The Statistician (2003)
52, *Part* 2, *pp.* 235–255

Book reviews

Books for review

If you would like to review a book, and thereby to retain it for your collection, please contact the Book Reviews Editor, whose details can be found by clicking on 'books currently available' in the information about *The Statistician* on the Royal Statistical Society's Web site:

`http://www.rss.org.uk/publications/`

Reviews in this issue

 0039–0526/03/52235

Probabilistic Risk Analysis: Foundations and Methods
T. BEDFORD AND R. COOKE, 2001
Cambridge, Cambridge University Press
xviii + 394 pp., £37.50
ISBN 0-521-77320-2

This book has developed from a Master of Science level course at Delft University of Technology. Three of the 18 chapters were co-authored by other writers. Although it contains about 60 pages describing distribution theory, and both Bayesian and classical methods of inference, it is aimed at graduates who already have a good knowledge of probability and statistics. Most chapters end with a collection of exercises whose solutions will be made available to *bona fide* teachers.

After the historical and motivating material, and a suitably compact account of the knowledge that is assumed, the meat of the book—'System analysis and quantification'—begins in Chapter Six. Diverse approaches are suggested for how to make an assessment of the chance that some system with interacting components will fail. Applications described include reactors, security systems, river barriers, protection from fire risk and computer software. As well as statistics and probability, the specific desired background includes competence in Boolean algebra, but it is plain that a good general mathematical training is expected. Throughout the book, the places of precise definitions of different modes and causes of failure, and of logical ways of combining information from several sources, are illustrated with genuine data.

The appearance and lay-out are very attractive, although I would have preferred that some ratios had been displayed as a numerator above a denominator, rather than make repeated use of '/'. The different symbols used for AND and OR gates are much too similar. Typographical errors and minor slips are few in number. Among them, it is Allais to whom the credit for exercise 2.4 belongs (and choice C1 should offer 1 million dollars); the formula in exercise 2.6 is not the answer to the problem stated; the definition (page 64) of partition is wrong; Fig. 6.6 has all its labels missing; the entry '3' in the network in Fig. 15.1 should be '1'. There are many acronyms but no index for easy reminder of what they stand for: this is very irritating.

The second edition of a book on the same broad subject, Wilson and Crouch (2001), has recently been published. Although Wilson and Crouch (2001) is far more concerned with medical and occupational risks than the book under review, and is avowedly much less mathematical, it is surprising how little ground the two have in common. Neither book mentions any work by the authors of the other: indeed, although both books have extensive bibliographies, only a handful of items appear in both. This prompts the fears that two schools addressing the same important area are developing almost independently, and that the benefits of the more abstract approach are not widely appreciated.

Often, the main purpose of risk analysis is to suggest actions to reduce risk. The present book describes how the public's perception of risk can be important in galvanizing action, and how the opinions of different experts can be combined. It notes the difficulties of finding agreement on objective measures of the benefits of risk reduction, and the importance of the base rate participation when seeking to quantify risk. Without offering pat solutions to difficult practical and ethical issues, it provides a readable mathematical framework within which these serious matters can be discussed. It can be recommended both for any specialist course on this subject and as a good source of practical applied probability in general.

Reference

Wilson, R. and Crouch, E. A. C. (2001) *Risk–Benefit Analysis*. Cambridge: Harvard University Press.

John Haigh
University of Sussex
Brighton

Techniques of Event History Modeling: New Approaches to Causal Analysis, 2nd edn
H.-P. BLOSSFELD AND G. ROHWER, 2002
Hillsdale, Erlbaum
x + 310 pp., £32.50
ISBN 0-8058-4091-5

This second edition is an attempt by the authors to provide an updated introductory account of modelling event history data since Blossfeld and Rohwer (1995). The objectives of this edition are the same as before:

(a) to demonstrate the usefulness of event history models in uncovering or mapping out possible causal relationships;
(b) to introduce the reader to the statistical package TDA ('transition data analysis'), which is used to analyse the examples presented;
(c) to supplement Blossfeld *et al.* (1989), which I have not read.

The book comprises 10 chapters and an appendix. Chapter 1 opens with a general introduction to event history data, providing examples from the social science field. This is followed by a brief outline of the material that is covered in later chapters. An

illuminating discourse on observation plans (cross-sectional, panel and event history, in that order) and their progressing importance in arriving at potential causal interpretations and relationships is next presented. The chapter ends by defining statistical terms such as transition rate, survivor function and cumulative hazard rate and showing their connections to one another.

Chapter 2 introduces the basic terminology for event history modelling. Concepts such as transitions, states, episodes, censoring and multistate and multiepisode processes are all mentioned. The way that event history data need to be organized and analysed in TDA is covered in detail.

Chapter 3 concentrates on nonparametric methods (life-tables, Kaplan–Meier curves and the log-rank test) used for describing and comparing survival curves for the special case of single-episode data (possibly with competing risks). I was a little disappointed that no explicit mention was made regarding the danger of overinterpreting survival curves. It would have been useful to have a paragraph advising the reader to be careful when interpreting a survival curve, especially when examining the right-hand side of the survival curve or looking at the fine details of the curve, as this information can be unreliable and inaccurate. (The overall survival pattern is more reliable!)

Also, I disagree with the authors when they suggest that one method of comparing survivor functions to determine whether they are statistically significantly different is to calculate 'confidence intervals for each of the survivor functions and then check if they overlap or not'. I would never advocate such an approach for formal assessment, but instead I would recommend constructing an appropriate test statistic (i.e. their second suggested method), such as the log-rank statistic. Besides, is not the confidence band for the difference in survivor functions (in the case of two curves) more appropriate than separate confidence bands for each of the survivor curves? Finally, no mention was made about the potential insensitivity of the log-rank test for curves that cross.

Exponential transition rate models, piecewise constant exponential models and exponential models with time-dependent covariates are introduced in Chapters 4, 5 and 6 respectively, with each of these chapters naturally building on and extending the previous one. The material that is covered in these chapters is interesting and, I believe, essential for appreciating and understanding the more complex event history models that are used regularly in the literature. Additionally, the examples presented there illustrate quite clearly the flexibility of these models to cope with time constant and time-varying covariates, multiple destinations (i.e. competing risks models) and multiple episodes data, and furthermore demonstrate the versatility of the TDA package. My only qualm was that the authors could and should have explicitly defined the terms exogenous and endogenous, rather than assuming that the reader (though probably a social scientist and thus already familiar with these terms) knew their meanings.

Chapter 7 discusses parametric models (such as the Weibull, the Gompertz, the log-logistic, the log-normal and the Sickle models) for duration data, which unfortunately I found rather repetitive. It could have been written more concisely. In Chapter 8, two graphical approaches to assessing the parametric assumption for the models described in Chapter 7 are discussed. The first is based on transformations of survivor functions (e.g. complementary log-survival curves) and the second is based on residuals.

Chapter 9 describes the very important Cox semi-parametric approach to modelling event history data. Topics such as stratification and assessing the proportional hazards assumption (both graphically and analytically) are addressed. However, it would have been useful for the authors to discuss more fully the extensions of the Cox model for survival analysis data to data sets with multiple events of the same type or different types, i.e. I would have liked to see the authors address more fully the issues regarding extending the Cox proportional hazards model to models that take into account the intraunit correlation. Discussions of the frailty and marginal approaches (Therneau and Grambsch, 2000; Andersen and Gill, 1982; Wei *et al.*, 1989; Prentice *et al.*, 1981), robust variance–covariance matrices and multiple timescales would have been appreciated.

The final chapter concludes the book by addressing issues to do with model (mis)specification. The issue of unobserved heterogeneity (due to omitted covariate information) and the topic of fitting parametric mixture models to account for this unexplained heterogeneity are critically covered in the majority of this chapter. In the discussion section other aspects of model specification (e.g. defining the state space) are also briefly summarized. Appendix A provides basic information on the TDA package.

Modelling of event history data is a very active area of research in medical and social science. This book is written for the social scientist (reflecting the authors' area of research), but all the ideas that are introduced in it can easily be translated to other areas of research. Overall, this book appears to have accomplished, at least, the first two of its aims. The third aim of being a supplement to Blossfeld *et al.* (1989) is difficult for me to judge, as I have not read

it, and therefore the details (technical or practical) that I found lacking in this second edition may have appeared there. I would recommend this book to social science researchers who are interested in event history modelling and those researchers who are interested in using the TDA package. The topics and examples that are covered would be accessible and relevant to this readership.

References

Andersen, P. K. and Gill, R. D. (1982) Cox's regression model for counting processes: a large sample study. *Ann. Statist.*, **10**, 1100–1120.

Blossfeld, H.-P., Hamerle, A. and Mayer, K. U. (1989) *Event History Analysis*. Hillsdale: Erlbaum.

Blossfeld, H.-P. and Rohwer, G. (1995) *Techniques of Event History Modelling*. Hillsdale: Erlbaum.

Prentice, R. L., Williams, B. J. and Petersen, A. V. (1981) On the regression analysis of multivariate failure time data. *Biometrika*, **68**, 373–379.

Therneau, T. M. and Grambsch, P. M. (2000) *Modeling Survival Data: extending the Cox Model*. New York: Springer.

Wei, L. J., Lin, D. Y. and Weissfeld, L. (1989) Regression analysis of multivariate incomplete failure time data by modelling marginal distributions. *J. Am. Statist. Ass.*, **84**, 1065–1073.

Brian D. M. Tom
Medical Research Council Biostatistics Unit
Cambridge

Statistical Consulting

J. Cabrera and A. McDougall, 2002
New York, Springer
xii + 390 pp., £62.50
ISBN 0-387-98863-7

There now seems to be a burgeoning little industry of statistical consulting texts; the authors of this book cite 12 other references on the subject. This book uses two approaches to the subject. In part one a broad approach is taken, putting statistical consulting in its historical and contemporary context, before looking at methodological issues and the life-cycle of a project. This sets the scene for part two, which consists of 20 case-studies set into three groups. Each group represents a level of complexity of analysis and design. They start with basic χ^2- and t-tests, analyses of variance and regression models. They then move on to logistic regression, time series and factorial designs. In the final group, mixed effects models, multivariate exploratory techniques, recursive partitioning and classification and regression trees are covered. There are appendices containing an outline of a statistical consulting course, fairly solid introductions to the SAS and S-PLUS programs and technical descriptions of probability distributions, statistical tests and power calculations.

This book covers much ground and is clearly aimed at maintaining undergraduates' awareness of how to apply the techniques that they are learning. More than anything, it should encourage them to think flexibly rather than to try to design every project so that it is another chance to apply their favourite statistical test. The breadth of areas and techniques covered can feel a little overwhelming, but then this is a reflection of how real world statistical consulting can feel. Achieving this breadth must represent a large amount of work by the authors. Given this breadth it is a book that really requires a teacher or lecturer to guide the student through it. The outline course at the back of the book gives a useful and realistic plan to do this. By assigning each of the three groups of cases-studies to an undergraduate year it is fairly easy to see how this book could see a student through their undergraduate course. This would achieve the highly desirable goal of helping to keep the student's mind clearly focused with regard to the applied, coal-face, end of statistics. At the same time the concise descriptions of statistical techniques should provide an effective reinforcement of what is being taught on the rest of the course. The structure could just as easily be applied to a three-term post-graduate course.

One of the book's greatest strengths is its chapter on communication; although whole books have been written about this subject, e.g. Derr (2000), this succinct little chapter gives one of the best introductions to report writing and graphics that I have come across. I hesitate to raise the issues of grammar and style, partly because my own will probably be scrutinized in the light of my comments. However, there are just a couple of pages of sensible advice that many undergraduates and new graduates could profit from following. On the subject of graphics, again following the advice of the two or three pages may lead to more students producing graphs that not only look impressive but also are properly annotated and so can be understood.

With its wealth of Web sites, the availability of data sets and broad coverage this would make a good recommended text-book for a graduate or post-graduate degree course with an emphasis on applied statistics. Its expense (unless the publisher intends to produce a paperback version) may be justified if it proves to be the student's companion throughout the course.

Reference

Derr, J. (2000) *Statistical Consulting: a Guide to Effective Communication*. Pacific Grove: Duxbury.

James Tighe
Oxleas National Health Service Trust
London

A Course in Probability Theory, 3rd edn
K. L. CHUNG, 2001
San Diego, Academic Press
xviii + 420 pp., £39.95
ISBN 0-12-174151-6

The first edition of Chung's classic *Course in Probability Theory* appeared in 1968, the same year as the third edition of Feller's *Introduction to Probability Theory and Its Applications*. Both books have been greatly influential, but in different ways. Whereas Feller stressed both theory and applications, Chung's intention was to write a mathematics book. He says,

> 'Although many notions of probability theory arise from concrete models in applied sciences, recalling such familiar objects as coins and dice, genes and particles, a basic mathematical text (as this pretends to be) can no longer indulge in diverse applications, just as nowadays a course in real variables cannot delve into the vibrations of strings or the conduction of heat'.

Some idea of the coverage of the book can be gleaned from the list of contents: 1, 'Distribution function'; 2, 'Measure theory'; 3, 'Random variable, expectation, independence'; 4, 'Convergence concepts'; 5, 'Law of large numbers, random series'; 6, 'Characteristic function'; 7, 'Central limit theorem and its ramifications'; 8, 'Random walk'; 9, 'Conditioning, Markov property, martingale'; 'Supplement on measure and integral' (new); 'General bibliography'; 'Index'. Despite the relationship between random walks and stochastic processes, there is no mention of birth–death or branching processes. Although martingales are regarded 'as an essential tool', the serious study of Markov processes, etc., is deferred to a subsequent course on stochastic processes.

The two earlier editions of the book presupposed a certain mathematical maturity of the student. Professor Chung has found, however, that many students from departments such as statistics, operations research and electrical engineering have needed to take a prerequisite course on abstract measure and integral theory. The defining addition in this new edition is a largely self-contained supplement on these topics.

Chung first taught a course based on this book, entitled 'Advanced probability', in 1966 at Stanford University. The back cover says,

> 'Since its publication by Academic Press [nearly 35 years ago], tens of thousands of students have taken a probability course using this classic textbook .... It has been used successfully at over 75 universities since its initial publication.'

One reason for its long-standing popularity is its flexibility. A short course built around Chapters 2–4 with selections from Chapters 5 and 6 and the early parts of Chapter 9 is one suggestion made by the author. Another suggestion, for mature mathematicians, is to 'begin with Chapter 3, go at once to Chapter 5 with a few glances at Chapter 4, skim through Chapter 6, and take up the remaining chapters seriously'.

If you or your department already have an earlier edition of this book, then consider buying a book on measure theory instead of this new edition. If you do not already have access to Chung's *Course in Probability Theory*, then try to ensure that there is a reference copy in your departmental library.

Freda Kemp
University of St Andrews

Analysis of Longitudinal Data, 2nd edn
P. J. DIGGLE, P. J. HEAGERTY, K.-Y. LIANG AND S. L. ZEGER, 2002
Oxford, Oxford University Press
xvi + 380 pp., £40.00
ISBN 0-19-852484-6

> 'Longitudinal studies are rare, other than among children, and it is wrong to think that a cross-sectional view ... can be a substitute, at least over a long period ...' (Rosenbaum, 1988).

This is a timely update of the authors' first book on the analysis of longitudinal data. The aim of this book is to describe statistical models for the analysis of longitudinal data and to update the current literature in this area. A recurring theme is the comparison of longitudinal data with that collected cross-sectionally. The statistical level of understanding is that of a post-graduate student in statistics. A strong emphasis is on applications for both the biological and the health care sciences. Indeed, most of the data sets are in the public domain.

The book is organized into 14 chapters plus an appendix detailing statistical concepts. The first three chapters give an overview of the subject. Indeed, the first sentence states that the defining characteristic of longitudinal data is that measurements are made repeatedly over time. This is in contrast with cross-sectional data where observations are measured once. The authors make an interesting statement saying that longitudinal data can be collected either prospectively (examining subjects over time) or retrospectively (via historical records). My experience is of the former. Longitudinal data require special statistical models as it is assumed that observations that are made over time are correlated.

What I tend to do when faced with repeated mea-

surements is to collapse the data into one or two summary measures according to Matthews *et al.* (1990). This approach is commented on briefly in Chapter 1. I was surprised that there was no reference to Matthews *et al.* (1990), as it is so well known. I guess that it is a function of how the approach to longitudinal data analysis has changed over the last decade. Matthews *et al.* (1990) argued for the simplicity of summary measures, an approach that has my sympathy. However, the analysis of repeated measurements has moved forwards since then and it is this that I now focus on.

Chapter 2 discusses design considerations with particular reference to sample size calculations. For cross-sectional studies five characteristics are needed to calculate sample size. These are the type I error rate, clinically worthwhile difference, type II error rate, measurement variation and drop-out rate. For longitudinal data two other things should be taken into consideration, namely the number of repeated measurements per person (this they say may be limited by study costs) and the degree of correlation between the repeated measures. The authors suggest that estimates of correlation may be available from the literature. I doubt this and imagine myself having to resort to guesswork. Examples of sample size calculations are given for both continuous and binary response data (assuming a type I error of 5% and type II error of 20%).

Chapter 3 discusses how to explore longitudinal data graphically. This is a natural starting-point for any study (longitudinal or not) and I found it a most useful reference. Tukey (1977) brought this back into fashion. The next few chapters discuss the general linear model for longitudinal data. One of the problems highlighted by the authors concerns sources of random variation. Three such sources are outlined, namely random effects, serial correlation and measurement error. The authors discuss different models for taking into account the variations outlined, including combinations of all three. The following chapters discuss special cases of longitudinal data and include binary response data, random-effects models and methods for categorical data. All the chapters give data examples as well as the detailed theoretical treatise. There is more on sample size estimation taking into account more than two groups.

Chapter 13 discusses the problems of drop-outs. This must be more a problem in longitudinal data when compared with a cross-sectional study. They make a nice distinction between purely missing data and data measured intermittently. The latter pose more problems statistically as a wider variety of non-response patterns need to be considered. A simple method to handle drop-outs is to analyse only complete data sets (this could be wasteful of data). An alternative method is to take the last observation forwards, which is used routinely in the pharmaceutical industry. If patients show improvement over time, this technique has the advantage of a conservative effect of treatment. This is a strong argument in its favour. However, the authors do not advocate this method of handling missing data, though they fail to justify their case. They prefer, instead, a random-effects model, arguing that a patient's pattern of responses is likely to depend on many characteristics (observed and not observed). The unobserved characteristics are included in any model as random effects and modelled as such. For me this is the best chapter of the whole book. It seems that there is much research to be done in this area. Missing values are not so much a problem if you adopt the approach of Matthews *et al.* (1990) of summary measures.

Is there anything missing? Bootstrapping is not covered though I expect that this is a novel field though I did note some correspondence on ALLSTAT recently re bootstrapping. Perhaps a history of longitudinal measurements could have been considered. There is also scant reference to statistical software. The authors argue that the software needs to catch up with the theory first and I accept this.

It is a long time since I undertook my post-graduate course in statistics. At the time methods for handling longitudinal data were in their infancy. I wish I had had this book then. It is excellent in all respects. What I need now to use the methods effectively is an appropriate training course. I recommend this book without reservation.

References

Matthews, J. N. S., Altman, D. G., Campbell, M. J. and Royston, P. (1990) Analysis of serial measurements in medical research. *Br. Med. J.*, **300**, 230–235.

Rosenbaum, S. (1988) 100 years of heights and weights. *J. R. Statist. Soc.* A, **151**, 276–309.

Tukey, J. W. (1977) *Exploratory Data Analysis*. Reading: Addison–Wesley.

Alan S. Rigby
University of Sheffield

Research Methods for Postgraduates, 2nd edn
T. GREENFIELD (ed.), 2002
London, Arnold
xiv + 270 pp., £29.95
ISBN 0-340-80656-7

This soft cover book consists of 42 chapters divided into 10 parts. There are 28 authors, most of whom are from the UK, but two are from Australia and one is from the USA. The diverse backgrounds of the authors, and the editor's presentation style, facilitate a light, refreshing, yet lively flavour to this

serious subject. The book aims to provide post-graduates with a concise, but thorough, reference text for research methods. Chapters on creativity, use of the Internet and software packages are new to this edition.

An introduction forms part I, in which research planning, time management, documentation, word processing tips and ethics are discussed. Part II covers writing research proposals, funding and other support for research. 'Tools of research', part III, presents topics on information technology, software, library services and the Internet. Part IV discusses creativity tools, including creative problem solving, software and Web sites for creativity. Part V, 'Research types', addresses specific types of research, such as randomized trials, laboratory and industrial experiments, agricultural experiments and survey research.

'Measurement', the largest section, is the topic of part VI, and 'Analysis' makes up part VII. 'Special tools', part VIII, first provides an introduction to mathematical models and then discusses the topics of deterministic models, stochastic models and simulation, and, lastly, optimization. Part IX, 'Presentation', gives valuable tips on writing, graphic display, oral presentation and Internet home page construction. The subjects of part X are protecting and exploiting technology, intellectual property rights and career opportunities. I feel that part V, 'Research types', is the section that could be most expanded, to enhance the next edition.

The main strength of the book is the extent to which it provides a vast interdisciplinary coverage of topics that compose the elements of research. It is backed by an abundance of references for focused detailed study of specific topics. The emphasis on the Internet and software are additional strengths. A minor weakness is that part V is a little succinct. It could be lengthened, so that the discussion includes additional types of research. The index consists of only four pages, but it is adequate. Overall, I am impressed by the richness of this book as a resource for post-graduate research, and I consider it an essential guide for post-graduates. The authors' aims have been achieved. I highly recommend it for university library purchase.

Mark A. Best
Eastview

Matrices for Statistics, 2nd edn
M. J. R. Healy, 2000
Oxford, Clarendon
x + 148 pp., £22.99
ISBN 0-19-850702-X

Michael Healy's *Matrices for Statistics* has been in print since 1986. This greatly rewritten new edition, like its predecessor, aims to give users of statistical methods involving matrices an understanding of the underlying matrix theory while avoiding aspects of the theory that they are unlikely to meet.

Chapter 1 is introductory. A useful feature is the index of terminology with section references with which it ends. Determinants, inverse matrices and the concepts of linear dependence and rank are the topics in Chapters 2, 3 and 4 respectively. Simultaneous equations and generalized inverses are studied in the context of least squares problems in Chapter 5. Chapters 6 and 7 move on to linear spaces and to quadratic forms and eigensystems. The final chapter is new. It contains a miscellany of topics that have been suggested by readers of the first edition; these include direct sums and products, vectorizing, matrix calculus, matrices with complex elements and quadratic forms in normal variates.

There are two appendices, on matrix calculation (including precision of results) and on matrix algorithms. The index is short yet adequate.

This new edition is deliberately less concise than the earlier one. Some material that was formerly given in the form of exercises for the reader is now discussed in the text. There remain many student examples; although there is no formal 'Answers' section in the book, many of the examples lead the reader towards their solutions.

This is a widely used text, suitable for statisticians, life scientists, behavioural scientists and engineers, that has been brought up to date and made more easy to read for non-mathematicians.

Freda Kemp
University of St Andrews

Nonlinear Statistical Modeling
C. Hsiao, K. Morimune and J. L. Powell (eds), 2001
Cambridge, Cambridge University Press
xviii + 452 pp., £65.00
ISBN 0-521-66246-X

The Cambridge series 'International symposia in economic theory and econometrics' aims

> 'to provide *refereed* journal-quality collections of research chapters of unusual importance in areas of currently highly visible activity within the economic profession'.

This is the 13th volume in the series. It is dedicated to Takeshi Amemiya and contains 16 papers inspired by his work on topics as diverse as dependent variables, discrete choice, non-linear estimation, panel data and simultaneous equation models.

The book begins with a unification of the treatment effect literature and the latent variable literature by J. J. Heckman and E. J. Vytlacil. Chapter 2, by N. E. Savin and A. H. Würtz, compares the properties of hypothesis tests for limited dependent variable models using

(a) critical values based on first-order asymptotic approximations and
(b) critical values calculated by bootstrapping.

Simulation estimation of sample selection models is L.-F. Lee's topic in Chapter 3. A new approach to the attrition problem in longitudinal studies is presented by K. Ryu in Chapter 4. The next three papers, by F. Goto, by J. L. Powell, and by Y. Nishiyama and P. M. Robinson, are all concerned with semiparametric estimation problems.

In Chapter 8, C. F. Manski studies bounds on relative risks and attributable risks by using auxiliary distributional information. The selection of regressors in a regression model is a long-standing research topic; D. McFadden's contribution to this book is to show that prescreened least squares is equivalent to direct estimation by non-linear least squares. T. E. MaCurdy presents a new estimation method for regression and simultaneous equations that uses the second and higher moments of the disturbances. In Chapter 11, T. W. Anderson and L. You prove that the existence of the second moment is sufficient for weak convergence of the standardized spectral distribution when the process consists of independent and identically distributed variables.

Unit root tests are studied by H. Lütkepohl, C. Müller and P. Saikkonen and by K. Morimune and M. Nakagawa in Chapters 12 and 13. A simultaneous switching autoregressive model for the analysis of financial time series data is developed as a generalization of the Tobit model by N. Kunitomo and S. Sato. The last two chapters, by T. Lancaster and by S. J. Yhee, J. B. Nugent and C. Hsiao, deal with problems arising in the analysis of panel data.

The book ends with Takeshi Amemiya's *curriculum vitae* and an index. These papers form a worthy tribute to him on the occasion of his 65th birthday.

Freda Kemp
University of St Andrews

Generalized, Linear, and Mixed Models
C. E. McCulloch and S. R. Searle, 2001
New York, Wiley
xxii + 326 pp., £70.50
ISBN 0-471-19364-X

I very much enjoyed reading this excellent book. The book is written for graduate statistical and mathematical students, and for practising statisticians; it can be very useful for statisticians in different environments and with different levels of experience.

In Chapter 1, the basic ideas of fixed and random factors, and mixed models are introduced. It gives a brief discussion and definition of generalized models, the generalized linear model (GLM) and the generalized linear mixed model (GLMM). If all parameters are considered as fixed then the model is a GLM. If the model includes fixed and random effects then the model is a GLMM. In addition, a few examples from clinical trials are presented to give a better understanding of different models with fixed or random effects. There is also a brief description of the analysis for such models.

Chapters 2 and 3 introduce all the main ideas in the remainder of the book, using the two simple contexts of one-way classification and linear regression. Chapter 2 deals with data which are either normally or Bernoulli distributed. Chapter 3 covers simple regression, i.e. regression with a single predictor. For a simple linear regression, the restrictive assumptions that y is a linear function of a predictor x is only a crude approximation or the first step in a more detailed analysis. Sometimes, it is for only a short interval of time that y is a linear function of x; over the entire range of time, the relationship can be non-linear. In some cases, we must transform y and/or x before the linearity assumptions are met even approximately. A different approach would be to model some known function of the mean of y as linear in x. Maximum likelihood (ML), as an estimating method, and likelihood ratio techniques are introduced for both balanced and unbalanced data. The authors suggest that these first three chapters, with a minimum of emphasis on generality of results and notation, together with some additional material, can be used for a one-semester Master's course.

Chapter 4 describes the general ideas of linear models (LMs). First, the general model and its notation are introduced; then an LM for fixed effects is discussed. The maximum likelihood method is used to estimate the vector slope β at the same time as the variance. Using this method, there are many solutions, but none of them is unbiased for β. The impracticality of that solution for β is avoided with the introduction of an estimable function of β. The ML estimator of the estimable function is unbiased and based on the sufficient statistics; it is a uniform minimum variance unbiased estimator. Examples of one-way and two-way classifications are presented with the following linear hypothesis tests: likelihood ratios, t-tests and confidence intervals.

In Chapter 5, 'Generalized linear models (GLMs)', a brief discussion that GLMs can han-

dle probit or logistics regression, Poisson regression, log-linear models for contingency tables, variance components estimation from analysis-of-variance mean squares and many other similar problems is presented. The building of GLMs is considered with respect to the following three aspects: the distribution of data, the function of the mean that will be modelled as linear in the predictors and the predictors. A review of tests of hypothesis such as likelihood ratio tests, Wald tests and associated confidence intervals is presented. Quasi-likelihood as an inferential method which works as well or almost as well as ML, but without having to make specific distributional assumptions, is discussed at the end of this chapter.

Chapter 6 is an extension of Chapter 4. The starting model in Chapter 4 is an LM with fixed effects whereas in Chapter 6 the random effect is added inside the model. This model is called a linear mixed model (LMM). A brief description of the general nature of longitudinal data in connection with LMMs is given. The estimation of fixed effects for known and for unknown variance (matrix) is presented and discussed; next the prediction of random effects for known and unknown variance is shown. The analysis-of-variance estimation of variance components and restricted ML for balanced and unbalanced data are presented at the end of this chapter.

Chapter 7 gives an extension of the longitudinal data modelling that is introduced in Chapter 6. The following models for both balanced and unbalanced data are examined: models for uncorrelated subjects; models that are uncorrelated between and within subjects; models that are uncorrelated between and autocorrelated within subjects; models that are correlated between but not within subjects. A generalized estimating equation model is presented at the end of this chapter.

The use of random effects is not restricted to LMMs. How should the random effects be incorporated in non-linear models? To answer this question, the GLMM is introduced in Chapter 8. This chapter explores the consequences of adding random factors and studies estimation by ML. The following topics are covered: marginal *versus* conditional models; other methods of estimation such as generalized estimating equations; penalized quasi-likelihood; conditional likelihood and tests of hypothesis such as likelihood ratio tests, asymptotic variances, Wald tests and score tests.

Chapter 9 deals with three different but interrelated methods of prediction: best prediction, best linear prediction and LMM prediction (best linear unbiased prediction). Because these methods do not all require the same assumptions, the assumptions required are discussed separately.

In Chapter 10, the authors identify some of the common methods that are used for likelihood calculation and maximization: the EM algorithm for ML and for restricted ML; the Newton–Raphson method for LMMs; numerical quadrature (Gauss–Hermite quadrature and likelihood calculations); Markov chain Monte Carlo algorithms; stochastic approximation algorithms for GLMMs. In addition, they mention and give references to some current research topics.

Chapter 11 discusses the topic of non-linear mixed models. The GLMMs are a proper subset of non-linear mixed models. Two examples (agriculture and pharmacokinetic models) are presented and discussed.

I would certainly recommend this book. It is definitely worth having *Generalized*, *Linear*, *and Mixed Models* as a part of your library.

Lada Mitchell
Novartis Pharma
Basel

Applied Statistics in the Pharmaceutical Industry, with Case Studies in S-Plus

S. P. MILLARD AND A. KRAUSE (eds), 2001
New York, Springer
xvi + 514 pp., £79.95
ISBN 0-387-98814-9

This book aims to give a general guide to statistical methods that are used in the pharmaceutical industry, and to show how to use the S-PLUS software package to implement these statistical methods.

There are 19 chapters which are organized to reflect the natural sequence of drug development: preclinical studies, early studies in humans, three phases of clinical trials (phases II, III and IV) and manufacturing or production. With 32 contributors to the book, from a wide variety of backgrounds (North American and European researchers from academic institutions and industry) it is not surprising that each chapter has a very different narrative. Thankfully, however, most chapters have a common structure that helps to orientate the reader. Each chapter contains an introduction to the topic, including the relevant statistical theory and regulatory considerations, followed by worked examples that illustrate the analysis and interpretation of each statistical problem. There is also an appendix to each chapter that describes all relevant S-PLUS programming code.

Overall, the book meets its aim, covering a wide range of statistical methodology and dealing with many situations that are encountered in pharmaceutical statistics. The authors should be congratu-

lated for showing how S-PLUS can be used in each situation although some readers may not like the reliance on user-written functions to do this. Also, there are some situations that are either not ideally suited to S-PLUS (Gibbs sampling in sample size re-estimation) or rely on output from other statistical software (use of NONMEM output in an example of Monte Carlo simulation), or cannot currently be implemented in S-PLUS (the continual reassessment method, a Bayesian-based dose escalation rule). Although most areas are covered in sufficient statistical detail, I would have expected binary and survival data to be covered in the power and sample size section, which is restricted to the non-central t-distribution. Nevertheless, the strength of the book is in the detail of the examples that illustrate the power of S-PLUS in interactive analysis and its ability to display graphical results in innovative and revealing ways. I particularly liked the examples showing the use of multipanel graphs and contour plots. Therefore, the book succeeds in tempting this died in the wool SAS user, and novice to S-PLUS, to try S-PLUS in more situations, especially given the useful Web site, which covers all examples in the book and allows users to download S-PLUS programming code easily.

I would recommend this book to any pharmaceutical statistician who wishes to explore alternative software for interactive data analysis and visualization. It may also be useful as a teaching text for potential medical statisticians because of the numerous real life worked examples and their software solutions.

Nigel Brayshaw
Deal

Applied Stochastic Modelling
B. J. T. Morgan, 2001
London, Arnold
xxii + 242 pp., £19.99
ISBN 0-340-74041-8

Modelling is fun to do. I hope that we all know the satisfaction in finding a suitable model to explain the patterns in an interesting set of data. In appropriate places, the model itself can shed light on the underlying mechanisms which give rise to the data.

The author's enthusiasm for his subject shines through this book. There are plenty of interesting example data sets, many from the field of wildlife studies and ecology, with others from many fields and an intriguing example of inflorescence patterns of strawberry plants, which is a topic I had never given any thought to before. Many of these examples have arisen from the author's own consulting activities, and the impetus for the book seems to come from these experiences and from courses run at the University of Kent and at Pfizer.

The book covers much ground in quite a short space. The full table of contents can easily be found on the Web, so I shall not give it here. The main themes are the use of likelihood methods and simulation. There is a very elegant presentation of the use of profile likelihoods and likelihood ratios to define multiparameter confidence regions. The simulation theme runs through many major areas, including randomization tests, bootstrapping, the EM algorithm and Bayesian analysis.

Data sets and code for MATLAB® and S-PLUS are supplied on the book's Web page. Each chapter has a good set of exercises.

In conclusion, I like this book and strongly recommend it. It covers many of my favourite topics. In another life, I would have liked to have written it, but Professor Morgan has made a better job of it than I would have done.

Tim Auton
Protherics
Macclesfield

Sources of Non-official UK Statistics, 5th edn
D. Mort, 2002
Aldershot, Gower
xii + 358 pp., £67.50
ISBN 0-566-08449-X

This fifth edition in seven years confirms the continuing search and demand for statistics of all kinds. Minimal information of almost 900 sources from over 400 organizations carries with it telephone contact numbers, Web site addresses and indications of whether some or all of the information is free. The structural format of the information and the comprehensive index assist the user to identify, clearly and quickly, possible sources. However, unlike the guide that is produced for National Statistics, the information is limited and rarely carries the sample size details of surveys.

Delving into the detail of the index revealed a few errors both of classification and of referencing. It was surprising to find only one reference to Wales though information about Wales can be found from the sources indexed under 'Region'.

This book will continue to be a valuable pointer both to data sources and to organizations. Market researchers will find it a handy reference to be kept close to hand. Libraries should stock this book for students who are always on the lookout for odd statistical series. The addition of a compact disc containing the current information, with a means of

updating the information between editions, would be a great advantage for serious researchers.

Ed Swires-Hennessy
Local Government Data Unit—Wales
Cardiff

Generalized Linear Models with Applications in Engineering and the Sciences
R. H. MYERS, D. C. MONTGOMERY AND G. G. VINING, 2002
New York, Wiley
xiv + 342 pp., £66.50
ISBN 0-471-35573-9

This book aims to present an introductory treatment of generalized linear models (GLMs), for students who already know the basics of regression, model fitting and statistical inference procedures. The authors' stated aim is to make the book accessible to a wide audience, including engineers and scientists as well as statisticians, and to present both the theory and its application. In accordance with this aim, the book includes a wide range of illustrative examples, many end-of-chapter exercises based on real data and illustrations of the use of popular statistics software (MINITAB, SAS and S-PLUS) to fit the models that are discussed.

The detailed contents of the book are, perhaps, not exactly what we would expect from the title. This is not in itself a bad thing, but potential readers and users should be aware of what they will find. After a brief introductory chapter, there is an extensive chapter on linear regression models, and another on non-linear regression. As in the rest of the book, this material is written with clarity and enthusiasm, and with a good balance between theory and application. Any comprehensive book on GLMs must contain some material on what one might call 'non-generalized' linear models, which are after all special cases of GLMs and which provide crucial background for the models that are discussed later; however, these two chapters make up almost a third of this book.

The middle third of the book, roughly, contains the central core material on GLMs. This begins with a chapter on logistic and Poisson regression: models, approaches to inference, ways of dealing with overdispersion and many interesting and rewarding examples. After that, the authors move on to the general definition of GLMs, illustrated principally by examples using the gamma distribution.

The final third of the book deals with more extensions and further applications. There is a chapter on generalized estimating equations, with examples of their use in longitudinal studies and in split-plot experiments. The final chapter covers a selection of more advanced topics, including aspects of experimental design, simulation studies of the properties of inference procedures (including the situation where the link function is misspecified) and generalized additive models.

This is a book with both strengths and weaknesses. On the positive side, the authors achieve a writing style that is at once clear and enthusiastic. Their use of examples shows that these techniques and ideas have real practical value. Yet theory is not neglected; the balance between theory and application seems just right for the intended audience of graduate level statisticians, scientists and engineers, and on the European side of the Atlantic it would perhaps be appropriate for some advanced undergraduate courses as well.

On the negative side, there were some omissions that rather surprised me. I rather expected to see something on ordinal response data, and/or on log-linear models. These are both areas where the interpretation of GLM analyses can be tricky. (However, no book in this wide and still developing area could cover everything.) I was also rather disappointed with the way that the computer examples were used and presented. The book is certainly not free of misprints, and there are several in the data tables. Any reader wanting to work through the example analyses for themselves would need to key in the data (because they do not seem to be available on a linked Web site), work out in some cases where the misprints are and work out for themselves actually how to run the analyses in SAS or S-PLUS. The book's publicity materials and preface both promise rather more than the book actually delivers in terms of computer guidance and illustration. For instance Chapter 5 (on the general family of GLMs) is said to contain 'several illustrations ... employing both SAS PROC GENMOD and S-PLUS', while in fact all the numerical results come from SAS, and the use of S-PLUS seems to be limited to producing various diagnostic plots. (Indeed S-PLUS is never explicitly mentioned in the chapter.) Generally, apart from a brief comparison of `SAS PROC LOGISTIC` and `PROC GENMOD` in logistic regression, there is no mention of the strengths and weaknesses of different statistical packages for the analyses described.

Finally, the book shows some inconsistencies in detail. For example, the very last section of the chapter on logistic and Poisson regression (Chapter 4) introduces the idea of classification (`CLASS`) variables in SAS, by referring back to a previous example where they could have been used but were not. In fact, the SAS output in that example makes it clear that a `CLASS` variable was used; indeed the output might make little sense to a reader who did not

already know what such a creature was. (In passing, I should note that the authors clearly do not take the view that marginality constraints in factorial models are essential.)

In summary, this is a useful and interesting book, with some flaws. It provides a valuable reference and practical advice, and could provide support for a taught course; but a reader who is new to the subject, working through the book on their own, might have some difficulties. I recommend it for library purchase.

Kevin McConway
The Open University
Milton Keynes

Modeling in Medical Decision Making: a Bayesian Approach
G. PARMIGIANI, 2002
Chichester, Wiley
xiv + 268 pp., £45.00
ISBN 0-471-98608-9

Published in Wiley's 'Statistics in practice' series this book is primarily aimed at students and practitioners of biostatistics. For practitioners who are involved in medical decision-making, this book provides a glimpse of the challenges in the use of quantitative modelling in decision-making. For statisticians, this book offers challenging statistical problems encountered in medical decision-making.

The book has two parts: part I, 'Methods' and part II, 'Case studies'. Part I has three chapters. The first chapter is on 'Inference', wherein fundamentals of Bayesian inference are discussed with examples. The topics discussed here include conditional probability, specificity, sensitivity, Bayes factors, odds, nomograms, conditional independence, prior and posterior probabilities and predictive values. Examples from medical diagnosis, genetic counselling and chronic disease modelling are used as illustrations for some of the methods introduced. The second chapter is 'Decision making'. This chapter contains fundamentals of Bayesian decision theory and foundations of expected utility theory. Among the topics that are discussed in this chapter are quantifying the value of health outcomes to individuals by using utility theory, decision-making in health care, the Bayesian approach to prediction and expected utility, statistical decision theory and loss functions. The third chapter in part I is 'Simulation', which includes discussions on Markov chain Monte Carlo methods, using simulation for prediction, calculating expected utility, probabilistic sensitivity analysis and searching for strategies.

In part II there are three chapters also. Chapter 4 illustrates the use of Bayesian meta-analysis in developing probability distributions on the magnitude of the effects of a medical treatment. The case-studies in this chapter include an overview of clinical trials of adjuvant tamoxifen for women with breast cancer and clinical trials about treatments for migraine headache. Chapter 5 is on 'Decision trees'. The modelling of multistage decision problems by using decision trees is illustrated in this chapter with a case-study of axillary lymph node dissection, a surgical procedure that is commonly used in early breast cancer. 'Chronic disease modelling' is the title of Chapter 6. A decision model exemplifying the type of multicomponent approaches that are needed to address complex decisions in chronic disease modelling is presented, using a study of the appropriate frequency of examinations in screening for breast cancer.

Most of the examples used in the book are from research papers in which the author of this book is one of the authors. The prerequisite for reading this book is exposure to elementary statistics and probability. No claim is made about the book being self-contained and readers are referred to original papers and other books for mathematical details. This may be discouraging for an uninitiated reader though the author has tried his best to present complex material in a readable fashion.

There are quite a few typographical errors and duplications of words, and also mistakes in citations. The typographical mistakes are not always transparent; for example, 'Uncertainly' should be 'Uncertainty' (page 189), 'chromic' should be 'chronic' (page 195) and 'spend' should be 'spent' (page 200). The references to 'Feinleib (1977)' (page 5), 'Oddoux *et al.* (1998)' (page 19), 'Weinstein *et al.* (1979)' (page 31), 'van Hout *et al.* (1994)' (page 108), 'O'Hagan *et al.* (2000)' (page 108) and 'Wiener *et al.* (1979)' (page 256), *inter alia*, are either missing or need correction. On page 105, 12th line, the meaning of the sentence 'Generate . . . if the first is not to D' is not clear.

The book will be useful to research students in biostatistics as it contains case-studies and also discusses the use of Bayesian inference techniques in medical decision-making. The statistical problems arising in the various case-studies are challenging. Since the case-studies that are discussed are topical, the book is a welcome addition to any undergraduate library in statistics as it gives a flavour of the difficulties that are encountered in medical decision-making.

Ravi Sreenivasan
University of Mysore
and University of Sheffield

Multivariate Permutation Tests: with Applications in Biostatistics
F. PESARIN, 2001
Chichester, Wiley
xxviii + 408 pp., £55.00
ISBN 0-471-49670-7

Most practising statisticians take a particular interest in techniques of data analysis that promise to be of broad applicability. Permutation testing is one such technique that has become much more widespread since software to implement the required intensive computations has been developed. Methods of permutation testing for univariate problems are reasonably well understood and covered extensively in the literature. Applying permutation tests to multiple end points is the subject of this text by Pesarin. The general approach that is proposed here is to apply a suitable permutation test to each individual component of the multivariate problem, and then to combine the individual inferences.

Much of the material here has not been published before and largely represents the author's own research on this subject. Accordingly, many formal proofs and derivations are included, and for a full understanding of the results a reasonably high level of statistical and mathematical knowledge is required. As well as descriptions of the author's methods, there are descriptions of applications to various biostatistical problems, but this is not the main thrust of the book. Despite the author's claims that the book will be of interest to practitioners (and also to undergraduate students), my feeling is that it will in fact be of value mainly to graduate students and researchers who have an interest in permutation testing and in nonparametric methods.

The first six chapters of the book introduce some general principles of permutation tests and also the author's methods for multivariate problems. Chapters 7–12 describe some general and some specific applications of the methodology. Software to implement some of the methods that are described in the book is available via the Internet, and Web site addresses are given in the introduction to the text.

In summary, this is a book for the specialist reader.

L. W. Huson
GNB Limited
Chislehurst

A Guide to First-passage Processes
S. REDNER, 2001
Cambridge, Cambridge University Press
x + 312 pp., £55.00
ISBN 0-521-65248-0

The first-passage problem (called also the first-crossing or first hitting time problem) is among the most challenging problems in stochastic processes. The present book provides a unified account of models of complex random phenomena that are described by stochastic processes and are such that it is of primary importance to know the behaviour of the system until some special event occurs, e.g. failure, ruin or reaching a critical value. Any such a situation leads to a specific first-passage problem. The aim of the author is to use a moderate mathematical background and to present as much as possible of the material that has accumulated in this area.

The author treats a variety of models: discrete models such as random walks and continuous models based on diffusion-type processes. Let us list the titles of the chapters: 1, 'First passage fundamentals'; 2, 'First passage in an interval'; 3, 'Semi-infinite system'; 4, 'Illustrations of first passage in simple geometries'; 5, 'Fractal and nonfractal networks'; 6, 'Systems with spherical symmetry'; 7, 'Wedge domains'; 8, 'Applications to simple reactions'.

Each chapter consists of sections dealing with a particular model (mostly from physics, in particular from electrostatics and current flows in resistor networks) and a first-passage problem followed by the solution. The author provides hints about the derivation of the formulae or solutions and makes comments on their use in particular cases. Among the mathematical tools that are widely used are recursive relations, the Laplace transform, partial differential equations, generating functions and asymptotic analysis of functions. It seems that, to understand deeply the content of the problems and the results, the reader needs good knowledge in several fundamental branches of mathematics. The style of the author is clear and encouraging.

Graduate and post-graduate students in some areas of physics and chemistry can benefit greatly from reading and using this book. Perhaps researchers in the area will find this book useful in their work, either as a reference source or as a source of ideas and techniques that can help them to study other complicated phenomena under randomness.

Jordan Stoyanov
University of Newcastle

Approximation Theorems of Mathematical Statistics
R. J. SERFLING, 2002
New York, Wiley
xiv + 372 pp., £59.50
ISBN 0-471-21927-4

This volume is a paperback version of Serfling's book from 1980. No changes (that I could spot) have been made to the contents. A range of standard

approximate and asymptotic results in mathematical statistics is presented in this text-book aimed at students with some experience in probability theory and mathematical statistics.

Notation, various types of convergence, limit theorems and other tools from probability theory and analysis are introduced in Chapter 1. Chapter 2 concerns sample statistics and their asymptotic representations. Transformations of normal statistics, vectors and multivariate vectors are discussed in Chapter 3. Chapter 4 considers asymptotic theory in parametric inference. Chapter 5 studies U-statistics, and Chapter 6 von Mises differentiable statistical functions. Chapters 7, 8 and 9 review M-estimates, L-estimates and R-estimates respectively. Asymptotic relative efficiency is discussed in Chapter 10.

The strength of this book is that it covers most classical asymptotic results without using advanced mathematics. The weaknesses are examples and exercises: the examples neither seem to clarify the theory nor to illustrate the applicability of the approximate results to practical situations, and the exercises consist mainly of completing proofs of the results that are covered in the book. However, supplemented with extra examples and exercises, this could be a good, if rather expensive, course book for students with no background in measure theory.

Pia Veldt Larsen
University of Southern Denmark
Odense

The Handbook of Parametric and Nonparametric Statistical Procedures, 2nd edn
D. SHESKIN, 2001
Boca Raton, Chapman and Hall–CRC
xxxii + 982 pp., £93.00
ISBN 1-58488-133-X

The author's aim in writing this book is

> 'to provide researchers, teachers, and students with a comprehensive reference book in the areas of parametric and nonparametric statistics'.

The envisaged readership is wide—people from mathematics, statistics, the biological and the social sciences, business and education. The reader is told that

> 'The Handbook of Parametric and Nonparametric Statistical Procedures is designed to be used by those who have a basic familiarity with descriptive statistics and experimental design ... although it is not a prerequisite'.

To remedy any deficiencies that the reader may have, Sheskin provides a lengthy introductory chapter giving an overview of statistics at the level of a first-year service course. The chapter is strong on 'how'. There are no proofs, but there are some interesting heuristic discussions, e.g. on parametric *versus* nonparametric procedures. This introductory material has been extended from 29 to 37 pages and now contains illustrative computations concerning the variance, skewness and kurtosis. It is followed by an outline of the wide range of inferential statistical tests and measures of correlation and association that are described in the book. The chapter ends with guidelines and decision tables concerning the choice of appropriate procedures. These have been revised to accommodate the additional material in the book.

The 26 chapters in the earlier edition have become 32. The additional ones are as follows: the single-sample test for evaluating population skewness; the single-sample test for evaluating population kurtosis; the Kolmogorov–Smirnov goodness-of-fit test for a single sample; the Kolmogorov–Smirnov goodness-of-fit test for two independent samples; the Moses test for equal variability; the van der Waerden normal scores test for k independent samples. Over 70 procedures are now covered.

The underlying philosophy is based on hypothesis testing using classical (non-Bayesian) methods. I am glad that this new edition contains a 15-page overview of computer-intensive methods; although a good set of references is provided, I would not, however, call the coverage in this addendum 'comprehensive'. There is also some material on point estimation, confidence intervals and correlational research. Multiple regression is treated under the heading of measures of association and correlation. The focus throughout is on small data sets—modern methods of analysing large data sets with many regressors (generalized additive models, generalized linear models, etc.) are not covered.

The following standard format is provided for the great majority of the tests: I, hypothesis statement and background information; II, example; III, statement of the null *versus* the alternative hypotheses; IV, test computations; V, interpretation of the test results; VI, additional analytical procedures for the test and/or related tests; VII, additional discussion of the test; VIII, additional examples illustrating the use of the test; IX, an addendum about one or more related tests not discussed in VI; references; end notes.

The appendix of tables at the end of the book now contains three more. It would be helpful to have the tables on an accompanying diskette. Indeed it would be useful to have the whole book

made available on a compact disc, particularly if one tried to adopt the author's suggestion of using the book as a text-book in undergraduate and graduate courses.

This book is a useful and comprehensive compendium of the type of material that was the mainstay of courses in applied statistics before the advent of the specialized packages that have been written for today's powerful computers. Even if students are no longer doing all their calculations by hand, however, they should still be familiar with the *modus operandi* of their computations. If yours is a department in a new institution that has not inherited a statistical library from the 1980s, then this book with its carefully prepared and clearly written content would be a wise investment. And, if your department is not new, still try to add it to your library collection but, because of its weight and the flexibility of its binding, make sure that it is kept in the reference-only section.

Freda Kemp
University of St Andrews

Probability for Statisticians
G. R. SHORACK, 2000
New York, Springer
xviii + 586 pp., £55.00
ISBN 0-387-98953-6

Students of statistics in the UK are unlikely to consider this a text that is suitable for their needs (despite the author's inclusion of 'highly statistical optional' chapters on asymptotics via empirical processes and asymptotics via Stein's approach). And few UK lecturers in statistics will feel able to devote enough lecture hours to enable this to be a suitable course text at any undergraduate or graduate level. The book was written and has been used for graduate probability classes (more than 85 hours) at the University of Washington in Seattle. Three or more hours per week on probability theory is a large amount of time in the UK context, where courses in applied statistics provide very strong competition in nearly all taught Master of Science statistics programmes.

The author claims, nevertheless, that 'There is a thin self-contained textbook within this larger presentation'. The first five chapters cover the measure theoretic material that is needed later in the book. Chapter 6, on topology and Hilbert spaces, is 'presented only for reference'. Chapters 7–9 use statistical theory to introduce some basic probability concepts, including distribution functions, quantile functions, conditional expectation and statistical distribution theory (in matrix notation where appropriate).

The presentation of probability theory begins in earnest in Chapter 10—from here on a 'pick-and-mix' approach to ways of presenting material is possible. Chapter 10 deals with the weak and strong laws of large numbers, the law of the iterated logarithm, the strong Markov property, martingales and maximal inequalities. Convergence in distribution, Brownian motion and characteristic functions are the leading topics in Chapters 11–13. Central limit theorems have already been developed in Chapter 11 via Stein's method; they reappear in Chapter 14 using characteristic function methods. Infinitely divisible and stable distributions are studied in Chapter 15. The highly optional Chapters 16 and 17 are on asymptotics via empirical processes and asymptotics via Stein's approach. Martingales receive coverage in depth in Chapter 18. Chapter 19 is on convergence in law on metric spaces. The appendix covers the gamma and digamma functions and maximum likelihood estimation.

Galen Shorack comments that it is his hope 'that even those well versed in probability theory will find some new things of interest' here. Although I am very doubtful about the possibility of using this book as a course text in the UK, I would like to see it in the hands of lecturers giving probability courses, on the principle that the lecturer's understanding of what he is teaching should be more mature than the level of understanding that he or she is trying to impart.

Freda Kemp
University of St Andrews

Introduction to Nonparametric Item Response Theory
K. SIJTSMA AND I. W. MOLENAAR, 2002
Thousand Oaks, Sage
viii + 168 pp., £18.99
ISBN 0-7619-0813-7

This book is volume 5 in the series 'Measurement methods for the social sciences' and is aimed at students and researchers in the social and behavioural sciences rather than at statisticians. An elementary knowledge of statistics is assumed.

The presentation concentrates on the design and analysis of tests to determine certain properties of a single latent characteristic such as an attitude or an ability. The emphasis is on situations where participants are required to respond to a series of items, usually presented in ascending order of difficulty, and each item is scored 1 for a correct response

and 0 for a nil or an incorrect response. The proposed nonparametric methodology is focused on the total score attained. Extensions of this 0–1 dichotomy to totals of scores for each item on a scale from 0 to p (> 1) are considered in the final two chapters.

A strength of the book is the care that is taken by the authors, obvious enthusiasts for the methods that they are expounding, to point out the difficulties in choosing and deciding the order of presentation of items to be included in experiments to investigate single latent characteristics such as arithmetic ability, verbal ability or attitudes towards a controversial subject such as abortion. Many of the points are illustrated by fully worked examples.

Two aspects of the presentation are less helpful to users than they might be. The first is the abbreviation of names of 13 of the most commonly used terms to initials. Thus, an item response function becomes an IRF and an item response test an IRT. Given this information you might or might not guess what an NIRT is, but I had to refer several times to the list of (so-called) acronyms in the appendix before readily recalling that IIO meant invariant item ordering. Use of such jargon by consenting adults regularly working in the area may have merits, but it is not helpful to newcomers.

The second shortcoming concerns the ordering of some of the material. In particular, I found much of that on scalability in Chapter 3 impossible to understand until I had read Chapter 4. The brief 'theoretical side-step' on the topic on page 36 does not explain how to compute the scalability coefficients that are used in Chapter 3—that is only explained in Chapter 4.

The authors justify their emphasis on the investigation of one latent characteristic at a time on the grounds that this is what many researchers in education and psychology do. Although there may be good reasons to use this approach I would have thought that often the simultaneous investigation of two or more latent traits, and how they relate to each other, would be a more fruitful field of experimentation. However, within the constraints imposed the methods described in this monograph are clearly useful.

In summary, the book is a good introduction for social scientists to important practical aspects of item response analysis and may also prove a useful reference for statisticians who are interested in, but have no specialist knowledge of, this type of investigation.

P. Sprent
Wormit

Statistical Inference in Science
D. A. Sprott, 2000
New York, Springer
xvi + 246 pp., £48.00
ISBN 0-387-95019-2

Many undergraduate and post-graduate programmes in statistics expect students to carry out an extensive and examinable project in applied statistics in their final year. Professor Sprott has had many years' experience teaching at the University of Waterloo, Ontario, and at the Centro de Investigacion en Matematicas, Guanajuato, Mexico. Some of his students have become statistical innovators; many are now statistical practitioners. This book is especially suitable for the latter group of students.

It has a slightly old-fashioned air, with its use of relatively small data sets (many drawn from medical research) and its avoidance of reference to any statistical computer package. Nevertheless it is ahead of much current statistical practice with its emphasis, not so much on using sample information to obtain estimates, their marginal distributions and their optimal properties, but instead on methods for separating sample information into components that address different aspects of an investigation. For example, the division of the sample information for a Poisson model with $t = \Sigma_{i=1}^{n} y_i$ and assumed mean $n\theta$ is given as

$$f(\{y_i\}; \theta) = \prod \frac{\theta^{y_i} e^{-\theta}}{y_i!} \equiv \left[\frac{(n\theta)^t e^{-n\theta}}{t!}\right] \cdot \left[\frac{t!}{\prod y_i!} \prod \left(\frac{1}{n}\right)^{y_i}\right],$$

where the first factor enables (marginal) inferences about θ to be made, whereas the second factor is the multinomial consequence of the Poisson model; if the data cast doubt on the multinomial distribution, then the Poisson model and consequently the inferences about θ from the first factor are also in doubt.

Professor Sprott begins with a discussion of repeatable experiments in Chapter 1. In Chapter 2 he describes at length, with many examples, the use of the likelihood function and relative likelihood. Chapters 3 and 4, on the division of sample information, form the core of the book. Conditional likelihood, marginal likelihood, pivotal likelihood and profile likelihood all receive attention. The next four chapters contain material on estimation statements, tests of significance, the location–scale pivotal model and the Gauss linear model. Notes and references for Chapters 1–6 appear at the end of Chapter 6; notes and references for Chapters 7 and 8 are at the end of Chapter 8.

Maximum likelihood estimation is the topic in Chapter 9. The author's approach, unlike that of Barndorff-Nielsen and Cox's *Inference and Asymptotics*, 1994 (section 8.4), is

> 'to obtain a linear pivotal that is efficient when assumed to have a $N(0, 1)$, t, or $\log F$ distribution, that is, the resulting pivotal likelihood will approximate the observed likelihood function'.

The chapter has notes and references and three appendices; they are on degrees of freedom of $\log(f)$, observed information of the profile likelihood and symmetrizing the likelihood function.

Chapter 10 is on controlled experiments. It ends with a cautionary note concerning randomization and random assignment. All the problems for students are contained in Chapter 11:

> 'It seemed that putting them at the end of a given chapter would limit and prejudge how they could be approached. In addition ... [some] problems would appear at the end of a number of chapters.'

There are no hints or solutions.

The book is strong regarding the large number of illustrative examples that use data from published scientific papers. The presentation of analytical findings is taken seriously, with much emphasis on the use of likelihood plots. For detailed treatments of generalized additive models, generalized linear mixed models, etc., a more specialized text is needed.

This book is intended not only for class teaching at advanced undergraduate and post-graduate levels but also as a reference text. Supervisors of applied statistics projects should ensure that their students have access to it. It is as least as essential for their departmental libraries as books on 'Statistical inference for scientific data using the package XYZ'.

A. W. Kemp
St Andrews University

Biometrika: One Hundred Years
D. M. Titterington and D. R. Cox (eds), 2001
Oxford, Oxford University Press
viii + 384 pp., £45.00
ISBN 0-19-850993-6

Over its first 90 years, *Biometrika* essentially had three editors only. The first issue appeared in October 1901, and Karl Pearson, with early assistance from Weldon, Galton and Davenport, was in charge until his death in 1936. His son Egon occupied the position for 30 years, David Cox served from 1966 to 1991, but Mike Titterington is already the third editor since Cox.

The first 216 pages of this celebration are a reprint of papers that were commissioned and published as the first issue of the journal in 2001. The intention was to describe the impact of *Biometrika* across diverse fields, and the choice of authors to make these surveys is admirable. Cox offers a history of the journal, Anthony Davison reviews its influence on statistical theory, Anthony Atkinson and Rosemary Bailey share the duties for experimental design and David Oakes assesses the papers on survival analysis. Peter Hall carries the torch for nonparametric statistics, Fred Smith covers sample surveys and Howell Tong has sifted through the time series work.

These accounts have plainly been constructed with great diligence. Accepting the invitation to hunt through a century of journal issues, and to assess the influence of all papers in a given area, could not have been a light decision. The authors have also described instances of interaction, when work begun in *Biometrika* was further developed elsewhere, and vice versa, but about 80% of all papers cited have this journal as their origin.

There are gems tucked away among the scholarly accounts. Atkinson and Bailey offer evidence of the difficult personal relationships between Fisher and Karl Pearson to which Cox alludes. Davison notes Wilks's prediction, when reviewing Jeffreys's *Theory of Probability*, that it would persuade few practitioners to change their approach. Smith remarks that ways of obtaining without-replacement samples with probability proportional to size were described in the 1960s but were felt unusable because of their computational complexity, an objection that no longer carries weight. He also points out that the journal has plainly not developed in one of the ways that its founders envisaged, that of helping to make the evolutionist a sort of registrar-general for all forms of life: hardly any work on sampling in biology has appeared. Oakes describes how Pearson himself, in the very first issue of the journal, overlooked a trap in comparing the mean lifetimes of older and younger siblings. Hall speculates on the future: that parametric and nonparametric techniques will

> 'work in tandem, as aspects of a single methodology for the quantification of information'.

He believes that the computer-associated developments of the last 25 years have seen changes in nonparametric statistics that are comparable in importance with changes in the second quarter of the last century when the foundations of parametric statistics were laid down. Tong also looks to the future: he calls for developments in time series to be linked

more closely with general work in stochastic processes. He also remarks that, in the list of 61 'breakthroughs in statistics' compiled by Kotz and Johnson, seven appeared in this journal. That seems an honourable proportion.

The rest of the book consists of reprints of 10 papers, from the journal's first 70 years, chosen as significant landmarks in statistics. Excluding the more recent work has meant that the papers selected have had time to demonstrate real impact on the subject. (Contrast this perspective with that of the current 5-yearly research assessment exercises, which seek to judge the worth of research almost as soon as it has appeared.) Some long papers that were strong candidates on grounds of importance have been omitted for brevity: Harold Hotelling on canonical correlations, the series of papers associated with the Durbin–Watson test and Kirstine Smith's 1918 experimental design paper, '30 years before its time', are among the unfortunates. Seven of the papers selected come from 1950–1971, nicely complementing the two neighbouring quarter-centuries picked out by Hall as important for other reasons.

Egon Pearson's perspective on the interactions between his father, Fisher and Gosset, and Maurice Quenouille's short paper, that developed the jackknife which he had introduced elsewhere, are both included. So is Frank Yates's thoughtful 1939 design paper, full of real data and arguing that the object of most agricultural experiments is the estimation of the magnitude of treatment effects, and their standard errors, not just testing for their existence. Curiously, although Henry Daniels's 1944 paper on correlation appears here, I find no reference to it in the seven commissioned surveys. Statistics within this journal is even broader than might have been supposed.

Because the entire contents of this book are already available in *Biometrika*, albeit scattered, personal sales may be limited. But libraries should be encouraged to acquire a copy, as many statisticians will wish to read the authoritative overviews, or to have easy access to the older reprinted seminal papers.

It is a happy academic tradition to publish a *Festschrift* to mark a significant birthday in the life of a distinguished scholar. Many such scholars continue to publish valuable work long after being honoured in this fashion; and we can expect *Biometrika* serenely to accept this present tribute, and to continue its role in the development of statistical theory and practice.

John Haigh
University of Sussex
Brighton

Conquering Statistics: Numbers without the Crunch
J. H. WEAVER, 2000
Cambridge, Perseus
xii + 236 pp., £12.99
ISBN 0-7382-0495-1

This book left me with a disturbing thought. Either there is or is not a worse way to begin learning about our subject than by reading *Conquering Statistics*. If there is, I fear for the readers of that book (or students in that course). If as I suspect there is not, perhaps I might save others a few hours and much irritation.

Conquering Statistics is intended to convey the great ideas of statistics to a general audience. I liked the selection of topics—samples, populations, probability, variability, the normal curve, the law of large numbers and statistical significance to name a few—but found the breeziness of the writing annoying, from the subtitle onwards. Sentences like

> 'The real tragedy of [the neglect of statistics] is that statistics, like broccoli, is a very good thing even though both may have a slightly malodorous aftertaste'

(page 3) are fine in a Dave Barry newspaper column but become tiresome well before the end of a 248-page book. It is unfortunate that statistics is neglected, but the book's fabricated illustrations (quality control at the Zippy Cola Company, anyone?) will not convince many of the tragedy. The conspicuous exceptions (some of the difficulties in estimating the attendance at the 1995 Million Man March in Washington) prove the rule.

Add to the awkwardness of style many errors of fact. Histograms are not 'bar-charts' even if they have bars (page 46). Subtracting the average score of the class from Bruno's examination score and dividing by the number of students in the class does not tell you how many standard deviations above the mean Bruno was (page 64). The chance of obtaining an even 50–50 split between heads and tails does not increase with the number of coin tosses (page 71), as an exercise in one of Weaver's cited references (Freedman *et al.*, 1978) makes clear. And so on.

Speaking of Freedman *et al.* (1978)—pick up a copy of that book, or another outstanding introductory statistics text, to obtain an accurate exposition of the ideas mangled in *Conquering Statistics*. Or try Gigerenzer's (2002) *Reckoning with Risk* (published as *Calculated Risks* in the USA) for a general interest introduction to statistical thinking. But, however you choose to conquer statistics, stay away from *Conquering Statistics*.

References

Freedman, D., Pisani, R. and Purves, R. (1978) *Statis-*

tics, 1st edn. New York: Norton.
Gigerenzer, G. (2002) *Reckoning with Risk: Learning to Live with Uncertainty*. London: Penguin.

Wesley D. Eddings
Johns Hopkins University
Baltimore

Elementare Grundbegriffe einer Allgemeineren Wahrscheinlichkeitsrechnung, vol. I, Intervallwahrscheinlichkeit als Umfassendes Konzept
K. WEICHSELBERGER, 2001
Heidelberg, Physica
xiv + 684 pp., £57.00
ISBN 3-7908-1411-3

This book, currently unfortunately only available in German, provides a detailed presentation of interval probability. In the preface, two further volumes are announced which are currently in development. The second volume will mainly focus on concepts of independence and conditional interval probability; the third volume will consider general sample spaces and basic theory of corresponding statistical methods. These two volumes will be very important for a complete and useful theory; I hope that they will appear soon. This first volume introduces interval probability, with key concepts and motivations clearly presented. A concise overview of some key concepts has been presented by the author in an English language paper (Weichselberger, 2000).

Central to the theory is the following generalization of Kolmogorov's axioms to interval-valued probabilities. For a σ-field $\mathcal{A}$ of random events in a sample space Ω, a set function $p(\cdot)$ defined on $\mathcal{A}$ is named a K-function (or K-probability or classical probability), if it obeys the three axioms of Kolmogorov. Additional axioms for an interval-valued set function $P(\cdot)$ are

(a) $P(A) = [L(A);\ U(A)], \quad \forall A \in \mathcal{A}, \quad$ with $\quad 0 \leqslant L(A) \leqslant U(A) \leqslant 1, \forall A \in \mathcal{A},$
(b) the set $\mathcal{M}$ of K-functions on $\mathcal{A}$ with $L(A) \leqslant p(A) \leqslant U(A), \forall A \in \mathcal{A}$, is not empty and
(c) $\inf_{p \in \mathcal{M}}\{p(A)\} = L(A)$ and $\sup_{p \in \mathcal{M}}\{p(A)\} = U(A), \forall A \in \mathcal{A}.$

An interval-valued set function which satisfies axioms (a) and (b) is called R-probability (for 'reasonable'), or F-probability (for 'feasible') if it also satisfies (c). Straightforward properties of F-probability, which do not need to hold for R-probability, are $U(A) = 1 - L(A^c),\ \forall A \in \mathcal{A},\ U(\emptyset) = 0$ and $L(\Omega) = 1$. The set $\mathcal{M}$ in (b) is called the 'structure' of the R-probability field, which is a fundamental concept in this theory. In addition, an important role is played by the concept called 'prestructure': any set $\mathcal{J}$ of K-probabilities is named a prestructure of an F-probability field if $\inf_{p \in \mathcal{J}}\{p(A)\} = L(A), \forall A \in \mathcal{A}$. Prestructures can be useful for calculations, and in situations where sets of classical probabilities are used without the properties of a structure, e.g. as in many often used models in robust (Bayesian) statistics.

After preparing for this theory in Chapter 1, which also includes a concise history of generalized probabilistic concepts with discussion of their motivations, the main concepts are defined, together with many interesting results in Chapter 2, for situations where all interval probabilities of interest are actually determined. This chapter also includes a large section on decision-making under F-probability, focusing on order relations for events in terms of F-probability, and including a detailed analysis of Ellsberg's paradox and an interesting example on insurance, a natural area of application of interval probability, with differences between upper and lower probabilities related to risk averseness of individuals and profit margins of insurance companies. Although the main concepts presented in this book are generally applicable, most of the development and presentation is with restriction to finite sample spaces, on which the concept of R-probability coincides with Walley's (1991) 'avoiding sure loss', and F-probability coincides with Walley's 'coherence'. A main difference from Walley (1991) is that Walley explicitly generalized the subjective framework of de Finetti to allow for different buying and selling prices for bets on an uncertain event A and—presupposing minimax strategies—related $L(A)$ directly to the highest buying price for a bet on A and $U(A)$ to the lowest selling price. Weichselberger's book explicitly does not build the theory from a particular interpretation of probability but takes a fully axiomatic approach and studies all consequences. Therefore it also allows for personal strategies deviating from minimax.

Chapter 3 considers partially determined interval probability, where the interval probabilities are not actually provided for all events. An interesting and relevant case appears with partial specifications which generalize the classical cumulative distribution functions. Chapter 4 presents results which particularly hold for finite sample spaces, to emphasize this material before more general sample spaces are considered in the later volumes. There is detailed attention to linear optimization for identification and construction of R- and F-probability fields. Classical statistical methods often build on a concept of equal probability for events, following the 'principle of insufficient reason'. This concept of symmetry is generalized, leading to important differences between 'K equal probability', which is suit-

able in cases of physical symmetry, and '*R* equal probability', which is suitable in cases of epistemic symmetry. The book ends with several appendices, presenting some further theory and a brief look at the use of upper and lower density functions, which provide important methods for statistical models, to be discussed in more detail in the third volume.

This is an exceptional book which offers a wide range of new concepts and related theoretical results. The relevance of many of the concepts may only become apparent after a detailed study, which is enabled by the careful descriptions and motivations and clear examples. The timing of this book is excellent, as many researchers in statistics and other fields, most noticeably perhaps artificial intelligence, have been developing theory related to generalized probabilities along these lines for several years now (for example see `www.sipta.org`). This theory promises to play the same role for these important developments as Kolmogorov's work played for modern statistics and probability theory, so it is crucial that an English language version of this book becomes available soon, as well as English versions of the two further volumes that the author has announced.

References

Walley, P. (1991) *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

Weichselberger, K. (2000) The theory of interval-probability as a unifying concept for uncertainty. *Int. J. Approx. Reasng*, **24**, 149–170.

F. P. A. Coolen
University of Durham

First Steps in Statistics
D. B. WRIGHT, 2002
London, Sage
x + 148 pp., £15.99
ISBN 0-7619-5163-6

I frequently recommend introductory statistics books to people who have diverse backgrounds and minimal statistics knowledge. In general these texts tend to be longer than I would like them to be and usually I end up giving some of my introductory statistics notes instead. On seeing that this book was only 147 pages long I felt that it was worth reviewing to see whether I would put it on my recommendation list.

As there is no preface it is difficult to say to whom this text is aimed. Instead I shall quote from the back cover. The text aims to offer a

> '... comprehensive introduction to the most common statistical techniques taught on first year research courses ...'.

The author aims to use his '... typical friendly and humorous style' using 'real-life' examples to

> 'guide the student comfortably through their first experience of doing basic statistics at undergraduate or postgraduate level'.

So in summary it is aimed at students who are about to experience statistics for the first time and they will be aided by the author's informal approach.

The text covers summary statistics (central tendency and spread), graphical methods for looking at data distributions (e.g. histograms and scatterplots), random sampling, interval estimation, hypothesis testing, analysis of variance, simple linear regression (and correlation), simple nonparametric methods (e.g. Wilcoxon) and contingency tables. In summary, this text seems to offer what most introductory texts offer with the exception of probability.

The author uses informal language to explain statistical techniques. The two main approaches appear to be to motivate a technique by using an example and then to lead the reader into the methodology or to explain the concepts of a technique and then again to use an example to lead the reader into the methodology. Graphs and tables are used extensively to illustrate the presentation of data and interpretation. To illustrate some specific points boxes highlighted in grey are used with a detailed example. All chapters end with a summary and exercises for the reader. Answers to the exercises are not given in the text but on the back cover it is stated that these can be obtained from the publisher.

The author uses a variety of examples and exercises that include real life situations, real life examples and some imaginative examples. The real life exercises and examples are primarily driven by the author's personal interests in psychology and perhaps in quantum physics. An example of an imaginative example uses combinations of two pizza toppings to illustrate populations and sampling. There are five toppings so the population contains 10 possible pizzas. Simple random sampling is described in the context of randomly selecting from the pizza population which appeared to be odd or imaginative as choosing a pizza is not really a random process. The author's style is to use many examples even when explaining one given technique. For example, to calculate the mean the author uses the height of children and rainfall data, and for the median he uses the height-of-children data, the rainfall data and salary data. For the mode (and proportions) he uses data from a question with categorical response. So for three measures of central tendency (the mean, median and mode) he uses four different examples.

Most statistical techniques that can be expected

in a standard introductory text are presented. I was concerned that the normal distribution is given minimal attention in favour of the use of *t*-tests for all tests involving continuous data. Several times the reader is referred back to Chapter 2 for the following sole explanation of the normal distribution.

> '"Bell-shaped" curves have a special place in statistics, so it is worth briefly mentioning them here. Figure 2.4 shows what is called a normal distribution (sometimes called the Gaussian distribution). It is shaped like one of those old-fashioned bells. In many of the statistical tests covered later in this book, it is assumed that they follow roughly this distribution. In reality most do not (Micceri, 1989).'

Another notable omission is the binomial distribution. This is noticeable when the numbers for a single binary variable are tested for equality by using the χ^2-test. Such examples are usually natural candidates for testing proportions using the binomial distribution.

There are some errors in the text. For example the formula for an odds ratio is stated as AD/BD rather than AD/BC. This error is repeated in the accompanying text and a nearby table. Another example is in Chapter 8: graphs of the data that are used to illustrate simple linear regression do not correspond for one of the outlying points.

In summary the author tries to use everyday language to explain statistical methods by using different examples and introducing the algebra after this. The use of long (and sometimes complicated) examples and the frequent change of examples does not make the text flow naturally. Using a smaller number of simple examples and introducing some of the algebra earlier could have helped the flow. Trying to keep algebra out of the explanation and not always using standard notation is not good practice. This could lead to the reader being at odds with other statistical texts if they should wish to learn more.

Before my final comments on the text I would like to quote from the text:

> 'Sometimes I read manuscripts where the authors try to use the longest words possible, and present their ideas in the most complicated manner. These are signs of poor writing.'

In applying this judgment to this text I would replace 'longest words' with 'longest examples'.

The aim of adding this text to my list of recommend texts for introductory statistics has not been met and hence I cannot recommend it. The author aims to use many examples to help to explain the methodology, which is natural for this type of text, but this has not been achieved. Primarily this is because examples change frequently (including during an explanation of methodology) and can be more difficult to understand than the methodology. It is also because examples may describe scenarios that are not naturally associated with the statistics. A couple of other major points in not recommending this text are the lack of detailed discussions about the normal and binomial distributions, and my concerns relating to the errors that I spotted (and those that I did not).

Saghir Bashir
SBTC
Cambridge

This review of Dennis Helsel's book on "Statistics for Censored Environmental Data Using Minitab and R, 2nd edition" is provided in the Book Review section of Integrated Environmental Assessment and Management. Helsel's book demonstrates why substitution methods are inappropriate, in most cases, for analyzing nondetect data, and he provides better statistical techiques to use in these cases. Read more. Article. Review of the book: "Category Theory". Pedro Pascual Gainza. Book Review Category Theory, Second Edition Steve Awodey Oxford Logic Studies 52. Writing a book review is a great way to let fellow readers know about an exciting new page-turnerâ€”or give a heads up that a book might not meet expectations. Whether youâ€™re reviewing a book on a site like Goodreads or on your personal blog, youâ€™ll want your review to be informative and helpful for your audience. Read on for our essential tips on how to write an engaging book review. Hereâ€™s a tip: Want to make sure your writing always looks great?Â Sharing these details will help your audience form their own opinion of whether they would enjoy reading the book. For exampleÂ Imagine if it were your book being reviewed, how would you want a reader to express their critique? What are your favorite tips for writing a great book review? Books for review If you would like to review a book, and thereby to retain it for your collection, please contact the Book Reviews Editor, whose details can be found by clicking on â€˜books available for reviewâ€™ in the information on the Royal Statistical Societyâ€™s Web site: http://www.rss.org.uk/main.asp?page=1977. Reviews in this issue Atkinson, A. C., Riani, M. and Cerioli, A. Exploring Multivariate Data with the Forward SearchÂ 458 Book Reviews. Exploring Multivariate Data with the Forward Search A. C. Atkinson, M. Riani and A. Cerioli, 2004 New York, Springer xxiv + 622 pp., Â£69.00 ISBN 0-387-40852-5. This is a companion to an earlier book (Atkin-son and Riani, 2000), both of which feature many informative graphs.