we have seen in the area of neural networks, applied researchers will find the language of computer science simpler and more direct without going through the danger of using the right assumptions, like independence, distributions, or functional form.

Some research oriented people might get also some ideas as how to extend the proposed methods. The authors started an interesting approach but soon had to realise that a genuine graphical approach to statistics might be highly desirable but cannot be realised with the current state of the art of statistical and computational skills and software.

Wolfgang Polasek: *polasek@ihs.ac.at*
Department of Economics and Finance
Institute for Advanced Studies
Stumpergasse 56, 1060 Vienna, Austria

Editor's note: Wolfgang Polasek passed away the day after he submitted the previous two reviews.

An Accidental Statistician: The Life and Memories of George E. P. Box
George E. P. Box
Wiley, 2013, xxiv + 275 pages, £23.50/€28.20/$34.95, hardcover
ISBN: 978-1-118-40088-3

Table of Contents

Readership: Biographical fans of the life story of a very famous statistician.

Joining the British Army when World War 2 broke out in 1939 would not, at first sight, appear to provide a suitable initial career path for a man who would become one of the

world's leading and most famous statisticians. Yet that is exactly what happened to George Box. After becoming involved with army experiments to combat the effects of poison gas, he realised the need to understand the principles of good experimental design in order to determine accurate ways to improve product development. He requested permission from army authorities to visit R. A. Fisher for statistical guidance. Since the army had no standard procedure for a sergeant to visit a professor, he was given a warrant that said he was taking a horse to Cambridge! After the war, he earned a Ph.D. at London University, studying under E. S. Pearson and H. O. Hartley. This led on to a position as a statistician in the Manchester Dyestuffs Division of Imperial Chemical Industries, a major chemical company of the time in the United Kingdom. He spent a year visiting North Carolina State College at Raleigh, where he supervised the thesis of his later co-author J. S. (Stu) Hunter, and then to a wider study of statistics in the US academic worlds of Princeton University and the University of Wisconsin. George rose rapidly to fame throughout the statistical world for his amazing advancement of several important areas of statistical theory and practice, including experimental design and time series analysis. He was a continuing source of inspiration to a large number of former graduate students, to his colleagues and his co-authors, and to many others worldwide. His weekly 'Beer and Statistics' seminars at his home, featuring a variety of informal speakers with real data problems, and followed up by discussion and suggestions for how best to proceed, remain a fond memory of many University of Wisconsin students and faculty. The students swore they learned more in the seminars than they ever did in class! In this superb book, George tells his fascinating story with his own wry and inimitable humour.

   George Box died on Thursday March 28, 2013 at age 94, just as his book became available. It is a fitting tribute to a superb statistician and popular man. His memorial service was attended by many of his former graduate students. Many of them brought copies of the book to the service in silent tribute.

Norman R. Draper: *draper@stat.wisc.edu*
Department of Statistics, University of Wisconsin–Madison
1300 University Avenue, Madison, WI 53706-1532, USA

Physiology, Biophysics, and Biomedical Engineering
Andrew W. Wood with contributions by Anthony Bartel, Peter Cadusch, Joseph Ciorciari, David Crewther, Per Line, John Patterson, Mark Schier, and Bruce Thompson
Taylor & Francis/CRC, 2012, xix + 762 pages, £49.99/$99.95, hardcover
ISBN: 978-1-42-006513-8

Table of Contents

Readership: Suitable for biomedical engineers, statisticians interested in novel engineering applications to medical and biological studies of the human body.

This book is in the series Medical Physics and Biomedical Engineering, blending the three topics listed in the title of the book - very, very different from all other books that I have reviewed in the past and more recently. At least partly this may be because biomedical engineering is a

relative late comer and, by its very nature, is very interdisciplinary, drawing on physics, chemistry, physiology, biology (including molecules, cells and proteins), and engineering, including transport of material into and out of cells, and cell to cell signalling.

Perhaps someday it will become a part of systems biology or at least overlap with it a lot, but not so yet. Currently, the other hard sciences and engineering dominate, though the goal is to model biological and medical phenomena in a very explicit manner. This reviewer is fascinated by the effort but is not convinced that biology is yet ripe for such applications. My reasons for this view are explained below, they are not meant to be negative. In fact, I am very positive about the subject. Let me explain what I mean. The authors restrict themselves to those biological phenomena that can be described by the physical phenomena of electricity, pressure, and flow.

The goals of the subject are already very ambitious and are probably being realised in the medical sciences. This emerging subject is already contributing to monitoring of cardiac functions (via imaging and otherwise), heart–lung machines, recording and interpretation of electrical signals in the human body, with possible applications and feedback not only to advances of medical science but also to building of artificial organs.

My only fear is that biology may not yet or ever be suitable for such deterministic methods. The most important current field of research in biology is systems biology, where we have high dimensional modelling problems with lots of uncertainty. It does seem such uncertainty and complexity will always remain part of biology. Hopefully as biomedical engineering progresses and advances take place in signal processing, both in the brain and other parts of the human body, uncertainty and feedback control will be part of the paradigm.

The book is strongly recommended to statisticians who wish to work in this relatively new, still emerging field of engineering, which will have to draw on both biology and statistics in the future.

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

Handbook of Statistical Bioinformatics
Henry Horng-Shing Lu, Bernhard Schölkopf, Hongyu Zhao (Editors)
Springer, 2011, ix + 627 pages, € 266.43/£224.50/$339.00, hardcover
ISBN: 978-3-642-16344-9

Table of Contents

Part II. Expression Data Analysis

8. Experimental design and ANOVA for microarray data (*Richard E. Kennedy, Xiangqin Cui*)

9. The MicroArray Quality Control (MAQC) project and cross-platform analysis of microarray data (*Zhining Wen, Zhenqiang Su, Jie Liu, Baitang Ning, Lei Gou, Weida Tong, Leming Shi*)

10. A survey of classification techniques for microarray data analysis (*Wai-Ki Yip, Samir B. Amin, Cheng Li*)

11. Statistical analysis of single nucleotide polymorphism microarrays in cancer studies (*Pierre Neuvial, Henrik Bengtsson, Terence P. Speed*)

12. Computational analysis of ChIP-chip data (*Hongkai Ji*)

13. eQTL mapping for functional classes of *Saccharomyces cerevisiae* genes with multivariate sparse partial least squares regression (*Dongjun Chung, Sündüz Keleş*)

14. Statistical analysis of time course microarray data (*Lingyan Ruan, Ming Yuan*)

Part III. Systems Biology

15. Kernel methods in bioinformatics (*Karsten M. Borgwardt*)

16. Graph classification methods in chemoinformatics (*Koji Tsuda*)

17. Hidden Markov random field models for network-based analysis of genomic data (*Hongzhe Li*)

18. Review of weighted gene coexpression network analysis (*Tova Fuller, Peter Langfelder, Angela Presson, Steve Horvath*)

19. Liquid association and related ideas in quantifying changes in correlation (*Ker-Chau Li*)

20. Boolean networks (*Tung-Hung Chueh, Henry Horng-Shing Lu*)

21. Protein interaction networks: protein domain interaction and protein function prediction (*Yanjun Qi, William Stafford Noble*)

22. Reverse engineering of gene regulation networks with an application to the DREAM4 *in silico* network challenge (*Hyonho Chun, Jia Kang, Xianghua Zhang, Minghua Deng, Haisu Ma, Hongyu Zhao*)

23. Inferring signalling and gene regulatory network from genetic and genomic information (*Zhidong Tu, Jun Zhu, Fengzhu Sun*)

24. Computational drug target pathway discovery: a Bayesian network approach (*Seiya Imoto, Yoshinori Tamada, Hiromitsu Araki, Satoru Miyano*)

25. Cancer systems biology (*Elana J. Fertig, Ludmila V. Danilova, Michael F. Ochs*)

26. Comparative genomics (*Xuhua Xia*)

27. Robust control of immune systems under noise: stochastic game approach (*Bor-Sen Chen, Chia-Hung Chang, Yung-Jen Chuang*)

Readership: This collection provides a marvellous feast to workers in this field. Recommended strongly for all research workers in bioinformatics.

The editors of the volume observe that diverse types of high dimensional throughput data are being generated by biotechnology, revealing different aspects of biological processes at the genomic level. Statistical bioinformatics is a study of these data using the best modern high dimensional statistical methods, many of which were discovered specifically for these problems.

Very broadly speaking there are three major themes, namely, sequence analysis, analysis of gene expression data, and systems biology. Systems biology studies problems of different types of networks to throw light on major problems like cancer systems biology, control of immune systems under noise, and comparative genomics.

Both microarrays and their modern version, where the RNA is replaced by short sequences from the DNA, offer approximation to the study of many genes simultaneously. Indeed it is in these areas and variable selection that statistical analysis has played a major role. Statistics and probability have also played a major role in understanding evolution and population genetics. A more recent sequencing related problem is to assess how accurate are specific short segments of the DNA used as genes.

We discuss in slightly more detail one or two examples from each of these three broad areas, which are called Parts. An article within a Part is called a Chapter.

We discuss Chapters 5 and 7 from Part 1, Chapter 11 from Part 2, and Chapters 19 and 24 from Part 3, to give a flavour of the diversity and depth of the articles in the collection. It is interesting that the authors are truly international, they work in prestigious institutions in Beijing, Berkeley, Los Angeles, Tokyo, Tempe, University Park, Columbia etc.

In Chapter 5, Zheng et al. study how variable selection may be made in the context of genes. The object is not just to make prediction but progress towards leaning pathways. The method is pioneered by Lo and Zheng, and a very general theory covering a broad area was constructed under the leadership of Chernoff.

In Chapter 7, Chen, Li, Rosenberg, and Lindsay survey the two major approaches (or families of approaches) for phylogenetic trees for modelling of evolutionary trees of populations. Then they describe a new method, called mixture trees, which 'produced a tree very similar to established methods' through some form of clustering.

Chapter 11 is a study of SNPs, which are positions on a genome showing substantial variation across individuals and hence are very useful for genetic studies. They have been used as markers for many epidemiological studies. The SNPs have also proved very important in cancer studies because of their copy numbers, which measure the loss of the aforementioned heterozygosity due to cancer cells and hence indirectly measure the abundance of cancer cells. The chapter, written by Neuvial, Bengtsson, and Speed provides an excellent introduction to copy numbers, how to calculate them from SNP array data, SNP technologies, and how to use them in cancer studies.

In Chapter 19, Ker-Chau Li provides an innovative measure of association, which is very different from visual statistical correlation because it depends on the state of a third variable. Genes in genetic pathways will often be correlated in this way. Li gives an interesting example. In one of his lectures at Purdue on this topic, he used Stein's inequality!

In Chapter 24, Imoto, Tamada, Araki, and Miyano, all from the Institute of Medical Science, Tokyo, provide models for gene networks based on microarray gene expression data, which can be either of two types, namely, gene knock down or time course. They combined nonparametric regression and Bayesian networks to model nonlinear relations between genes. They also selected network structure through a kind of BIC (Bayesian Information Criterion).

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

Stationary Stochastic Processes: Theory and Applications
Georg Lindgren
Chapman & Hall/CRC, 2013, xxvii + 347 pages, £57.99/$89.95, hardcover
ISBN: 978-1-4665-5779-6

Table of Contents

Readership: Graduate students and researchers interested in stochastic processes.

This is a very nice book about second order stationary stochastic processes for readers who have already some knowledge about stochastic processes and want to deepen and extend their knowledge in this field. The book focuses mainly on one-dimensional processes but discusses also vector-valued processes and random fields. The different topics are usually treated quite theoretical and several proofs are provided in great detail. For all readers to be able to follow the math also a nice appendix about probability theory is part of the book. Nevertheless the book is not too dry and theory is always motivated by interesting examples coming from engineering and physical sciences. This book is in my opinion also a very good basis for an advanced course on this topic as it is logically structured and as each chapter also contains many exercises.

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

Ensemble Methods: Foundations and Algorithms
Zhi-Hua Zhou
Chapman & Hall/CRC, 2012, xiv + 222 pages, £49.99/$79.95, hardcover
ISBN: 978-1-4398-3003-1

Table of Contents

1. Introduction
2. Boosting
3. Bagging
4. Combination methods
5. Diversity
6. Ensemble pruning
7. Clustering ensembles
8. Advanced topics

Readership: For researchers and students not yet familiar with ensemble methods but interested in them.

This book assumes that the reader is already familiar with basic statistical learning methods and discusses then how to combine single learners to an ensemble which usually performs better. While the book is rather written for a machine learning and pattern recognition audience the terminology is well explained and therefore also easily understandable for readers from other areas. In general the book is well structured and written and presents nicely the different ideas and approaches for combining single learners as well as their strengths and limitations. Focusing on the ideas however means also, that no method is discussed in great detail but usually many references as well as further reading suggestions are provided instead. Therefore, I think this book is suitable for readers who want to get an overview over the ideas behind the different ensemble methods but maybe not for readers who want to study a specific method in depth. While I consider this book very informative I missed something like a case study which would describe for some real data which single learners should be combined in which way.

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

Introduction to Hierarchical Bayesian Modeling for Ecological Data
Éric Parent, Étienne Rivot
Chapman & Hall/CRC, 2013, xxi + 405 pages, £57.99/$89.95, hardcover
ISBN: 978-1-58488-919-9

Table of Contents

Readership: Biostatisticians, environmental scientists, statisticians.

This book will either delight you or frustrate you, as the ecological data referred to in the title merely includes applications in quantitative fisheries science. A better title might have been: *Introduction to hierarchical Bayesian modelling with fisheries applications.*

This noted, over the past few years, fisheries science has proved incredibly productive for the development of statistical modelling applications, with the increased emphasis on risk-based approaches to management and the explicit incorporation of uncertainty in the setting of fishing quotas.

The authors believe that hierarchical Bayesian modelling is like a *golden key* to be given to practising scientists - providing a tool (i.e. *access*) that frees their creativity in designing statistical models. The book is well-written and easy to read, and the material presented deserves a greater exposure in taught statistics courses. I thoroughly recommend the book and believe that the statistical techniques and their application to quantitative fisheries science could ideally complement a short undergraduate course in applied statistics.

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

Statistics for Bioengineering Sciences: With MATLAB and WinBUGS Support
Brani Vidakovic
Springer, 2011, xvi + 753 pages, €96.25/£81.00/$129.00, hardcover
ISBN: 978-1-4614-0393-7

Table of Contents

Readership: Students in engineering, bioengineering and biomedical engineering.

This text has resulted from the author's teaching of introductory statistics to engineering students in the USA. Dealing both with the theoretical aspects of statistical methods and the need to implement software that engineers are familiar with, the book is a delight to read.

The text is well-written for the modern day student and uses colour coded text (blue type face for MATLAB codes and outputs, red-brown for WinBUGS programs) and icons (one icon for data sets, one for MATLAB codes, and one for WinBUGS code). Each chapter starts with a *What is covered in this chapter* and ends with *Exercises* and *References*. Of interest to historians, the author has included photographs throughout the text of leading statisticians and mathematicians who have shaped modern day statistics. I recommend the book to any one intending to use either MATLAB or/and WinBUGS for statistical modelling and analysis.

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

Biostatistics with R: An Introduction to Statistics Through Biological Data
Babak Shahbaba
Springer, 2012, xvi + 352 pages, €58.80/£49.99/$64.95, softcover
ISBN: 978-1-4614-1301-1

Table of Contents

Readership: Undergraduate students in biostatistics, and natural and physical sciences.

The plethora of books on using R is difficult to keep up with and the market may be close to saturation point. This book is, however, unique and distinct from others in the field that it presents R-Commander - a useful tool for teaching basic statistics to students who do not have a computer programming background. The Appendix A provides the simple instructions for its installation on a computer and introduces the reader to the R-Commander window.

The author presents statistical methods by their direct application to numerous past scientific studies that are publicly available. In addition, the various data sets are available online and can be downloaded which makes the text ideal as a teaching aid.

A balance has been struck between mathematical rigour and readability; with more complex concepts that are of limited application in everyday analyses relegated to the end of each chapter in *Advanced* sections.

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

Programming in C++ for Engineering and Science
Larry Nyhoff
Chapman & Hall/CRC, 2012, xiii + 730 pages, £57.99/$89.95, softcover
ISBN: 978-1-4398-2534-1

Table of Contents

Readership: Students in computing, engineering and science.

This book grew out of the author's many years of teaching courses in computing; originally using the FORTRAN language and now almost exclusively using C++. The language allows efficient programs to be written and developed using the object-oriented paradigm which aids the reuse of computer codes.

C++ is an enhancement to the C programming language developed in Bell Labs in 1979. A programming language standard for C++ was late in adoption - originally in 1998 and later revised in 2003. The ISO/IEC148821998 forms the basis for this text. The book is lavishly illustrated with examples and exercises which would make it both an ideal course companion and a book for private study.

The author's abilities to explain briefly the history of computing and to write an engaging text are to be commended. If you buy only one text on programming in C++ then this should be the one for you.

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

Applied Operational Research with SAS
Ali Emrouznejad, William Ho
Chapman & Hall/CRC, 2012, xiv + 270 pages, £59.99/$93.95, hardcover
ISBN: 978-1-4398-4130-3

Table of Contents

Readership: Undergraduate and postgraduate students in operations research and engineering.

Operations research (or *operational research* as preferred by the authors) was introduced during World War II to facilitate the optimal use and allocation of limited war material and resources. Its success relies on a scientific decision-making tool that involves the use of a mathematical programming model. The common feature that these programming models have is that they all involve optimisation, often under constraints.

Algorithms and methods for linear programming, integer linear programming and goal-programming are now common place but their mathematical treatise can often lead to difficulties when translating into analytical codes for computers. This book successfully bridges the gap between theory and practice.

Written primarily for SAS (Statistical Analysis System), the authors present a wide range of optimisation problems to demonstrate how the SAS/OR$^{®}$ procedures work. The text is well-written, easy to follow, and will appeal both to those who have an interest only in concepts and to those merely interested in solving practical problems.

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

Six Sources of Collapse: A Mathematician's Perspective on How Things Can Fall Apart in the Blink of an Eye
Charles R. Hadlock
Mathematical Association of America, 2012, xiv + 207 pages, $50.00, hardcover
ISBN: 978-0-88385-579-9

Table of Contents

Readership: Readers who wonder how 'things can fall apart in the blink of an eye'.

Shooting wild pigeons in Northern Louisiana in 1875 was pretty easy judging by the print on page 3 of this wonderful book. The sky was black with them. Within 30 years, the entire population was zero. How did they collapse so suddenly and so completely? Other questions you might ponder: Where is the Roman Empire? What happened to Polaroid and Kodak? Hear of a case of smallpox lately? How many pet rocks do you own? When do you next fly on TWA? The fact is, big things often seem to just disappear! This book discusses, in a most interesting way, how such processes might be modelled statistically.

Figure 2.14 makes an interesting point, not discussed in most elementary statistics classes. Outliers in a data set can stretch out a fitted normal distribution so that the fitted peak falls below the data peak. The solution for collapse data is to fit a so-called 'stable distribution'. What's that? Didn't your professors teach *you* that, either? In fact, it has four parameters and the normal is a special case. If you teach statistics, look it up now and tell others about it.

This is an utterly fascinating book and will enthral you with its large range of examples. There are three praising reviews on the back cover. The book is even better than those reviews say it is. It provides an entirely satisfying read, cover to cover.

Norman R. Draper: *draper@stat.wisc.edu*
Department of Statistics, University of Wisconsin–Madison
1300 University Avenue, Madison, WI 53706-1532, USA

Time Series Modeling of Neuroscience Data
Tohru Ozaki
Chapman & Hall/CRC, 2013, xxv + 548 pages, £63.99/$99.95, hardcover
ISBN: 978-1-42-009460-2

Table of Contents

Readership: Suitable for experts in time series modelling, especially for neuroscience (fMRI) data.

Professor Ozaki is a devoted friend and former student of the late Professor Akaike, famous for his outstanding contributions to time series and model selection. Professor Ozaki writes in the same style, deep and innovative, but perhaps relying more on intuition than a fully rigorous treatment. He also expresses his thanks to his friend Pedro Valdes, who has helped him in his work on neuroscience time series modelling.

This is a very unusual book on time series, with much that is new, innovative , and usually not found in other books on time series, for example multivariate AR models, multivariate dynamic

models, causal analysis and the Doob decomposition, and so on. Among the major pleasures of browsing through the book are the acquaintance with 'Laplace's Demon', seeing Pearsonian and multimodal distributions as stationary distributions for dynamic models, Einstein's inductive use of Boltzmann entropy - to mention just a few of the novelties.

But the hard core of the book is about state space modelling and its application to neuroscience data. The pages 331 through 351 are a richly textured but precise and detailed introduction to state space modelling. Here is a lovely summary by Ozaki that I have not seen elsewhere - it deals with time series dynamics where some of the variables are 'unobservable as data', and the inference problems include estimation of the unobserved state variables assuming the model is known and inference about the model from time series data. Of course Kalman filters are discussed in detail.

An application to fMRI data appears towards the end of the book, though this topic is discussed several times throughout the book. In the case of this particular data, an earlier analysis found strong causal relations, but it is based on a somewhat restrictive assumption of diagonal innovation variance. A better analysis via state space modelling avoids this restriction and provides better insight. More such examples of neuroscience data, with some brief description of modelling and analytical issues, at a suitable secure location on the Internet would make this book even more valuable.

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

Applied Medical Statistics Using SAS
Geoff Der, Brian S. Everitt
Chapman & Hall/CRC, 2013, xv + 541 pages, £57.99/$89.95, hardcover
ISBN: 978-1-4398-6797-6

Table of Contents

Readership: Medical students; consultant statisticians.

A recent request to a statistical professional body by a doctor seeking help with analysing data they had collected, was greeted with derision by some of the members of that body. The implication was that the request was about as sensible as asking a medical association for help with a medical procedure that was half finished.

The doctor in question may have been better served by simply purchasing this wide-ranging and accessible book. Medical students would also appreciate the range of topics addressed.

However section 3.4.2 describes such things as box plots and t tests as 'familiar', so students would be advised to be reading this book after, or at the same time as, their introductory statistics course.

I think consultant statisticians would also appreciate the refreshers/introductions to statistical techniques and the SAS code for each. Indeed SAS code is liberally scattered throughout the text, and a couple of SAS macros are referred to in the meta-analysis chapter.

The authors do not shy away from sections of theory such as the matrix representation of a regression model and the associated least squares estimator. On the other hand I was slightly disappointed to see the 'quick and dirty' approach to longitudinal data still accorded two entire sections, one for continuous responses and one for binary responses. The 'quick and dirty approach' summarises each subject's profile then applies, for example, a one-way ANOVA to the transformed variable. At least there is a whole chapter on mixed-effects models in there as well.

Amongst the more fashionable sections that caught my eye were: a short section on propensity scores in Chapter 9; a chapter on Bayesian methods and the SAS implementation of MCMC within the genmod procedure; and a chapter on multiple imputation and the mi procedure.

The text is supported by ten pages of references and a sizeable index. The code and example data sets can be downloaded from the SAS website.

Alice Richardson: *alice.richardson@canberra.edu.au*
Faculty of Education, Science, Technology and Mathematics
University of Canberra, ACT 2601, Australia

Mathematics Galore! The First Five Years of the St. Mark's Institute of Mathematics
James Tanton
Mathematical Association of America, 2012, xv + 271 pages, $50.00, hardcover
ISBN: 978-0-88385-776-2

Table of Contents

Readership: School teachers, school and undergraduate students, explorers of mathematics.

The book is a compilation of activities used at the St. Marks Institute of Mathematics in Massachusetts. This Institute attracts both able school students and interested community members to classes, and also offers newsletters and email puzzles.

There are twenty-six chapters or topics, from arctangents to weird multiplication, organised alphabetically. Each topic consists of a puzzler or several, tidbits that enlarge on the puzzler, and concludes with commentary, solutions and thoughts. Eight appendices deal with theorems in slightly less detail than the main chapters. An index groups the chapters into themes comprising calculus, combinatorics, geometry, graph theory, number theory, probability, and trigonometry.

Of most interest to statisticians are likely to be the chapters on Benford's Law and perhaps The Middle of a Triangle, which could be linked to the concept of the mean. Readers looking for activities for Pi Day would also find the chapter on Playing with Pi and the appendix on Bending Buffon's Needle useful.

References are provided for each chapter, and they range from classic research papers of the early 20th century to websites and mathematics education papers of the 21st century.

As an academic it is easy to become bogged down in assisting the struggling students in the class. A book like this is a delightful reminder of the stimuli that high achievers can use to really fly.

Alice Richardson: *alice.richardson@canberra.edu.au*
Faculty of Education, Science, Technology and Mathematics
University of Canberra, ACT 2601, Australia

Risk Modelling in General Insurance: From Principles to Practice
Roger J. Gray, Susan M. Pitts
Cambridge University Press, 2012, xiv + 393 pages, £45.00/$70.00, hardcover
ISBN: 978-0-521-86394-0

Table of Contents

Readership: Students in actuarial science, insurance and risk management.

This textbook provides a detailed, yet easy to follow summary of five major areas in general insurance: modelling of risk, model-based pricing, risk-sharing, ruin theory, and credibility theory. In each topic there are plenty of suitable worked examples and exercises. The statistical package R is used to demonstrate how simple code and functions can be used profitably in an actuarial context. Prerequisites include first courses in calculus, probability and statistics. The presentation of the material is carefully thought through.

The authors of actuarial risk theory books have their traditions. In 1968 Beard, Pentikäinen, and Pesonen stated in the preface of *Risk Theory* that the modern theory of risk can give an actuary concrete assistance in the form of practical applications. Respectively, the authors of this book say that a key feature is the inclusion of three detailed case studies that illustrate how the concepts are used in practice. Both books balance rigour and intuition and more than half of the topics of the two books are the same. Many other actuarial risk theory books have appeared since *Risk Theory*. This book serves as a useful addition to the existing actuarial literature because of its elegant style and detailed case studies.

Lasse Koskinen: *lasse.koskinen@modelit.fi*
Model IT Oy
Unioninkatu 13, FI-00130 Helsinki, Finland

Handbook of Computational Finance
Jin-Chuan Duan, Wolfgang Karl Härdle, James E. Gentle (Editors)
Springer, 2012, xi + 804 pages, €249.00/£224.50/$339.00, hardcover
ISBN: 978-3-642-17253-3

Table of Contents

Readership: Academic: researchers in statistics and finance; Industry: banking, investments and insurance.

This Computational Finance handbook is the fourth in the *Springer Handbooks of Computational Statistics* series. Many of the leading internationally recognised experts in the field

have been assembled to write chapters on this important area of modern Statistics. This handbook provides a carefully chosen survey of the concepts and methods of computational finance, ranging from basic background material through the current frontier of research, e.g. fitting high-dimensional copulae to data.

After the introductory chapter, the book is divided into three parts (II–IV) concerned with, respectively, asset pricing models, statistical inference in financial models, and computational/numerical methods. These parts cover modern financial tools, computationally efficient algorithms, the pricing of complex products and risk behaviour. Several chapters use simulation to illustrate the points to be discussed. Emphasis is put on the applications and statistical aspects of the methods, so many examples are presented. Being a collective work, some minor repetitions and lack of ties between chapters are unavoidable but as a whole this is consistent and self-contained presentation.

The chapters of the Part II give a theoretically oriented picture of assets pricing models including both elementary and more advanced ones. Part III covers a wide range of divergent topics where statistical inference, testing, and especially estimation can be used. Nearly half of the chapters belong to Part IV. They address a large variety of computational issues on topics such as filtering methods, numeric solutions to nonlinear and stochastic differential equations, lattice approach and implied trees, dynamic hedging and approximations, and option pricing using Fast Fourier Transformation. More specific applications are stress testing, portfolio optimisation and bankruptcy prognosis. In addition, the book includes a very brief introduction to software tools Matlab and R.

This handbook is an authoritative and valuable account of an important field. I am sure that it will be an important reference source for researchers and practitioners.

Lasse Koskinen: *lasse.koskinen@modelit.fi*
Model IT Oy
Unioninkatu 13, FI-00130 Helsinki, Finland

Computational Methods in Finance
Ali Hirsa
Chapman & Hall/CRC, 2013, xxix + 414 pages, £59.99/$89.95, hardcover
ISBN: 978-1-4398-2957-8

Table of Contents

Readership: Advanced level students, researchers and practitioners wanting to learn more about computational methods in finance. The book is geared towards useful numerical and computational algorithms.

The depth and breadth of this stand-alone textbook on computational methods in finance is astonishing. It brings together a full-spectrum of methods with many practical examples. The

aim of the book has not been to get into details of stochastic calculus. Instead, the purpose of the book is to aid the understanding and solving of current problems in computational finance.

This book is composed of two parts. The first part describes various methods and techniques for the pricing of derivative contracts and the valuation of variety of models. The second part focuses on model calibration and estimation. Computer algorithms are given, not for the analysis of data, but so that the reader can develop an understanding by working through examples. As a whole this book provides an excellent synthesis of numerical methods needed for solving practical problems in finance.

This book provides plenty of exercises and realistic case studies. Those who work through them will gain a deep understanding of the modern computational methods in finance.

This uniquely comprehensive and well-written book will undoubtedly prove invaluable to many researches and practitioners. In addition, it seems to be excellent teaching book.

Lasse Koskinen: *lasse.koskinen@modelit.fi*
Model IT Oy
Unioninkatu 13, FI-00130 Helsinki, Finland

Computational Statistics, Second Edition
Geof H. Givens, Jennifer A. Hoeting
Wiley, 2012, xviii + 471 pages, €100.20/£83.50/$125.00, hardcover
ISBN: 978-0-470-53331-4

Table of Contents

Readership: Advanced graduate or graduate students in statistics, statisticians, and practising scientists involved in computational statistics.

The book is a major revision, and considerable extension, of the successful 2005 monograph. The authors have updated and broadened the coverage by adding new topics related to MCMC and statistical dependence, and provided R code and datasets at the book website. A couple of quotations from the preface serve to convey the style and purpose of the book: 'Our choice of topics reflects our view of what constitutes the core of the evolving field of computational statistics', 'We have aimed to provide a reasonably up-to-date survey of a broad partition of the field, while leaving room for diversions and esoterica'.

The content is presented in four parts, offering a very apt blend of theoretical development and illustrative examples. The book spans a broad range of modern and classic topics including optimisation, integration and simulation, density estimation and smoothing, and bootstrapping. Algorithms are explained both conceptually and by using step-by-step descriptions. Combinatorial algorithms, including Tabu Algorithms are dealt with in detail. Examples are based on

real-world applications from various fields ranging from genetics to network systems and each chapter ends with plenty of exercises.

This book is full of interest for anyone undertaking computational statistics, and nicely conveys the power and applicability of the presented methods.

Lasse Koskinen: *lasse.koskinen@modelit.fi*
Model IT Oy
Unioninkatu 13, 00130 Helsinki, Finland

Generalized Linear Mixed Models: Modern Concepts, Methods and Applications
Walter W. Stroup
Chapman & Hall/CRC, 2012, xxv + 529 pages, £57.99/$89.95, hardcover
ISBN: 978-1-4398-1512-0

Table of Contents

Readership: Graduate students in statistics, statistics professionals seeking to get up to speed, reference and a continuing education text for researchers and practitioners.

This book provides a modern introduction to linear models (LM). It kicks off with a 'big picture' (chapters 1–3), illustrating and motivating the development which is to follow. Rather than developing an elaborate theory for working with the traditional LM, the author starts with the generalised linear mixed model (GLMM) from the outset. The GLM, LM and LMM are treated as special cases of the GLMM. Chapter 1 addresses the questions like 'What is a model? What should a model do? What are the essential elements of a model?' Chapter 2 focuses on developing understanding the interplay between design principles and modelling. Chapter 3 covers issues like conditional versus marginal modelling, broad versus narrow inference space and data versus model-scale inference.

The 'theory' chapters 4–6 cover the estimation and inference essentials. The remaining ten chapters focus on common applications of GLMMs. This includes multifactor split-plot, clustered, and multilevel models for Gaussian and non-Gaussian data, models for repeated measures and spatial data, count data and overdispersion, models for multinomial categorical data, and other applications of the GLMM.

The numerous examples in this textbook are anchored to SAS® PROC GLIMMIX. The thought processes embedded in GLIMMIX reflect the author's way of thinking about GLMMs. It is an advantage that one syntax applies in a consistent manner to all of the GLMMs covered in this text. However, it would be highly useful to have a website which would display the corresponding examples using R.

The book focuses on data-driven modelling and design processes, and it provides a context for extending traditional linear model thinking to generalised linear mixed modelling. This is a very sound text which teachers of any course on GLMMs should consider adopting.

Erkki P. Liski: *erkki.liski@uta.fi*
School of Information Sciences
FI-33014 University of Tampere, Finland

Paradoxes in Scientific Inference
Mark Chang
Chapman & Hall/CRC, 2013, xviii + 273 pages, £25.99/$39.95, softcover
ISBN: 978-1-4665-0986-3

Table of Contents

Readership: Puzzle and paradox lovers.

'A paradox is a statement or phenomenon that seems contradictory but in reality expresses a possible truth. It can be simply something that is counterintuitive.' (Page vii.) Three examples:

A drunken man with a ring of 10 keys arrives at his door. He picks a key at random and tries it. He repeats this until he finds the right one. Of the possible number of trials, 1, 2, 3, . . . which number is most likely? Would you believe 1? It's true. (p. 44)

Here is a square of numbers:

96	64	37	45
39	43	98	62
84	76	25	57
23	59	82	78

The sums for the rows, columns and diagonals (extended in the usual way to four numbers, of course, for example $64 + 98 + 57 + 23$) are 242. Now write down another square which has all of the 16 numbers rewritten with digits reversed. What are the various sums now? Try it; you will be astonished! (p.40)

A series of n candidates for marriage present themselves, one by one, and each one rejected becomes no longer available. How can we select the best candidate if the current one can be compared only with the ones previously rejected? You will get the best candidate with a probability of nearly 30%, even for large n, if you let go the first 37% (100/e more precisely) and then select the first one better than any previous candidate. If none is better, select the last. (p. 47)

This is a very, very entertaining book. It is extremely well written, extremely clear, and very highly recommended.

Norman R. Draper: *draper@stat.wisc.edu*
Department of Statistics, University of Wisconsin–Madison
1300 University Avenue, Madison, WI 53706-1532, USA

Martin Gardner in the Twenty-First Century

Michael Henle, Brian Hopkins (Editors)
Mathematical Association of America, 2012, xiii + 297 pages, $40.00, softcover
ISBN: 978-0-88385-913-1

Table of Contents

Readership: All who love Martin Gardner type mathematical puzzles.

Beginning in about 1958, Martin Gardner, who died in 2010, became famous for the fascinating mathematical problems he presented and discussed. This wonderful collection of such problems, published by the Mathematical Association of America, beautifully demonstrates the 'strange kind of unearthly beauty' that mathematics can provide.

Here is a simple example from Chapter 24: Three coins are placed in a row. 'Both heads and tails must appear, otherwise the trick ends before it begins.' The blindfolded or back-turned magician orders 1. 'Flip the left coin'; 2. 'Flip the middle coin'; and then asks 3. 'Do all three coins show the same face?' If yes, the trick is accomplished. If no, the magician says 'Flip the left hand coin.' Magically, the 'all same face' objective *is now* achieved! How this always works is demonstrated by drawing a cube whose corners contain all eight possible arrangements. Simple and yet astounding! Try it.

This book, which contains many such beautiful things, varying from the simple to the complicated, would make a wonderful present for anyone who likes puzzles. In particular, try to think of a recipient close by who will lend it back to you!

Norman R. Draper: *draper@stat.wisc.edu*
Department of Statistics, University of Wisconsin–Madison
1300 University Avenue, Madison, WI 53706-1532, USA