

Book Reviews

This section will review those books whose content and level reflect the general editorial policy of *Technometrics*. Publishers should send books for review to Eric R. Ziegel, BP Naperville Complex, Mail Code C-7, 150 West Warrenville Road, Naperville, IL 60563-8460. (ziegeler@bp.com)

The opinions expressed in this section are those of the reviewers. These opinions do not represent positions of the reviewer's organization and may not reflect those of the editors or the sponsoring societies. Listed prices reflect information provided by the publisher and may not be current.

The book purchase program of the American Society for Quality can provide some of these books at reduced prices for members. For information, contact the American Society for Quality, 1-800-248-1946.

Practical Strategies for Experimenting

G. K. Robinson *Ned M. Gibbons* 486

Reliability: Modeling, Prediction, and Optimization

Wallace R. Blischke and
D. N. Prabhakar Murthy *David Howell* 486

Applied Statistics for Engineers and Scientists

Jay Devore and Nicholas Farnum *Diane K. Michelson* 487

Reliability Theory: With Applications to Preventive Maintenance

Ilya Gertsbakh *M. Tortorella* 487

Statistical Modelling With Quantile Functions

Warren Gilchrist *Jeffrey Glosup* 488

HALT, HASS and HASA Explained: Accelerated Reliability Techniques

Harry McLean *Necip Doganaksoy* 489

Conditional Specification of Statistical Models

Barry C. Arnold, Enrique Castillo,
and José Maria Sarabia *M. Bhaskara Rao* 490

Probability Applications in Mechanical Design

Franklin E. Fisher and Joy R. Fisher *Lianfen Qian* 490

Wavelet Methods for Time Series Analysis

Donald B. Percival and
Andrew T. Walden *Michael R. Chernick* 491

Nonlinear Time Series Analysis

H. Kantz and T. Schreiber *Katherine Campbell* 491

Multivariate Statistical Methods: A First Course

George A. Marcoulides and
Scott L. Hershberger *Steven M. LaLonde* 491

Practical Geostatistics 2000

Isobel Clark and William V. Harper *Donald E. Myers* 492

Robust Bayesian Analysis

David Rios Insua and Fabrizio Ruggeri (editors) *S. K. Katti* 493

Analysis of Customer Satisfaction Data

Derek R. Allen and Tanniru R. Rao *Ashok K. Singh* 493

The Process Evaluation Handbook

Donald J. Wheeler *John D. Barrett* 494

Measuring the Software Process

William A. Florac and Anita D. Carleton *Marcey Abate* 495

Environmental Statistics With S-PLUS

Steven P. Millard and Nagaraj K. Neerchal *Thomas Lumley* 495

Mastering Data Mining

Michael J. A. Berry and Gordon S. Linoff *Eric R. Ziegel* 496

A Handbook of Statistical Analyses Using Stata

Sophia Rabe-Hesketh and
Brian Everitt *Ronald D. Fricker, Jr.* 496

The Theory of the Design of Experiments

D. R. Cox and N. Reid *Dechang Chen* 497

Editor Reports on New Editions, Proceedings, Collections, and Other Books

Introduction to Statistical Quality Control

Douglas C. Montgomery 497

SAS® System for Regression

Rudolf J. Freund and Ramon C. Littell 498

Principles of Multivariate Analysis

W. J. Krzanowski 498

Visualizing Categorical Data

Michael Friendly 498

Geostatistics for Environmental Scientists

Richard Webster and Margaret A. Oliver 499

Statistical Methods in Spatial Epidemiology

Andrew B. Lawson 499

A Course in Time Series Analysis

Daniel Peña, George C. Tiao, and Ruey S. Tsay (editors) 500

Seasonal Adjustment With the X-11 Method

Dominique Ladiray and Benoît Quenneville 500

Optimization Heuristics in Econometrics

Peter Winker 500

Customer Centered Six Sigma

Earl Naumann and Steven H. Hoisington 500

Forthcoming Reviews

Practical Strategies for Experimenting, by G. K. Robinson, West Sussex, U.K.: Wiley, 2000, ISBN 0-471-49055-5, vii + 265 pages, \$65.

The author states in the preface that he believes "the non-statistical aspects of planning and conducting experiments are more important than the formal design and analysis." Furthermore, his book "concentrates on questions for which there are no uniquely correct answers." He follows this direction, in general, throughout his book, giving a broad range of examples that support his belief. This book is most beneficial for an experienced practitioner.

The table of contents outlines his process.

Chapter 1: Introduction
Chapter 2: Clarify the Objective
Chapter 3: Summarize Beliefs and Uncertainties
Chapter 4: Decide on a Strategy
Chapter 5: Plan a Single Experiment
Chapter 6: Design the Experiment
Chapter 7: Collect the Data
Chapter 8: Update Beliefs and Uncertainties
Chapter 9: Revisit the Objective
Glossary
References
Index

In the Introduction, the author gives his own version of the sequential nature of deduction and induction as it applies to scientific progress and, in particular, a circle-within-a-circle checklist for sequential experimentation. An alternative title for Chapter 8 might be "Analysis and Interpretation." The Glossary omits explanations, for example, of confidence, coding, and significance. The References are extensive and up to date.

Each chapter section has observations drawn from personal experience. The author asks specific questions that need to be answered so that the process runs smoothly. These are generally followed by one or more examples that illustrate his point. Examples are based on the literature or based on his own or his colleagues' work at the Commonwealth Scientific and Industrial Research Organisation in Australia (www.csiro.au). For readers of *Technometrics*, there are few nonchemical industrial examples.

The format is unusual since this book is an outgrowth of a manual. Chapter sections are in bold, subsections are italicized, and many bullet items are spread throughout. A few sections seem to be still in draft form. Continuity is jagged at times. There are some incomplete sentences. Typos are rare. Spelling is British—for example, "programme" instead of "program."

A few sections, from my viewpoint, need more attention. The author could go into more detail in Section 3.1 in discussing if an experimental design is necessary. Perhaps, this could be done by naming typical, nonstatistical alternatives, such as failure modes and effects analysis for reliability applications. He is distracted from his purpose in Section 4.7 by discussing at length his own "simulation game." He evaluates four (statistical) strategies in determining the best approach to finding a simulated, nonlinear global optimum. In Section 8.3, he seems to be unaware of factor coding; he centers factor levels but does not scale. He omits in Section 8.6 a discussion of "practical significance" versus "statistical significance." This is hinted at in Chapter 9 but only as a bullet item. There are no problems. Each chapter might have some "what ifs" or "discussion points" as thought provokers.

This book is not an introductory textbook. It assumes that the reader is an experienced practitioner who has encountered some or many of the problems discussed. The author suggests that it might be read "during the planning of a substantial experimental programme checking that important issues are not being overlooked." It should also be valuable as a supplementary textbook in a consulting course.

Ned M. Gibbons
General Motors Corporation

Reliability Modeling, Prediction, and Optimization, by Wallace R. Blischke and D. N. Prabhakar Murthy, New York: Wiley, 2000, xxvii + 812 pp., \$89.95.

This is an excellent book, distinguished by excellence of exposition, breadth and comprehensiveness of topical coverage, relevance, number and impor-

tance of examples, depth of references, and quality of exercises. This will become one of the standard references on reliability for engineers, operations managers, and applied statisticians.

This book is laid out in five parts: "Context of Reliability Analysis," "Basic Reliability Methodology," "Reliability Modeling Estimation and Prediction; Reliability Management, Improvement, and Optimization," and an Epilogue. There are 20 chapters and four appendixes.

"Context of Reliability Analysis" contains two chapters. The first chapter, an overview, opens the book with general examples, historical background, and the multiple perspectives of reliability. The second chapter presents 27 reliability cases. These cases provide the motivating examples for the subsequent development of material. The data for each case are real, and many of the cases are previously unpublished. These cases are relevant, important, and well developed.

"Basic Reliability Methodology" contains the next three chapters. They present the basics of data collection, data description, common and not-so-common reliability models, and basic statistical analysis. The chapters are well written and topical. For example, I found the section on classifying distributions by failure-rate behavior short but very much to the point. There is a good but brief discussion of maximum likelihood and moment estimators, but no discussion of the frequently used rank regression technique. There was a very good coverage of the usual life-data distributions and many of the more exotic models, such as the inverse Gaussian, Gumbel distributions (with truncation), and exponentiated Weibull.

"Reliability Modeling, Estimation and Prediction" constitutes the heart of the book, comprising Chapters 6–11. There is an excellent chapter on modeling component failures, with a very nice description of failure mechanisms and their modeling implications, including environmental effects. The chapter on multicomponent systems includes reliability block diagrams, fault trees, FMECA analysis, Markovian models, and a discussion of failure dependency and interactions. The chapter on advanced statistical analysis covers grouping, censoring, Bayesian concepts, aggregated data, and system models. Software reliability deserves and gets a separate chapter, including discussion of design, modeling, and testing of software reliability. The chapter on design of experiments and analysis of variance is also well written, but the exclusion of censored data from the discussion limits the usefulness of the material for reliability applications. The chapter on model selection and validation covers both graphical and goodness-of-fit methods. The discussion of the selection of a life distribution aptly characterizes the mixture of art and science involved and offers a useful checklist.

"Reliability Management, Improvement, and Optimization" is the fourth part of the book, consisting of Chapters 12–18. Reliability management from a qualitative perspective is presented in Chapter 12. This includes sections on strategic issues as well as a nice summary of the various government and commercial reliability programs and their standards. Warranties, which are discussed in detail in Chapter 17, are presented briefly. "Reliability Engineering," Chapter 13, covers the engineering process for the design, control, and testing of reliability. There are many useful checklists and flowcharts to support the description of this process. Chapter 14 covers reliability prediction and assessment and includes a short description of accelerated testing. Chapter 15 covers reliability improvement. There is a good discussion of redundancy and reliability growth, including all of the common reliability growth models and switching issues in standby systems. Maintenance of unreliable systems receives a quick treatment in Chapter 16, with metrics based on cost and availability, supported by the appropriate renewal equation analysis. This leads to "Warranty and Service Contracts," Chapter 17. This is an excellent discussion of different warranty and service policies, supported by stochastic and economic analysis. These issues in turn lead to a general discussion of reliability and optimization, Chapter 18, covering optimal reliability allocation, optimization during manufacturing, optimal warranty policies, and optimal maintenance policies.

"Epilogue" includes two complete case studies in Chapter 19 and an extensive list of resource materials in Chapter 20. Software manufacturers will want to see how their products are described in the section on reliability software. There are four terse appendixes, covering probability, stochastic processes, statistical tables, and stochastic optimization. Twenty-six pages of references (over 500) are included.

The exposition is uniformly excellent—clear, concise, and supported by high-quality graphics, mostly generated by Minitab. I found very few typographical errors for a work of this size (kudos to the Wiley support staff and the authors). The price is very reasonable, especially compared to some other

recent offerings by competing publishers. There are numerous and thought-provoking exercises—457 by my count. Many are based on the example cases from Chapter 2.

The basic statistical groundwork for reliability textbooks has been well laid by Leemis (1995), Lawless (1982), and Nelson (1982), among others. Meeker and Escobar (1998) have, in my opinion, the best general survey. Recent books like that by Modarres, Kaminskiy, and Krivtsov (1999) are extending the material into new, focused applications. The reviewed book does not compete with the 1998 book by Meeker and Escobar, which it cites often and which covers the statistical issues completely. Instead, it places reliability analysis in a much broader and more complete framework. It is so comprehensive that it would provide material for more than a two-semester graduate sequence in reliability. It integrates several threads of operations research—statistics, probability, stochastic processes, and optimization—and it will find a warm welcome in the best graduate programs. I strongly recommend it.

David Olwel
U.S. Navy

REFERENCES

- Leemis, L. M. (1995), *Reliability: Probabilistic Models and Statistical Methods*, Englewood Cliffs, NJ: Prentice-Hall.
- Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: Wiley.
- Meeker, W. Q., and Escobar, L. A. (1998), *Statistical Methods for Reliability*, New York: Wiley.
- Modarres, M., Kaminskiy, M., and Krivtsov, V. (1999), *Reliability Engineering and Risk Analysis*, New York: Marcel Dekker.
- Nelson, W. (1982), *Applied Life Data Analysis*, New York: Wiley.

Applied Statistics for Engineers and Scientists, by Jay Devore and Nicholas Farnum, Pacific Grove, CA: Brooks/Cole, 1999, ISBN 0-534-35601-X, xiv + 577 pp. + CD, \$81.95.

This book is an excellent textbook for a one-semester statistics course for students in engineering or the natural sciences. It consists of 11 chapters, statistical tables in an appendix, answers to odd-numbered exercises, and an index. The book follows the current practice of beginning with data, summary statistics, and graphs, then proceeding to probability and distributions, and continuing with confidence intervals and hypothesis testing, regression, and analysis of variance. Students are expected to be familiar with integral calculus, but calculus is not required for an understanding of the book.

A CD is included that contains data for all the examples and exercises in Excel, SPSS, JMP-IN, and MINITAB format. Installation was easy and fairly quick. I was able to look at the data in both Excel and JMP. Some data manipulation may be required because the data are not always in standard format, with columns as variables and rows as observations. However, data in real life do not come in standard format, and data manipulation in software packages is one of the first things an engineer needs to learn.

The chapter titles are as follows:

1. Data and Distributions
2. Numerical Summary Measures
3. Bivariate and Multivariate Data and Distributions
4. Obtaining Data
5. Probability and Sampling Distributions
6. Quality Control
7. Estimation and Statistical Intervals
8. Testing Statistical Hypotheses
9. The Analysis of Variance
10. Experimental Design
11. Inferential Methods in Regression and Correlation

The authors provide a bibliography after each chapter, including a short statement about each book. They include standard textbooks that are readily available.

I was pleased with the inclusion of many topics that engineers using statistics need to know but that are not usually included in a one-semester course. These topics include measurement systems analysis, sampling distributions, control charts, and design of experiments, including factorial and fractional

factorial designs. Given the time restraints of a one-semester course, I would expect the treatment of most of these topics to take no longer than a day or two. However, this book is an excellent reference to the basics of most statistical techniques that are used by scientists and engineers. I hope students will use this book as a reference in their future studies and work.

It is important for students of statistics to see how data are collected as well as how they are analyzed. The authors meet this requirement by including many real-world examples with citations. They also use output from different software packages throughout the book, including MINITAB, S-PLUS, SAS, and JMP.

I did notice one slight error in the chapter on control charts. The Western Electric out-of-control rules “two of three in Zone A or beyond” and “four of five in Zone B or beyond” were presented in the accompanying graph as crossing the centerline. The rules actually only apply to one-half of the control chart.

There is quite a bit of advanced material in this book, including kernel density estimation, bootstrapping, maximum likelihood estimation, and likelihood ratio tests. Of course, the coverage is not thorough, but it will give students the idea that there is theory behind the applications.

Diane K. Michel
Sematech

Reliability Theory: With Applications to Preventive Maintenance, by Ilya Gertsbakh, Berlin: Springer-Verlag, 2000, ISBN 3-540-67275-3, xii + 219, \$74.

Preventive maintenance planning is one of the few applications of reliability theory in which the close connection, via economics, of the mathematical models to the needs of decision makers is immediately apparent. Ilya Gertsbakh, a contributor to the recent Special Issue of the *IIE Transactions on Quality and Reliability Engineering* devoted to Reliability Economics (Gertsbakh 1998) and coauthor with K. Kordonskiy of a widely respected reliability-theory textbook (Gertsbakh and Kordonskiy 1969), has written a full-length treatise on preventive maintenance (PM) for engineering systems. I think you will agree with me that this book is a welcome addition to the resources our profession possesses to learn from, teach from, and refer to.

The first three chapters—“System Reliability as a Function of Component Reliability,” “Parametric Lifetime Distributions,” and “Statistical Inference from Incomplete Data”—form a review of the mathematical and statistical theory of reliability. The coverage is broad but not overly deep and the emphasis is on topics that will be used later in the development (I might have called the book *Preventive Maintenance Models in Reliability Engineering* because its real weight seems greater in the later chapters than in these first three). The topics here go by pretty quickly (Gertsbakh is eager to get to the applications!), so how could this affect a course? The instructor will have to fill in details for beginning students. For example, the discussion of failure rate is very straightforward, but the instructor needs to augment it because there is not enough coverage of the differences in usage between nonrepairable systems and repairable systems, where the same words mean two entirely different concepts. So here’s a clue that this is not a book for beginners, at least without the guidance of an experienced coach. But there are still wonderful nuggets in these first three chapters—the treatment of the dodecahedron network at the end of Chapter 2 that makes important points, shows an important technique, and discusses approximations all in the space of just over a page; the derivation in Section 2.3.2 that shows why the Weibull distribution is so popular among reliability engineers; and the development of system reliability characterization by the path of static probability to fixed-time probability to lifetime calculus to maximum-minimum operations are just three examples. Those who are ready to tackle the remaining four chapters of the book on their own probably will not need the review, but in the hands of an instructor who has considerable modeling experience in addition to a thorough technical command of reliability theory, the review can be a powerful teaching aid.

The meat of the book is in the last four chapters covering preventive maintenance models based on the lifetime distribution, preventive maintenance based on parameter control, best time scale for age replacement, and preventive maintenance with learning. The literature on preventive maintenance models is large but scattered. Gertsbakh has done a good job of collecting important applications (many of which are original contributions of his) and giving them a unified treatment. There is a nice mixture of statistics in with the probability models. That is, examples are given with real data, not just magically

appearing values of the parameters of some distribution, and completing the modeling task includes estimation of the relevant parameters before computing the optimal PM plan. This gives a more complete picture of an analyst's real activities when it comes time to implement a PM solution.

Chapter 4 starts with a review of renewal theory and reward processes. The high spots are hit, but again the review is brief and is most suited for those who already know the subject. The nice interplay of the probability models and the statistics is maintained, for example, in Section 4.1.3, which discusses stationary availability. One odd point is that the renewal functions (4.1.9) for the gamma distribution and (4.1.10) for the (truncated?) normal distribution are hardly closed forms! Here the author misses the (few) closed forms besides that for the exponential distribution that are known—the shifted exponential and truncated exponential (Barlow and Proschan 1964) and a mixture of exponentials (Tortorella 1993). In fact, Barlow and Proschan pointed out that Parzen (1962) showed that (4.1.9) can be integrated and produces the finite sum displayed after (2.14) of Barlow and Proschan (1964). The basic block replacement model is then described. Here the basic strategy of PM modeling is clearly set out, if not explained outright (it would be good if an instructor were to fill this in by explicitly listing the steps involved). This serves well for the remainder of the chapter, which explores many variations on the basic model, including minimal repair as well as replacement (renewal), age replacement, and multielement systems. Useful qualitative results are also provided. One particularly appealing contribution is Section 4.3.5 on what to do when the input information you have to base your model on is uncertain or of poor quality. This is a situation engineers face every day, but there is very little elsewhere in the pedagogical literature to help. As is characteristic of this book overall, the treatment is very thorough but it is also very economical, even dense in places. The instructor will have to augment the book with alternate views and more extensive explanations for beginners.

Chapter 5 covers PM for systems in which an indicator of intermediate condition between perfectly functioning and completely failed is available. This is called preventive maintenance with parameter control. The educational example Gertsbakh gives is the deterioration of concrete construction, a system that fails by means of accumulating damage. The simplest treatable example, though, is a system comprising many identical units, where the “accumulated damage” is the number of failed units. The development brings in dynamic programming and gives a nice illustration of Bellman's principle. After introducing Markov and semi-Markov processes, opportunistic replacement in a two-component system and systems with multidimensional state descriptions are covered. At the end of the chapter, Gertsbakh gives a lovely synthesis that shows how the maintenance action framework is neatly summarized as a control problem for a certain stochastic process related to the deterioration (or state changes) of the object. Those versed in sequential analysis may note that Figure 5.12 has a familiar feel.

Chapter 6 focuses on age replacement and, in particular, on how age replacement should be carried out when there is more than one sensible time scale to describe the age of the object. For example, an automobile's age can be described by calendar age or mileage. The strategy Gertsbakh proposes is simple: Form a new time scale that is a convex combination of the original time scales and optimally choose the coefficients by minimizing the coefficient of variation of the time-to-failure distribution in the combined time scale. An illustration is given from a fatigue life application, and algorithms for determining the coefficients are given under nonparametric conditions, parametric conditions, and censoring. This is an interesting contrast to George's approach of looking at the two-dimensional distribution of calendar age and mileage in a population of automobiles that have accumulated a certain percentage of failures (George 1998).

The final chapter deals with PM with learning. This again is an attempt to make PM models correspond more faithfully to what happens in real engineering. The arrival of new data as a system continues to operate should cause us to reexamine the PM decisions we made based on the data we had before the new arrivals. Gertsbakh creates a sequential framework in which to formally capture these notions and uses a Bayesian approach to updating the PM strategy with the arrival of new data.

There are five appendixes. Appendix D gives normal and Weibull probability paper that you can copy and use. It also gives *Mathematica* code for normal and Weibull probability plots. The inclusion of probability paper, while quaint, may also be taken as a homage to the reliability engineers of the past who had to undergo the tedium of probability plotting by hand. In this appendix, a reference to one of the many commercial life-data-analysis programs would have been useful because these software packages permit

probability plotting for many more distributions. Appendix E provides a table of the renewal function for a Weibull distribution with shape parameters 1.5, 2, 2.5, 3, 3.5, and 4. As explained on page 71 of the text, the values in the table were obtained by summing the first five terms of the infinite series for the renewal function, with the convolutions being performed recursively via some (unstated) quadrature rule. A maximum error of 0.001 is claimed. Readers wishing to learn more about approximations for renewal functions may consult Tortorella (2001) and the many references therein.

The exercises are practical, challenging, and do illustrate well the points in the text. Solutions are provided for all the exercises. Some instructors may not like this (a rich source of exam problems is closed off!), but students who do the exercises conscientiously will rapidly gain facility by learning from the well-organized solutions—for example, by comparing their efforts with the published solution. Many of the solutions also contain *Mathematica* code for the required operations. The aphorisms opening each chapter are enjoyable and mostly apt. There are a few typos (I estimate about a dozen based on my count of four in the parts of the book I read thoroughly), but I did not spot anything that a few minutes of thought would not make right (maybe creating deliberate “typos” is an interesting way to add “hidden” exercises—on second thought, maybe it is just too dangerous).

Overall, this is a wonderful book. The theory is sound, the examples are practical, real concerns are treated, and the exposition, although economical, is well organized and logically laid out. The emphasis is on modeling; many proofs are omitted. Anyone surviving a solid course based on this book would truly be an expert on PM and, more importantly, on how to create and use new PM models. Beginners will need the guidance of a good instructor. As a resource for self-study and as a reference for the professional, this book is an example for others to follow.

M. Tortorella
Bell Laboratories

REFERENCES

- Barlow, R. E., and Proschan, F. (1964), *Mathematical Theory of Reliability*, New York: Wiley.
- George, L. L. (1998), “Convert From Calendar Time to Operating Time Reliability,” *Reliability Review*, 18, 15 ff.
- Gertsbakh, I. (1998), “Parallel Time-Scales and Two-Dimensional Manufacturer and Individual Customer Warranties,” *IEE Transactions on Quality & Reliability Engineering*, 30, 1181–1189.
- Gertsbakh, I. B., and Kordonskiy K. (1969), *Models of Failure*, Berlin: Springer-Verlag.
- Parzen, E. (1962), *Stochastic Processes*, San Francisco: Holden-Day.
- Tortorella, M. (1993), “Comparing Renewal Processes, With Application to Reliability Modeling of Highly Reliable Systems,” *Technometrics*, 31, 357–364.
- (2001), “Numerical Solution of Renewal-Type Integral Equations,” unpublished manuscript.

Statistical Modelling With Quantile Functions, by Warren Gilchrist, Boca Raton, FL: Chapman & Hall/CRC, 2000, ISBN 1-58488-174-7, xx + 320 pp., \$69.95.

The use of quantile functions in the description and modeling of probability distributions is the focus of this book. The key idea is that, by using quantile functions, a methodology for estimating probability models from observed data can be developed that is analogous to regression methods used to estimate the expected value of a random variable. This methodology springs from the fact that quantile functions can be combined through the operations of addition and multiplication to produce other more complex and useful quantile functions. Models for the quantile function are equivalent to a model for the probability distribution of stochastic data since the quantile function is the inverse of the cumulative distribution function.

The methodology developed in this book provides a fundamentally different approach to modeling the stochastic behavior of data in comparison to the standard statistical approach. In the standard approach, models are selected from a library of potentially useful models, with attention generally focused on a few standard models. As the author points out, when there are thousands of observations, standard probability models that are controlled by one or two parameter values may not fit the data very well, especially in the tail area

of the distribution. There are no such limitations on models built using the methodology presented here.

The goal of this book is to serve as an introductory textbook for statistical practitioners who wish to use quantile methods for data analysis. To provide a complete introduction, the book also provides an overview of the history and theory behind quantile methods, including a large bibliography of papers and books in which these methods have been developed. The book also develops a large toolkit of graphical and computational methods for using the quantile function and the related quantile density function and density quantile function for statistical data analysis.

The book is organized into 14 chapters and 3 appendixes. Chapter 1 serves as an introduction and grand overview of the topics covered throughout the remainder of the book. In this chapter, the author illustrates the usefulness of the techniques by a series of short examples employing quantile methods. The remaining material falls into four main categories—description of statistical data and probability distributions by quantile methods, foundational theory supporting the use of quantile methods, development of a framework for the practical use of quantile methods for the analysis of statistical data, and applications of the techniques in some standard statistical problems.

Chapter 2 is devoted to the overview of the use of the quantile function, the quantile density function, and the density quantile function for describing statistical data, especially through the use of graphical techniques. Summary statistics such as the quartiles and related measures of spread and skewness are discussed also. The extension of these statistics to descriptive population parameters is addressed in Chapter 3. Chapter 3 also begins the introduction of the foundational theory of quantile methods with the development of nine rules for the combination and transformation of quantile functions. Chapter 4 continues a description of the foundational theories with a discussion of order statistics and their distributions. Chapter 5 provides an overview of some of the probability distributions that can be used as building blocks in quantile methods. These include the uniform, exponential, Pareto, Weibull, power, extreme type I, and normal distributions. Chapter 6 presents and discusses methods that can be used to combine and transform these building blocks into more complex distributions. Chapter 7 explores some additional useful probability distributions such as the logistic, skew logistic, lambda, extreme type II and III, and Burr family of distributions.

Chapters 8, 9, and 10 are concerned with building a framework for the practical application of quantile methods in statistical data analysis. Chapter 8 addresses the identification of adequate models for experimental data through the use of graphical methods such as the quantile-quantile plot, skewness plot, and tail-shape plot. The chapter also develops a suggested sequence of steps that should be followed by researchers using quantile methods to identify probability models. The estimation of parameters for models developed through quantile methods is discussed in Chapter 9. The methods presented include moment-matching methods and maximum likelihood methods with discussion of methods for measuring the lack of fit of quantile-based models. Chapter 10 focuses on validating models developed using the techniques discussed in the previous two chapters with both numerical and graphical tests.

Chapters 11, 12, and 13 are concerned with a summary of a number of applications of quantile methods. Chapter 11 discusses some of the straightforward applications of the developed techniques to reliability, hydrology, and statistical process control. Chapter 12 discusses the use of quantiles in regression, discussing semiparametric and fully parametric models. Chapter 13 introduces bivariate quantile methods with attention specifically on circular and elliptical distributions. The final chapter, Chapter 14, provides some closing comments on the process of developing models for the probability distribution of statistical data. The three appendixes provide further detail on some mathematical functions useful to quantile methods, on the method of maximum likelihood, and on bivariate transformations of probability distributions.

Each chapter contains some exercises meant to guide the reader in understanding the techniques discussed. However, there are no answers provided to the exercises. The book also includes numerous short examples of the techniques being presented. I frequently found the examples included in the book to be too brief, leaving the reader to puzzle out too many of the details. In particular the chapters on model identification and estimation could have benefited from more examples and more detailed examples. I also think the book would have particularly benefited from a case study, which could have been used to illustrate the techniques in detail and highlight their usefulness. Overall though, I think this book provides a valuable starting point for anyone

interested in quantile methods and it makes a strong case for the adoption of these methods as part of the applied statistician's toolbox.

Jeffrey Glusup
Sun Microsystems Inc.

HALT, HASS and HASA Explained: Accelerated Reliability Techniques, by Harry McLean, Milwaukee: ASQ Quality Press, 2000, ISBN 0-87389-489-8, xvi + 152 pp., \$48.

HALT (Highly Accelerated Life Test), HASS (Highly Accelerated Stress Screen), and HASA (Highly Accelerated Stress Audit) are engineering approaches to reliability testing with the ultimate goal of improving, rather than statistically demonstrating or estimating, the reliability and robustness of products. Workshops and seminars on these topics have become a niche area for a small number of industrial consultants and trainers over the past few years. Despite their proponents' claims of success and popularity of such testing, there have been relatively few publications and case studies describing their use in applications. In this sense, this book fills an important void in the reliability engineering literature and thus it is a welcome contribution. Incidentally, a book with a similar title (Hobbs 2000) has also appeared recently, but it will not be reviewed here.

What exactly are HALT, HASS, and HASA? What are their underlying objectives and methodologies? How do they differ from each other and from other types of reliability testing such as ALT (Accelerated Life Test), ESS (Environmental Stress Screen), STRIFE (Stress-Life Test), RDE (Robust Design Experiment), RGT (Reliability Growth Test), RDT (Reliability Demonstration Test), and burn-in? If you are someone involved with reliability improvement projects and/or research but yet are unsure in your answers to these questions, you will benefit from this book.

The basic message of the book is sound and appropriate. HALT is used during the design phase of a product on pre-production units to generate forced failures under extreme stress conditions to be able to identify and remove their root causes, where it makes sense to do so, by design improvements. HASS is described as a reliability screen after production release in which 100% of the manufactured units are tested. HASA is similar to HASS except that it is applied on an audit basis. All three types of testing expose products to extreme environmental conditions, typically involving thermal and vibration cycling. The choice of stress variables and their levels needs to be tailored to the specifics of the situation. Proper implementation of these testing methodologies seems to be more of an engineering art than an exact science. The main emphasis of the book is on design of engineering tests so as to expose important weaknesses of the product. Guidelines for robust design of products even before any testing is initiated or approaches to identify root causes of failures observed during testing are outside the scope of the book.

There are seven chapters in the book:

Chapter 1: The Importance of High Reliability Products at Market Introduction—How and Why Do a HALT?

Chapter 2: Highly Accelerated Stress Screen—HASS

Chapter 3: Beyond the Paradigm of Environmental Stress Screening—Using HASA

Chapter 4: Refinements on Highly Accelerated Stress Audit (HASA)

Chapter 5: The Equipment Required to Perform Efficient Accelerated Reliability Testing

Chapter 6: How to Sell New Concepts to Management

Chapter 7: Some Commonly Asked Questions and Observations

The bulk of the 35 or so references cited in the book are from publications such as the *Proceedings of the Institute of Environmental Sciences* and the *Proceedings of the Annual Reliability and Maintainability Symposium*.

The target audience for the book is engineers who are responsible for design and implementation of reliability test programs. Most of the examples and illustrations involve electronic and electromechanical components reflecting the author's experience at companies such as HP and AT&T. The presentation on test design is driven primarily by engineering considerations, a consequence of the basic mission and philosophy of such testing. The author provides some highly specific and useful recommendations on technical topics ranging from the choice of test stress profiles (e.g., how much vibration and for how long?) to the type of material to use in building the test fixtures.

The contributions made by reliability statisticians to design and analysis of HALT, HASS, and HASA have been sparse and not sufficiently visible. This is reflected by the modest statistical content of the book. Chapters 3 and 4 are devoted to traditional attribute acceptance test plans in support of HASA. Citations to references such as Montgomery (1996) or Schilling (1982) would have been useful here for readers who wished to learn more about sampling plans. In Chapter 1, the author describes a procedure to establish ± 4 sigma limits (to assure no more than .0032% defects to escape HALT) to determine which HALT failures must be fixed via design improvements. Even though I do not consider myself a statistical purist, I had much difficulty following the statistical reasoning offered here due to imprecise definition of terms (e.g., sigma is used to mean "z" without defining it) and unclear assumptions about the random variables involved and their distributions.

The book is generally well written and easy to read. It will be a useful resource for reliability test engineers who are looking for practical engineering guidance on HALT and related test methods. The introductory sections of the book can also serve the needs of reliability-modeling and data-analysis specialists who wish to learn about the basic objectives and methodologies of the subject test methods.

Necip Doganaksoy
GE Corporate Research and Development

REFERENCES

- Hobbs, G. K. (2000), *Accelerated Reliability Engineering: HALT and HASS*, New York: Wiley.
Montgomery, D. C. (1996), *Introduction to Statistical Quality Control* (3rd ed.), New York: Wiley.
Schilling, E. G. (1982), *Acceptance Sampling in Quality Control*, New York: Marcel Dekker.

Conditional Specification of Statistical Models, by Barry C. Arnold, Enrique Castillo, and Jose Maria Sarabia, New York: Springer-Verlag, 1999, ISBN 0-387-98761-4, xvi + 411 pp., \$74.95.

There are several ways the joint distribution of two random variables X and Y can be specified. One way is direct: Spell the joint distribution in the form of a bivariate table if X and Y are discrete or in the form of a joint probability density function if X and Y are (absolutely) continuous. Of course, once the joint distribution is spelled directly, one can derive marginal and conditional distributions of every hue and color. Another way is somewhat indirect: Spell the marginal distribution of X and the conditional distribution of Y given $X = x$ for each x in the support of X . In a similar vein, one can spell out the distribution of Y and the conditional distribution of X given $Y = y$ for each y in the support of Y .

Suppose we specify two sets of conditional distributions: One set consists of conditional distributions of X given $Y = y$ for each y in the support of Y and the second set consists of conditional distributions of Y given $X = x$ for each x in the support of X . Is there a compatible joint distribution of X and Y that gives the two sets of conditional distributions with which we originally started? The answer, in general, is no. There are examples. Two different joint distributions might produce the same two sets of conditional distributions.

Consider the following two joint distributions for X and Y :

$$\begin{array}{c} \underline{Y} \\ \underline{X} \begin{pmatrix} 1/2 & 0 \\ 1/2 & 1/2 \end{pmatrix} \end{array}, \quad \begin{array}{c} \underline{Y} \\ \underline{X} \begin{pmatrix} 1/3 & 2/3 \\ 1/3 & 2/3 \end{pmatrix} \end{array}.$$

Under the first joint distribution of X and Y , the conditional distribution of X given $Y = 1$ is degenerate at $X = 1$, the conditional distribution of X given $Y = 2$ is degenerate at $X = 2$, the conditional distribution of Y given $X = 1$ is degenerate at $Y = 1$, and the conditional distribution of Y given $X = 2$ is degenerate at $Y = 2$. Under the second joint distribution, we have the same sets of conditional distributions.

The predominant theme of the book under review is an exploration of the interplay between conditional distributions and joint distributions. Two fundamental questions arise. (1) When are the conditional distributions of X for all possible values of Y and the conditional distributions of Y for all possible values of X compatible? In other words, will there exist a joint distribution for X and Y that gives the same conditional distributions with

which we started? (2) If compatibility is confirmed, is the joint distribution unique?

In Chapter 2, the authors meticulously examine the issues involved when both X and Y are discrete. Chapter 3 presents the case of normal conditionals in great detail. Chapter 4 presents general results regarding joint distributions whose conditionals are posited to be members of exponential families. Chapter 5 discusses a host of other conditionally specified families. The cases of Pareto, Cauchy, beta, Weibull, and logistic conditionals, in particular, are also included.

Once we venture beyond two dimensions, formulation of the conditional specification paradigm becomes more intricate. Extensions from two dimensions to higher dimensions can take off in several directions. These extensions have been explored thoroughly in Chapter 8.

It is generally hard to specify a multivariate model directly for a given dataset. One advantage of the conditional specification paradigm is that it enables the data analyzer to reduce the problem of specification to a specification of a series of one-dimensional conditional distributions. The authors succeed admirably in promoting this approach. For the benefit of the data analyst, the authors do focus on estimation problems in conditionally specified models. They have included Bayesian analysis to bring the Bayesians into their fold.

The mathematics involved is not hard. In many arguments, functional equations do crop up. Fittingly, the book is dedicated to Janos Aczel, one of the pioneers responsible for placing functional equations in the firmament of mathematics. Functional equations and characterizations go together hand in glove. I am sure that this book will delight those whose specialty is in characterizations.

The book is well written. The style is pleasant. The subject matter of the book should appeal to purists as well as data analysts. A substantial part of the book is based on the research carried out by authors themselves. A number of open problems are interspersed throughout the book. I enjoyed reading it. One could organize a graduate course based on this material that should stimulate inquisitive graduate students.

M. Bhaskara Rao
North Dakota State University

Probability Applications in Mechanical Design, by Franklin E. Fisher and Joy R. Fisher, New York: Marcel Dekker, 2000, ISBN 0-8247-0260-3, xiii + 276 pp., \$135.

This book is an introductory textbook in probability and statistics for a rather mechanical-engineering-oriented audience. The mathematical level is voluntarily kept minimum. References and exercises are given at the end of each chapter. It is an extension of Haugen (1980) with applications. It is in a format of class notes.

The book contains four chapters. Chapter 1 covers basic statistical method including histogram, confidence interval, and hypothesis testing. Two probability distributions, Gaussian and Weibull, are also introduced. Chapter 2 introduces basic concepts of probability and some basic probability laws such as additive law and Bayes theorem, first-order variance approximation for functions of independent random variables, coefficient of variation, and confidence limits. Chapter 3 covers concepts of optimization and geometric programming. Chapter 4 contains reliability analysis for failure-rate data.

The book contains a large amount of data from mechanical designs for exercises. I feel that more examples could have been provided for topics in each chapter. It would be particularly useful to credit the computer software used since there are numerous computer calculations. It can be seen through the many illegible sentences and mathematical mistakes that it was not carefully reviewed before publication.

Lianfen Qian
Florida Atlantic University

REFERENCE

- Haugen, E. B. (1980), *Probabilistic Mechanical Design*, New York: Wiley.

Wavelet Methods for Time Series Analysis, by Donald B. Percival and Andrew T. Walden, Cambridge, U.K.: Cambridge University Press, 2000, ISBN 0-521-64068-7, xxv + 594 pp., \$69.95.

The purpose of this book is to introduce wavelet methodology in time series applications. Emphasis is placed on the discrete wavelet transform and other discrete wavelets. There are many analogies to the spectral analysis of time series that the authors make in the book. However, they do not assume that the reader has any knowledge of Fourier analysis.

They cover the relevant aspects of the frequency-domain theory of time series in Chapters 2 and 7. Just as the frequency-domain methods decompose the variance of a time series into frequency components, so do wavelets decompose the variance according to scaling. The authors take great pains in demonstrating this with several specific wavelets including the discrete wavelet transform (DWT) and the maximal overlap discrete wavelet transform (MODWT).

Fast transform methods exist for wavelets as well as for Fourier transforms, and they are apparently simpler and faster for wavelets. This is a computational selling point for wavelets.

The contents of the book are as follows:

1. Introduction to Wavelets
2. Review of Fourier Theory and Filters
3. Orthonormal Transforms of Time Series
4. Discrete Wavelet Transform
5. The Maximal Overlap Discrete Wavelet Transform
6. The Discrete Wavelet Packet Transform
7. Random Variables and Stochastic Processes
8. The Wavelet Variance
9. Analysis and Synthesis of Long Memory Processes
10. Wavelet-Based Signal Estimation
11. Wavelet Analysis of Finite Energy Signals

The authors attempt to provide a self-contained book, and they start out at a very elementary level. They emphasize applications and the practical aspects of wavelets. The discussion is detailed and thorough, and they provide an excellent list of references. This is not a complete account of wavelet theory. The authors believe that to cover the plethora of available wavelets would be overwhelming for their audience. They concentrate on the wavelet filters described in Daubechies' work and restrict themselves to time series applications. It is aimed at advanced undergraduate and graduate students and is also a useful reference for researchers in statistics and various physical sciences.

A nice feature of the book is the demonstration of various techniques using some familiar real time series such as the atomic-clock data. Several practical examples are given in Chapter 5. I found the material fairly easy to follow in the beginning, but it did get more difficult to follow as they got to more advanced concepts. To me it is easy to see how a time series can be decomposed into periodic components and from that how the spectral distribution can represent a decomposition of the variation in the time series among its frequency components. The analogous result with respect to scales for wavelet transform is more difficult to grasp.

In reading the book it is important to review the conventions and notation at the beginning. Some of the conventions are very unique. Figures and tables are referenced to their page number rather than in a sequential order. This can be annoying at times. When more than one figure appears on a single page, they are listed as, for example, 98a and 98b for the figures on page 98.

Important aspects of the book that make it worth having for time series analysts are the material on the wavelet variance, the analysis of long-memory processes, and signal estimation through "thresholding" and "denoising." Moreover, it is reasonably priced for a book of nearly 600 pages.

Michael R. Chernick
Biosense Webster

Nonlinear Time Series Analysis, by H. Kantz and T. Schreiber, Cambridge, U.K.: Cambridge University Press, 1997, ISBN 0-521-55144-7, xvi + 304 pp., \$80.

Despite the similarity of the title of this book to the earlier one by Tong (1990), the contents of these two books are almost orthogonal. Tong

emphasized nonlinear extensions of autoregressive (AR) models (product-, fractional-, or amplitude-dependent AR models, doubly stochastic models, etc.), for which he is able to present a significant number of theoretical results. Kantz and Schreiber concentrate on computational algorithms that have been used, mostly by nonstatisticians, to analyze time series that are thought to represent low-dimensional deterministic chaos. There are no formal theorems, and definitions are also casual. For example, the definition of stationarity in the beginning of Chapter 2 provides an intuitive, nonmathematical interpretation of the concept and emphasizes its practical consequences. Even the later, nominally more advanced chapters do not contain theorems and proofs. Rather they sketch results and supporting arguments and point the reader to original and review papers for details.

The book is organized into two main parts. The first half introduces the main methods treated by the authors—phase space methods (embedding methods), nonlinear prediction, estimation of Lyapunov exponents and attractor dimensions, and hypothesis testing, particularly using surrogate datasets. Most of these topics are revisited at a more advanced level in the second half of the book, which also introduces a few additional topics such as chaos control. Always the perspective is that the concepts presented are useful only if the relevant quantities can be estimated based on finite (although frequently very long) time series samples, so computational algorithms are prominently featured.

The main value in this book lies in its intuitive presentation of these concepts and their abundant illustration by examples. The examples are based on about a dozen different datasets and develop intuition and practical ability. For statisticians who work with physical scientists—a category that presumably includes many readers of *Technometrics*—this approach provides an introduction to dynamical systems that is probably closer to the way in which those scientists understand the subject than is the statistical time series perspective offered by Tong [1990]. Its readers may be better prepared and motivated to continue with the theoretical literature found in the disciplinary journals, to which Kantz and Schreiber have also provided extensively annotated guidance.

At the same time, their data-analytic approach makes their book more accessible to statisticians than most of the deterministic chaos literature. Much of that literature does not concern itself with the inference problems that arise if one has data but not much theory about the dynamics of the process that produced those data. (This tends to be the case for applications in ecology, economics, or epidemiology, to name three disciplines in which these methods have become quite popular.) By contrast, Kantz and Schreiber continually remind the reader of the limitations of observational data, although they provide mostly ad hoc methods for dealing with them. Their book makes it clear that there is still much to be done to establish a statistical theory of inference for deterministic, chaotic dynamical systems, despite two decades of theoretical and computational development.

Katherine Campbell
Los Alamos National Laboratory

REFERENCE

Tong, H. (1990), *Non-linear Time Series—A Dynamical System Approach*, Oxford, U.K.: Clarendon Press.

Multivariate Statistical Methods: A First Course, by George A. Marcoulides and Scott L. Hersherberger, Mahwah, NJ: Lawrence Erlbaum, 1997, ISBN 0-8058-2572-X, xi + 322 pp., \$39.95 (softcover).

The authors' intended audience for this book is advanced undergraduate and graduate business students. The authors suggest, however, and I agree, that students—and, I would add, practitioners—from other disciplines will also find the book useful. Some more theoretically sophisticated readers will undoubtedly find the lack of mathematical rigor in this book somewhat alarming, but I suspect that many readers will find the pragmatic approach and helpful advice extremely refreshing. Each multivariate topic is introduced using univariate analogs wherever possible and concludes with a worked-out multivariate example that includes output from SAS (1979, 1989a,b), generally with the accompanying SAS syntax that produced the output. In between is practical advice on such things as choosing the number of components in a principal-component or factor-analysis problem or choosing the criteria for

comparing groups in a multivariate analysis of variance. It is interesting to note that, in both of these particular cases, the authors suggest, after introducing the various criteria for each problem, that if all the various criteria agree, it simply does not matter which you choose. They go on to conclude that, if the criteria do not all agree, one should look for those criteria that converge on similar conclusions. Such practical advice, although I'm sure is routinely followed by experienced practitioners, is rarely found in print. Many authors would rather spend time elaborating on the details of each of the criteria. In this context the introduction of the criteria, along with associated references, seems to suffice quite well.

The first chapter is an introduction to the book's approach, including a justification of their choice of SAS as the statistical package. It would have been useful to supply a general motivational example as an introduction to multivariate analysis in this chapter. The chapter basically parroted the information presented in the Preface. The second chapter is intended as an introduction to basic matrix algebra, but, as these chapters usually are, it is very simplistic and not a substitute for taking a decent course in linear algebra. However, in this book, unlike many others, there seems to be little lost by not having had a course in linear algebra because the reader here is spared most of the mathematical details anyway.

Each of the remaining six chapters covers a specific multivariate topic. At the end of each chapter there are problems on which the student can try out the techniques. Each chapter can be read independently of the others. The chapters are (3) "Multivariate Normal Distribution and Tests of Significance," (4) "Factorial Analysis of Variance," (5) "Discriminant Analysis," (6) "Canonical Correlation," (7) "Principal Components and Factor Analysis," and (8) "Confirmatory Factor Analysis and Structural Equation Modeling." The chapters average about 40 pages, including the example. SAS output is explained in some detail with each example. It would have been useful to annotate the actual output; the reader is left bouncing between the tables and the descriptions in the text.

The book may seem too heavily dependent on the use of SAS as the package of choice by non-SAS users. In choosing SAS the authors felt that "at times the output from these packages is so similar that only seasoned users are able to distinguish between them." If the reader has even limited experience with the SAS package, they will be able to follow the syntax used in the examples. However, if the reader is totally unfamiliar with SAS, he/she might better gain some preliminary experience before working on the examples because many of the more operational details of SAS are not covered. Readers using a package other than SAS may even find whole topics unavailable in their package. For instance, canonical correlation analysis and structural equation modeling are not available in MINITAB (Minitab 1997).

This book is a good introduction to multivariate statistical methods for someone who wants to understand the basic underpinnings of the aforementioned techniques and try them out on some data. The authors do not claim, or aim, to be a reference for the theory behind the applications. They do, however, provide a great deal of guidance for those people who, with the advent of readily available and very sophisticated statistical software, will be applying these methods to their own data.

Steven M. LaLonde
Rochester Institute of Technology

REFERENCES

- Minitab, Inc. (1997), *MINITAB User's Guide 2: Data Analysis and Quality Tools*, State College, PA: Author.
SAS Institute, Inc. (1979), *SAS User's Guide, 1979 Edition*, Raleigh, NC: Author.
—— (1989a), *SAS/STAT User's Guide, Version 6* (4th ed, Vol. 1), Raleigh, NC: Author.
—— (1989b), *SAS/STAT User's Guide, Version 6* (4th ed, Vol. 2), Raleigh, NC: Author.

Practical Geostatistics 2000, by Isobel Clark and William V. Harper, Columbus, OH: Ecosse North America, 2000, ISBN 0-9703317-0-3, 342 pp., \$60.

As noted by the authors, this book is an update of an earlier version in 1979; the principal changes pertain to the addition of material providing a more classical statistics background.

There are a total of 12 chapters, a bibliography, a collection of statistics tables, hard copies of the various datasets used in the discussions, and an index. Following the Preface, there is a two-page list of notations. See later comments for errors in this listing.

Chapter 1 describes the intended audience, provides a brief overview of the datasets that will be used, and refers to the software used in the book. The software is a simpler version of a commercial package produced by Clark's consulting company; a demo version of the software can be downloaded from http://uk.geocities.com/drisobelclark/PG2000_demo.html. The reader might be better advised to get a free copy of the GEOEAS software, which can be downloaded from the www.ai-geostats.org Web site.

Chapter 2 reviews basic descriptive statistics. Four datasets are used for illustration. Chapter 3 reviews some of the basic properties of the Normal distribution. Chapter 4 reviews the lognormal and three-parameter lognormal distributions. Chapter 5 is a very short review of the binomial, negative binomial, geometric, Poisson, and compound Poisson distributions. Chapter 6 reviews the basics of hypothesis testing—mean, standard deviation, difference of two means. Chapter 7 introduces regression and trend surfaces.

Chapter 8 is the first instance of a discussion of spatial data. Chapter 9 describes the estimation and modeling of variograms. Chapter 10 introduces ordinary kriging. Chapter 11 extends the point kriging in the previous chapter to block kriging—that is, estimation of averages over areas or volumes. Chapter 13 is a very brief discussion of lognormal kriging, kriging in the presence of a nonstationarity, indicator kriging, and rank transform kriging.

The strongest point in the book is the consistent use of several datasets throughout. The weak point is the presence of many errors.

The authors claim that the book is intended for beginners and in particular those without a strong mathematical background. In fact there is a lot of mathematics in the book, although it is mostly trivial such as algebraic simplification of the formula for a sample variance. The reader need not be burdened with a lot of the mathematical details—for example, derivations—but at the same time authors should avoid saying things that are incorrect. If a result is to be motivated by heuristics, then it should be clearly labeled as such. There are a great many errors in the chapters on statistics and many in those on geostatistics as well. The discussion of confidence intervals is mostly incorrect. It is claimed at one point that a variogram must be "conditionally positive definite"; in fact it is the negative of the variogram that must be conditionally positive definite; that is, the variogram must be conditionally negative definite. The discussion of the nugget effect has a number of mistakes; the authors do not properly distinguish between the variogram being nonzero at the origin and having a jump discontinuity. They do not distinguish between kriging with a variogram having a nonzero nugget and the variation of kriging that results in smoothing—for example, the thin plate spline versus the smoothing spline.

In the list of notations, e is identified as the natural anti-log function. In the preface, $2x$ is incorrectly identified as the differential of x^2 rather than as the derivative. χ is identified as the "distribution of sample variances from Normal samples" (not even χ^2 but χ). μ is identified as the "arithmetic mean of values in a Normal population." In several places it simply denotes the expected value of a random variable (not necessarily Normal). In Chapter 2 it is claimed that the sample standard deviation (of a set of heights) can be interpreted as "the typical difference between the actual height of any class member and the average height of the class." The kurtosis is defined as the fourth central moment, *which can then be standardized* by the square of the variance (italics added for emphasis).

The book is fraught with strange phrases. "For one thing it can not go negative—there is no such thing as a negative probability" (speaking of probability density functions). "Many distribution functions have strange looking constants in the formula, simply to ensure that the area under the curve is 1." "In particular, the Normal distribution is the only one which is always the same shape." "Statistical theory can show us what this new population will look like." "We find a 90% confidence interval is *contained by the Student's t value of*, so that" (italics added for emphasis). The t statistic is incorrectly defined on page 44 (the $_n$ is missing).

In summary, a knowledgeable reader will find some interesting examples and will be able to steer around the misstatements. It is not recommended for beginners since they will likely not recognize the mistakes.

Donald E. Myers
University of Arizona

Robust Bayesian Analysis, edited by David Rios Insua and Fabrizio Ruggeri, New York: Springer-Verlag, 2000, ISBN 0-387-98866-1, xiii + 422 pp., \$59.95.

Bayesian analysis is a most straightforward technique with the algorithm defined by the steps: (1) Write down your prior probabilities about the phenomenon being studied. (2) Before conducting experiments, write down the probabilities of the prospective outcomes, conditioned on your current prior probabilities. (3) Compute posterior probabilities. (4) To make a decision, write down your loss function. (5) Make a decision that minimizes the loss. It looks so sensible!

In actual applications, questions arise about the merit of the answers given to each request for input. What will happen if your prior was wrong? If I were a pharmaceutical agency, to what extent should I trust *your* prior? The same holds about the conditional probabilities. The story is no different about the loss function. There are no questions about posterior probabilities or the aspect of minimizing your loss.

On a personal basis, I make such probabilities, keep track of my posterior probabilities, and when a decision is due, I find out what my decision will be. I am most pleased with them. In the most recent time, I used it to assign grades to my students. You may send me an email to get a copy of it. My grades matched very well the grades that the department would have given using its own style. An additional advantage was that I had successive posterior distributions so that when a student would complain about his/her final grade being too heavily loaded by one of his/her missteps, I would show him/her how successive opinions converged to the final grade and that any misstep was just the final nail. The best way for any of us to promote Bayesian structure is to use it in our daily practice when we are taking a chance on ourselves and not on our employers!

This book catalogs the attempts to study the way the Bayesian decisions change as the priors differ by various margins. You can see that, if the priors are halfway reasonable, the data are going to overwhelm the prior and give a reasonable decision. Questions arise when the priors are such that a decision is made with only a few observations and people with differing priors come to entirely different conclusions. There is a vast amount of literature on this topic. You did not expect one final solution, did you? There is not one.

With this diverse literature generated by a large group and published in equally diverse journals, I was most thrilled that these editors are putting together the work of the more prolific of the authors in such a succinct way. J. O. Berger and the editors, in their opening article, write on page 7, "Thus, there is robustness against only change in probabilities, not with respect to joint changes in probabilities and losses."

Before you go negative on such robustness studies, recall the robustness studies about normality and other aspects of currently standard techniques in the Fisherian style. Many highly differing algorithms were considered. Success of the current technique in biostatistics and so forth is because in these areas one stays within certain bounds; things do not go berserk; when they do, they have their mechanisms for straightening them out, and the society is set up to handle the losses taken by the consumers.

When R. A. Fisher proposed his randomization, design of experiments with t tests, analysis of variance, and so forth, it took many successful projects to make them household words. I recall the experiments made by George Snedecor and followed up by T. A. Bancroft in the area of canning foods for military use in the European Theatre—as they used to call it—and the many other topics such as quality control of ball bearings and antibiotics that we take for granted.

A set of topics as given in the chapter headings illustrates the areas covered. It is most impressive and so is the total work published by the authors over the decades. Here is a quick list:

Global Bayesian Robustness for Some Classes of Prior Distributions.
Likelihood Robustness
Stability of Bayes Decisions and Applications
Robustness Issues in Bayesian Model Selection
Bayesian Robustness in Bayesian Nonparametrics
Linearization Techniques in Bayesian Robustness
VIII: Case Studies
Sensitivity Analysis in Ict Neo
Sensitivity of Replacement Priorities for Gas Pipeline Maintenance
Robust Analysis in Medical and Epidemiological Settings
A Robust Version of the Dynamic Linear Model With an Economic Application

After I got this book to review, I read it many times and I have new ideas on how to develop new methods and promote their use. I hope that you get equally excited and give voice to this methodology just as Snedecor et al. gave to Fisherian style, Wilcoxon and others to the use of nonparametrics—you list others who gave life to the work in decision theory and many other areas of statistics. The job of promoting use of statistics in the outside world of applications has just begun.

I am sure that this book will contribute much to this effort. I must submit that the authors' effort in putting together this voluminous material in this brief book of only 400 pages is admirable.

S. K. Katti
Professor Emeritus
University of Missouri

Analysis of Customer Satisfaction Data, by Derek R. Allen and Tanniru R. Rao, Milwaukee, WI: ASQ Quality Press, 2000, ISBN 0-87389-435-7, xiv + 243 pp., \$55.

The intended audience for this book is "advanced service quality managers and marketing researchers with more than a modest exposure to statistical data analysis." The book is described as a "comprehensive guide to multivariate statistical analysis in customer satisfaction, loyalty, and service quality research." I would have described the book as a guide to *interpretation* of multivariate statistical analysis in customer-satisfaction, loyalty, and service-quality research. The book has a good description of survey-instrument design and scale selection (Chap. 3) and also a good treatment of missing data in customer-satisfaction research (Chap. 4). The remaining five chapters in the book cover a wide range of multivariate data-analysis tools including multiple linear regression, causal analysis, principal components analysis (PCA) and factor analysis (FA), and structural equations. The book should be of use to managers who are interested in getting an overview of multivariate data-analysis tools applicable to customer-satisfaction research. However, Appendix A ("Matrix Algebra in Statistics") seems to serve no purpose—if the reader has some knowledge of matrix algebra, then this chapter gives no new information or insight about the material, and if the reader is not familiar with matrix algebra, then this chapter does not have sufficient detail to be of much use. Chapter 9 also has a brief discussion of matrix algebra that does not serve any purpose and could have been omitted. Some additional comments follow:

1. Page 73: Discussion of r does not point out that statistical significance of r only implies $r \neq 0$. No discussion of physical significance is included.

2. Page 80: The stepwise selection technique described in the book is, in fact, the forward stepwise method; the backward stepwise method is not discussed and the fact that the two stepwise approaches could lead to two completely different equations is not mentioned.

3. Page 121: The term "heteroscedasticity" is described as a condition wherein the residuals vary over the range of the predictor variables. Residuals, being random variables, always will vary over the range of the predictors; the term "heteroscedasticity" refers to the condition wherein the *variance* of the residual varies over the range of the predictors.

4. Page 207: There is a typographical error—the element in row 1, column 2, of the matrix product $A'A$ should be 35 and not 34.

5. Chapter 8 has a fair description of PCA/FA and also a good description of exploratory versus confirmatory analysis. This chapter does mention that PCA/FA can be performed using either the covariance matrix or the correlation matrix but fails to state that the correlation matrix is typically used when the variables ranges differ by orders of magnitude. On page 155, a statement is made that "in PCA, the diagonal of the correlation matrix is assumed to be composed of ones"; it should be noted that the diagonal of the correlation matrix is always composed of ones, and this is not an assumption in PCA. Moreover, interpretation of bi-plots (loadings and scores) is missing from this chapter (Johnson 1998).

6. The term "Discriminant Analysis" is not mentioned in the book but is included in the glossary. The glossary does not include a clear description of this term.

7. The term "eigenvalue" is included in the glossary (p. 217), but the book fails to define this term.

This is a well-written book, but it does have a few shortcomings—paucity of real examples from customer-satisfaction research, no mention of profile analysis—a potentially useful tool in customer-satisfaction research—and no mention of correspondence analysis, to name a few. The book is still a good resource for people working with customer-satisfaction data who have some background in mathematics and applied statistics.

Ashok K. Singh
University of Nevada, Las Vegas

REFERENCE

Johnson, D. E., (1998), *Applied Multivariate Methods for Data Analysis*, Belmont, CA: Duxbury Press.

The Process Evaluation Handbook, by Donald J. Wheeler, Knoxville, TN: SPC Press, 2000, ISBN 0-945320-55-8, xii + 250 pp., \$35.

What is *The Process Evaluation Handbook*? Or more specifically, what major issue does it address? What is its *raison d'être*? These are questions I attempted to answer while reading the book, and they are questions that remain largely unanswered. For starters, the name is inappropriate. The book only superficially covers process evaluation. It could be called *Statistica Mensa* (since it has more tables than Home Depot).

The single-page Preface and Introduction sections offer a nebulous description of what the book is about—saving money. The author stresses that the book is not about capability indices. For this we are fortunate because *capability* is not defined until the third chapter (and only implicitly there). There is no mention of the intended audience or the necessary background required for reading the book. A working knowledge of integral calculus is, however, absolutely necessary. Given the dearth of definitions and the succinctness of the chapters (none longer than 20 pages), a background in statistical process control (SPC) from a book at the level of Montgomery (1997) prior to tackling this one is strongly recommended. One can only speculate regarding the choice of textbook the author would recommend for this purpose; the book is devoid of citations.

Chapters 1–3 cover the basic motivation for the book, which is the ability to reduce costs by implementing some quantitative process improvement method(s). One can determine the potential savings to be achieved through the implementation and ostensibly determine whether or not to proceed. There is no prescription regarding the specific choice of quantitative method to be applied, though several are listed as potential candidates. The author apparently assumes that the reader is literate regarding methods such as algorithmic SPC and autoregressive integrated moving average models, an assumption that is rarely warranted in practice. Practitioners who are at a stage prior to SPC implementation (the stage at which this book is appropriate!) likely have little knowledge of $\bar{X}$ -bar and R charts, much less algorithmic SPC.

There are several problems (or to use quality-speak, opportunities for later improvement) in the first three chapters. The author employs a Taguchi-type quadratic function to represent the costs incurred as the measurements deviate from target. No theoretical or empirical justification is given for this choice. This quadratic function is used to compute all of the table values. If a linear (or other nonquadratic) function is more appropriate, the tables are useless. In the linear case, the quadratic function will grossly underestimate the cost between the target and the specs, and the results from using the approximation will be pernicious.

The possible occurrences of scrap and rework are not exhausted in the chapters and tables. It is assumed here that *all* out-of-spec items on a given side of the target will be either scrapped or reworked. It is entirely possible in many applications (several of which I have directly observed) that items closer to the spec might be reworked and those farther away scrapped. Furthermore, only a single quantitative measurement is being considered. Rarely are items scrapped or reworked based only on a single measurement, and usually several measurements are taken on each item. For example, we could measure fabric weight, thickness, tensile strength, color, and fuzziness. It is doubtful that simply working with weight will offer any meaningful process cost evaluation. It would only yield a poor first-order approximation. Perhaps it is deemed that multivariate concerns are beyond the scope of the current book, but some mention of this serious limitation should be made.

The discussion of process feedback and adjustment systems in Chapter 3 is oversimplified. Some time series processes should be controlled using automated process controllers, and others (such as iid processes) should not. This is never discussed here. Likewise, the discussion of current and upgraded processes is not lucid. In certain processes, the machinery is responsible for a large amount of variation. The formulas in the book cannot determine whether the current process variation is due to controllable or uncontrollable factors. Those due to uncontrollable factors (i.e., intrinsic machinery limitations) may be fixed only through process-wide changes. Changes in the process due to factors such as upgrades in machinery will also cause the process cost structure to change, and this fact is not considered.

Chapters 4–6 offer a continuation of the first three chapters covering the special circumstances of one-sided specifications and targets at a boundary. The tables included in the book suffice to cover these situations. Each of these chapters closes with helpful worksheets for the practitioner implementing the techniques.

The final chapter covers count data and process specifications. There is a problem in the introduction of this topic. The specifications on percent nonconforming are often determined (and should be) based on a trade-off between the cost of nonconformance and the cost of decreasing this percentage in the short run. (Note that the producer sets these specs.) The author takes these specs as a given and proceeds to find potential economic benefits. This creates a potential looping problem because the spec settings have economic ramifications of their own. The issues regarding setting your own specs versus the customer setting them (and the economic ramifications thereof) probably warrant a separate chapter.

The treatment of costs in the example here lacks clarity. The situation in question is a setting in which items are rejected if they weigh less than some threshold amount. The author asserts that “the excess cost function can be thought of as a step function.” Items that fail to meet the threshold will cost the amount of materials and production, but this cost decreases (probably linearly) as we move away from the threshold. Similarly, for weights above the threshold, the cost grows (again, probably linearly) as we move away from the threshold (by the amount the author refers to as “giveaway”). Although it is true that the cost jumps from 0 at the threshold to the cost of materials and production for a value just smaller than the threshold, this is not a step function. Perhaps it could be called a “displaced escalator.”

The remainder of the book is composed of tables and worksheets. There is a one-page discussion of $\bar{X}$ charts using the moving range. Mention is made of the mean squared successive differences (MSSD) estimator, but (surprise!) no definition is offered. Even many SPC-fluent practitioners have rarely heard of the MSSD estimator. The tables show the effective costs of production for the scenarios given in the previous chapters. Finally, there are a few tables for normal curve areas, bias-correction factors, and a few control charts. Given the background expected of the reader, it is likely that he or she will already possess most if not all of these table values.

The strength of the book lies in the examples and worksheets. They are all easy to follow (with the exception of the example in Chap. 7), and a practitioner will find the worksheets especially helpful and apropos given the various ISO requirements regarding documentation of statistical technique usage.

The major weakness of the book lies in the complete lack of references and definitions. There is no index in the back, so losing one's place creates an uneasy moment or two. It is highly likely that one would need to cross-reference other books while using this one.

It is my opinion that, in general, economic aspects of quality should be continually studied, and the current book offers that much. Other recent studies (e.g., see Hegji 2000) should also be given attention. Perhaps this area will undergo the Hegelian dialectic and produce some noteworthy works in the future.

Overall, *The Process Evaluation Handbook* offers only marginal benefits, and the weaknesses far outweigh them. (Of course, in this age \$35 for marginal benefits is not half bad.) I would recommend this book only for the consultant with the specific need for using these tables on an ongoing basis. The current book contributes little to the corpus of the quality literature and does so with limited detail. Very few practitioners will find this book useful.

John D. Barrett
University of North Alabama

REFERENCES

- Hegji, C. E. (2000), "On the Economics of Choosing Product Quality," *The Southern Business and Economic Journal*, 23, 112-123.
- Montgomery, D. C. (1997), *Introduction to Statistical Quality Control*, New York: Wiley.

Measuring the Software Process, by William A. Florac and Anita D. Carleton, Reading, MA: Addison-Wesley, 1999, ISBN 0-201-60444-2, xxii + 250 pp., \$54.95.

The concepts, methods, and practices associated with process management and continual improvement have recently gained wider acceptance within the software community. In this book, Florac and Carleton further this advancement by detailing the use of statistical process control (SPC) to manage and improve software processes, where software process refers to any process or subprocess used by a software project or organization. By establishing a framework and providing the necessary detail to implement SPC, this book demonstrates how quantitative insights into the behavior of software processes can be gained and used to achieve both business and technical goals.

Chapter 1 provides an introduction to the basic concepts of managing, measuring, controlling, and improving software processes and also presents the rationale behind the use of control charts. The first chapter concludes by outlining a framework for measuring and improving process behavior through activities that include clarifying business goals, identifying and prioritizing issues, selecting and defining measures, collecting, verifying, and retaining data, analyzing process behavior, and evaluating process performance.

Chapter 2 discusses those activities that are prerequisites to collecting and analyzing process performance data. These activities include identifying process management issues, selecting and defining the measures, and integrating the defined measures with the organization's software process.

Chapter 3 provides specifics associated with collecting, verifying, and retaining software process data. Methods and tools for obtaining data, execution of data-collection procedures, and the capture and recording of data are all provided.

In Chapter 4, the basics of Shewhart control charts, including the distinction between signals and noise, evaluation of process stability, the structure of control charts, and methods for detecting anomalous process behavior, are presented. Chapter 5 further expands on these concepts by providing the procedures and formulas for constructing standard control charts for both variables and attributes data. Chapter 6 continues by discussing a number of topics that need to be considered using control charts. These topics include considerations such as amount of data necessary to construct a control chart, rational sampling and subgrouping, and aggregation of data.

Although the previous chapters provided the mechanics for constructing and analyzing control charts, Chapter 7 gives guidance on appropriate actions to take after construction and analysis. These actions include the removal of assignable causes, process changes, and improvement methods.

The book concludes with Chapter 8, which provides 10 steps for initiating the application of SPC in a software environment. This chapter also includes the answers to some frequently asked questions about SPC.

Throughout the chapters, examples are provided to further clarify the use of SPC in software processes. Additional features of the book include tables for calculating control-chart limits, detailed discussions addressing SPC fundamentals, and an extensive list of references.

Florac and Carleton state that the book is appropriate for software managers or practitioners who have responsibility for product quality or process performance. By starting with the basics of SPC, this book allows those software managers or practitioners with minimal SPC experience to become familiar with the key issues, concepts, and techniques for applying SPC within their organization. For practitioners or statisticians who are experienced with the concepts and applications of SPC, the greatest use of this book will be the framework and the examples of appropriate metrics for measuring software processes.

Marcey Abate
Sandia National Laboratories

Environmental Statistics With S-PLUS, by Steven P. Millard and Nagaraj K. Neerchal, Boca Raton, FL: CRC Press, 2001, ISBN 0-8493-7168-6, 830 pp., \$89.95.

Environmental Statistics With S-PLUS is a textbook with a thorough coverage of basic data analysis, but it also ventures into more complex areas such as risk assessment and spatial statistics. Although the use of the relevant S-PLUS packages is described throughout, the book is basically about statistics, not software: It is statistics "with" rather than "in" or "for" S-PLUS. It is allegedly intended for just about anyone; in practice it would be suitable for many people with a background in engineering or environmental health.

The book starts with a discussion of sampling, pointing out common pitfalls and encouraging thought about the purpose of the sampling. This, together with the subsequent chapter on descriptive statistics, provides a good foundation. A substantial proportion of the rest of the book deals with tolerance, prediction, and two-sample comparison in a variety of situations, including censored and truncated data, followed by single chapters on linear models, time series, spatial statistics, and risk assessment. Examples and simulation studies illustrate the topics, and there is often discussion of the practical difficulties that are encountered in applying these methods. Each chapter contains a set of exercises. These are mostly based on datasets presented in the book, many of which are also conveniently provided in the S-PLUS extension packages.

Much of the discussion relies on parametric assumptions that might appear old-fashioned in many other areas of statistics. This is fairly common in environmental statistics and may be partly due to the pressure for conformity from litigation and regulation and partly due to the desire to have methods that at least appear to work for small numbers of costly samples. Nonparametric methods are provided, but their need for larger sample sizes is cited as a disadvantage (rather than as an indication that parametric methods may be overly optimistic).

Some good points are made that are usually missed in introductory books. In discussing two-sample tests, for example, the authors point out that the Wilcoxon test has poor power for detecting a shift in a subpopulation, a fact that more statisticians should be aware of. The difference between uncertainty and variability is emphasized in the chapter on risk assessment; this is reasonably familiar in environmental statistics but ignored in many other areas.

On the other hand, some opportunities are missed. The balance between Type I and Type II errors is discussed well in the introductory material and in other places. Given this, it is surprising when the authors discuss compliance monitoring, compute lower confidence limits for the 95th percentile, and baldly state (pp. 277 and 280) that the EPA guidance document "incorrectly computes a one-sided upper confidence limit." Surely this is just a question of whether the null hypothesis is "compliance" or "noncompliance" and thus whether the fixed Type I error corresponds to overregulating or underregulating. In another example, there is a good discussion, with simulations and graphs, of the central limit theorem, but its relationship to the robustness of the *t* test, analysis of variance, and linear regression is not mentioned. Instead, the reader is referred to an introductory biostatistics textbook for information on robustness.

This illustrates another deficiency. Many statistical facts and opinions are supported by simulation study or reference to original work, but for a number we are merely referred to another textbook. References to other introductory textbooks are indeed valuable when they are used to provide further information, but references to some actual evidence would be more useful for, for example, the degree to which the *t* test is sensitive to differences in variance or nonnormality.

In general the book is a useful one. Many textbooks share its deficiencies, but its virtues are less common. It has many references to EPA guidance documents as well as to other statistical resources, it describes what S-PLUS is doing as well as how to use it, and it gives much practical guidance. Other books would be needed in addition for any real coverage of linear models, let alone spatial statistics, but a useful introduction to some of the issues is given.

Thomas Lumley
National Research Center for
Statistics and the Environment
University of Washington

Mastering Data Mining, by Michael J. A. Berry and Gordon S. Linoff, New York: Wiley, 2000, ISBN 0-471-33123-6, xviii + 494 pp., \$44.99.

This book had an interesting origin: Previously the authors wrote Berry and Linoff (1997), reported for *Technometrics* by Ziegel (2001). As they note in the introduction (p. xvii) to this new book, a lot has changed with the technology in that period of time. In addition, they note, "We want to address the needs of practitioners on both the business and the technical sides" (ibid.). The book arrived as I was departing for an international conference and a site visit concerning data mining projects, so I read most of it on the trans-Atlantic crossing. It is an easy read. Technical details are not eliminated, but everything is done to make this book totally accessible to anyone in the organization who might need background or have any kind of participation in a data mining project.

The book has three parts. The first part gives all the background on the subject. The first of the four chapters here gives both the business and the technical basis for data mining. The second chapter presents the premise that organizations should treat data mining as a core competency. It should be noted that this perspective is set forth by two authors who are part of a consulting business. The third chapter gives an overview of the data-mining process. The fourth chapter is about customers. The subtitle for the book is *The Art and Science of Customer Relationship Management*, a perspective that influenced the selection of illustrations and case studies but otherwise did not detract from the value of the book for applications that do not involve customers.

The second part has four chapters that provide a readable overview of data-mining methodology. The first of these chapters discusses three major techniques—*k*-means clustering, decision trees, and neural networks. Each topic is handled by discussing the most common approach for using the technique. The next chapter considers the collecting, organizing, and managing of data. There follows a chapter on building predictive models. The two interesting topics in this chapter are the division of the data into test, training, and validation sets and the developing of multiple models for one application. The last chapter in this part has four case studies concerned with the setting up of data mining environments.

The last part of the book has seven case studies. All but one of these are devoted to customer-relationship applications. There is a chapter on improving manufacturing processes that discusses data mining projects at R. R. Donnelley and Time-Warner. Both applications involve the improving of problems in printing plants. These are moderately useful in helping the novice understand how to be a useful participant in a data mining project.

This book deals in passing with the relationship of both statistics and statisticians to the data mining process. It certainly promotes a role for the statistically knowledgeable participant in the data mining effort. This is actually a good background book for statisticians, despite the lack of any references of any kind. Statisticians will learn the process and find direct and indirect insights into their possible roles in data mining efforts.

Eric R. Ziegel
BP

REFERENCES

- Berry, J., and Linoff, G. (1997), *Data Mining Techniques*, New York: Wiley.
Ziegel, E. (2001), Editor's Report on *Data Mining Techniques*, by J. Berry and G. Linoff, *Technometrics*, 43, 114.

A Handbook of Statistical Analyses Using Stata (2nd ed.), by Sophia Rabe-Hesketh and Brian Everitt, Boca Raton, FL: Chapman & Hall/CRC, 2000, ISBN 1-58488-201-8, 210 pp., \$39.95.

This book is neither a statistical textbook nor a Stata software reference manual. Rather, *A Handbook of Statistical Analyses Using Stata* is more of a concise Stata "how-to" manual that takes readers through the analysis of a dozen different datasets, each requiring the application of one or more statistical methods. It progresses from producing simple descriptive statistics in Stata through the Stata commands for more advanced analytical techniques such as generalized linear models, survival analysis, and maximum likelihood estimation.

In its second edition, *A Handbook of Statistical Analyses Using Stata* is well written and nicely organized. It is designed for readers either already familiar with the statistical techniques it presents or for those sufficiently sophisticated to be able to learn the requisite statistical details from other sources. Readers with solid statistical training but unfamiliar with Stata will find this book to be a succinct and pleasant introduction to the software. As the authors state in their Preface, "Our hope is that this approach will provide a useful complement to the excellent but very extensive Stata manuals." The second edition is an update of the original edition from Stata 5 to Stata 6.

In the first chapter, the reader is given a concise overview (24 pages) of the Stata interface and syntax. This chapter provides all of the fundamental commands necessary to become immediately functional in the software. Then, each subsequent chapter presents a problem and dataset followed by a step-by-step analysis in Stata. The datasets are largely taken from clinical trials or epidemiological studies. Although each chapter does attempt to briefly describe the particular statistical technique used, the level of detail varies and is often relatively superficial. However, this allows the authors to get right to the point—presenting Stata in the context of real problems and real analysis.

The scope and focus of the book is best illustrated by its table of contents:

- Chapter 1: A Brief Introduction to Stata
- Chapter 2: Data Description and Simple Inference: Female Psychiatric Patients
- Chapter 3: Multiple Regression: Determinants of Pollution in U.S. Cities
- Chapter 4: Analysis of Variance I: Treating Hypertension
- Chapter 5: Analysis of Variance II: Effectiveness of Slimming Clinics
- Chapter 6: Logistic Regression: Treatment of Lung Cancer and Diagnosis of Heart Attacks
- Chapter 7: Generalized Linear Models: Australian School Children
- Chapter 8: Analysis of Longitudinal Data I: The Treatment of Postnatal Depression
- Chapter 9: Analysis of Longitudinal Data II: Epileptic Seizures and Chemotherapy
- Chapter 10: Some Epidemiology
- Chapter 11: Survival Analysis: Retention of Heroin Addicts in Methadone Maintenance Treatment
- Chapter 12: Principal Components Analysis: Hearing Measurement Using an Audiometer
- Chapter 13: Maximum Likelihood Estimation: Age of Onset of Schizophrenia

Readers interested in a further description of each chapter should consult the recent review by Lachenbruch (2000). Although this review was for the first edition, the chapter descriptions are still relevant. (See also Cox 2000.)

A Handbook of Statistical Analyses Using Stata could be useful in filling specialized pedagogical niches. Although the text assumes a level of technical statistical sophistication that makes it unsuitable for introductory or lower-level courses, it might form the basis for an advanced undergraduate or graduate course in applied data analysis. It could be particularly relevant in a curriculum in which students must first take a number of theoretical statistical courses, especially in schools of public health and other health-related settings, followed by a course focused on applied data analysis. It could also be used as a supplementary textbook for other courses that require the use of Stata.

Each chapter has exercises at the end. However, I found the exercises relatively simple and limited to posing minor variants of the analysis presented within the chapter, mainly to motivate the use of alternate Stata syntax. These exercises were sufficient in the context of learning to use Stata. However, I suggest that expanding the exercises would improve the book's usefulness in the classroom. In particular, I would like to see new exercises for each chapter based on a second dataset that is not used in the chapter. This dataset would be similar in scope and analytical requirements to the one presented in the chapter but also sufficiently different so that students could do their own analysis (prompted by the exercise questions) from scratch. This would allow the instructor to use the chapter dataset and analysis in a lecture and yet still have other material for student homework and exercises.

Readers should note that the book is focused on the "Intercooled" version of Stata for Windows. Moreover, in the never-ending race of software manuals to keep up with the latest software revisions, Stata 7 has recently been released, which may date some of the material presented. However, since I do not have access to Stata 7 as of this writing, I cannot judge how much different that version is and how much of an effect those changes will have on this

book. Nonetheless, this book provides an excellent, interesting, and efficient introduction to Stata that I can easily recommend. And, I am sure a third edition is already in the works.

Ronald D. Fricker, Jr.
RAND

REFERENCES

- Cox, N. (2000), "Book Review: *A Handbook of Statistical Analysis Using Stata*," *Biometrics*, 56, 650–651.
Lachenbruch, P. A. (2000), "Book Review: *A Handbook of Statistical Analysis Using Stata*," *The American Statistician*, 54, 153–154.

The Theory of the Design of Experiments, by D. R. Cox and N. Reid, Boca Raton, FL: Chapman & Hall/CRC, 2000, ISBN 1-58488-195-X, 323 pp., \$69.95.

Intended for anyone concerned with the theoretical issues in the design of experiments, this well-organized book of 224 pages (8 chapters) of text and 74 pages of appendixes provides a clear account of the major topics in the area. The underlying theory of design of experiments is set out systematically, illustrations from a broad range of areas of application are taken to motivate the discussion, and examples are given to show some detailed methods. Bibliographic notes, further results, and exercises are presented to sketch historical and recent developments.

Chapter 1 gives a concise description of some general concepts in the design of experiments. They include types of investigation, various forms of study such as prospective study and retrospective study, and requirements and key steps in design. A simplified model is given for the formulation of experimental design largely used in the book.

Chapter 2 focuses on the issue of bias removal. Randomization and retrospective adjustment, two commonly used methods in reducing bias, are discussed in detail.

Chapters 3 and 4 study the control of haphazard variation. Chapter 3 discusses five methods of reducing the effect of haphazard variability—to use more uniform material and more internal replication, to use more experimental units, to use the technique of blocking, to use retrospective adjustment, and to use special models. However, the emphasis has been placed on the method of blocking. Matched pair design and randomized block design are studied in detail in this chapter.

Chapter 4 is a continuation of the study of blocking techniques, with Latin squares, incomplete block designs, and crossover designs being covered.

Chapters 5 and 6 deal with factorial designs. In Chapter 5 the basic ideas involved in factorial experiments are discussed. They include main effects, interactions, contrasts, and extracting information from fractions of the full factorial designs. In Chapter 6, various more specialized topics related to factorial experiments are covered. They include confounding, factors at more than two levels, orthogonal arrays, split-plot designs, and response surface methods. Taguchi methods are also discussed.

Chapter 7 concerns optimal designs. It begins with some simple but motivating examples. Then the general equivalence theorem of Kiefer and Wolfowitz is presented. Algorithms for finding optimal designs are discussed. The issue of optimality under the framework of nonlinear designs, space-filling designs, and Bayesian designs is also addressed.

Chapter 8 talks about some additional topics. The issues of number of experimental units and sampling within units are studied first. Then, according to different situations, the following designs are studied: adaptive designs, sequential regression designs, designs for one-dimensional error structure, and spatial designs.

Three appendixes cover the material supplementary to the main text. Appendix A presents some important results on the linear model and analysis of variance. Appendix B introduces some key algebraic ideas. Appendix C discusses the use of S-PLUS by providing code for the main examples in the text.

In summary, this book is an excellent addition to the literature. It can serve as a cornerstone in a graduate student's exploration in the theoretical aspects of experimental design and is a valuable reference for statisticians working in medicine, agriculture, the physical sciences, and other areas of biometry and industry.

Dechang Chen
University of Wisconsin-Green Bay

Editor Reports on New Editions, Proceedings, Collections, and Other Books

This section reports on new editions of books previously reviewed in *Technometrics*, collections of papers and conference proceedings, and other statistics books that should have some interest for the readership. Selections and comments do not represent any perspective of the editor's employer or of the sponsoring societies.

Eric R. ZIEGEL
BP

Introduction to Statistical Quality Control (4th ed.), by Douglas C. Montgomery, New York: Wiley, 2001, ISBN 0-471-31648-2, xix + 796 pp., \$114.95.

The third edition (3E) of this book, which first appeared around 1984 as a very modest-sized introduction to the use of control charts and acceptance sampling, was given a full review in *Technometrics* (Gaudard and Ramsey 1997). As the reviewers proclaimed, "It offers improvements over the second edition of sufficient substance to warrant its purchase even by owners of the second edition." The second edition had greatly expanded the focus of the book. One hopes that those people who were motivated by that perspective will not be too distraught by my echoing the opinion of the reviewers of the 3E for this new book. This fourth edition (4E) is more than 100 pages larger than the 3E. Much of this new material does not appear in any other statistical process control (SPC) book. The author has also completely reorganized the 3E "based on my own teaching experiences and extensive feedback from other users of the text" (Preface, p. vii). The author's modesty probably inhibits his also mentioning much consulting time on location with a wide variety of industries.

The reorganization involves the second part of the 3E, which previously was simply called "Statistical Process Control" and contained the book's material on control charts and process capability. The last topic was also the subject of the last chapter in that part of the book. Now the fundamental components of control charting, plus a chapter called "Process and Measurement System Capability Analysis," have been collected in the new Part II "Basic Methods of Statistical Process Control and Capability Analysis." This part of the book is then followed by a new Part III, "Other Statistical Process Monitoring and Control Techniques." The chapter on cumulative sums (CUSUM's) and exponentially weighted moving averages (EWMA's) in the 3E was moved into this part of the book. It is still followed by the chapter "Other Statistical Process Control Techniques." Material on multivariate process monitoring and on engineering process control formerly in this chapter has been moved into two new chapters. There are nearly 100 pages of new material in these chapters. They do a great job of integrating into SPC methodology the two topics that are most important for the current engineering practice of SPC. Specific new topics include multivariate EWMA charts, regression adjustment, latent structure methods, and process control by feedback adjustment.

The remaining material on experimental design has been made more compact by combining the 3E chapters on factorial experiments and response surface methods into a single chapter on process optimization. The final two chapters on acceptance sampling and the appendixes are essentially unchanged from the 3E. Supplemental text material, PowerPoint slides, and datasets are available at a Web site and on a CD-ROM. The book uses MINITAB for most of its illustrations. The publisher will package the book with either MINITAB or a set of Excel add-ins for student use or for training classes.

This revision of the book is a fine model for continuing to keep a product viable. The author has done this by modernizing the methodology, adding new material that reflects current practice, condensing some of the existing material that is dated to maintain the book at a reasonable size, and utilizing modern technology, such as CD-ROM's and Web sites, to make additional material available. For similar efforts, see Montgomery (2001) for the author's classic design of experiments textbook and Montgomery, Peck, and Vining (2001) for his newly revised regression book.

REFERENCES

- Gaudard, M., and Ramsey, P. (1997), Review of *Introduction to Statistical Quality Control* (3rd ed.), by D. Montgomery, *Technometrics*, 39, 332.

Montgomery, D. (2001), *Design and Analysis of Experiments* (5th ed.), New York: Wiley.

Montgomery, D., Peck, E., and Vining, G. (2001), *Introduction to Linear Regression Analysis*, (3rd ed.), New York: Wiley.

SAS® System for Regression (3rd ed.), by Rudolf J. Freund and Ramon C. Littell, Cary, NC: SAS Institute (New York: Wiley), 2000, ISBN 1-58025-725-9 (0-471-41664-9), viii + 236 pp., \$59.95.

In the report on the second edition (Ziegel 1992), I encouraged anyone who did not own the first edition (1E) to "call SAS and order this new edition right away." I also noted that "if you have the First Edition, I don't know that you need to buy this new one." Now it has been 10 years since the 2E. The 1E was published in 1986. I will reiterate both comments. Any statistician or SAS user who does regression analysis must own this book. There also is still not a lot of additional material relative to the 1E or the 2E that would encourage the purchase of an updated version for your personal library. In part there is not a lot in classical regression methodology that is new in SAS, or elsewhere for that matter. In addition, there is a large companion volume (Allison 1999, reported by Ziegel 2000 for *Technometrics*), that presents the methodology for the one area in regression analysis that has received a lot of attention in the past 10 years. Another popular new regression procedure, at least in the world of chemistry, PROC PLS, is not discussed at all in the book. All of the SAS procedures that do regression analysis in some way are summarized in Chapter 1, but there is no explanation about what will or will not be discussed in the book.

The basic chapter on PROC REG, Chapter 2, adds a new section on the datasets that are created by the OUTFEST= and TABLEOUT options. There is a brief note about ODS output. Some of the other chapters also comment on ODS output. Chapter 5 was renamed from "Polynomial Regression" to "Curve Fitting" to allow the inclusion of material on nonparametric regression and PROC LOESS. Chapter 6, "Special Applications of Linear Models," has a new section for analyzing binary response variables using PROC LOGISTIC. There is a brand new Chapter 8, "Using SAS/INSIGHT Software for Regression," that goes through all the different regression capabilities of PROC INSIGHT. See Ziegel (1992) and the review for the 1E, Mason (1988), for more information on the content of this book.

Two things that I would have liked to have seen, in addition to discussion of PROC PLS, are more use of SAS/GRAPH for graphics and inclusion of SAS macros for some of the many other regression methodologies, like MARS, for which there are no procedures. The authors eliminated the appendixes in the 2E that had the code for the examples. One can always get that stuff from Web sites. Any further contemplation of the incompleteness of the book as a regression textbook simply reflects the aspects of the methodology that are absent as procedures in SAS.

REFERENCES

- Allison, P. (1999), *Logistic Regression Using the SAS® System*, Cary, NC: SAS Institute.
 Mason, R. (1988), Review of *SAS System for Regression*, by R. Freund and R. Littell, *Technometrics*, 30, 117–118.
 Ziegel, E. (1992), Editor's Report on *SAS System for Regression* (2nd ed.), by R. Freund and R. Littell, *Technometrics*, 34, 500–501.
 ——— (2000), Editor's Report on *Logistic Regression Using the SAS® System*, by P. Allison, *Technometrics*, 42, 323.

Principles of Multivariate Analysis (rev. ed.), by W. J. Krzanowski, New York: Oxford University Press, 2000, ISBN 0-19-850708-0, xxi + 586 pp., \$70 (softcover).

I have been critical previously of a number of second editions wherein very little has been changed from the first edition. This author has a whole new approach for that situation. No changes have been made at all. Instead, the author has added an 18-page appendix "to present a broad overview (of the new developments), tracing the main lines of development that have taken place, focusing on those areas that are likely to have the most endearing impact, and picking out the most useful features from the point of view of the practical user" (Appendix B, p. 544). No change was made to the book

because that "would detract from its original purpose and intended audience." Instead of revision or correction to the first edition, "opportunity has been taken to revisit any aspects of the original treatment that may need amplification or clarification." One would think the author had written the Holy Bible.

The first edition was reviewed for *Technometrics* by Ling (1990). He found the book "to be a useful reference on the subject of applied multivariate methods." He also noted, "I do not believe this book is at all suitable for students encountering multivariate ideas for the first time." Ling (1990) gave a complete summary of the book. As for the new appendix, it is really just a discussion. There are only a few equations scattered among what is otherwise just text.

This book was old-fashioned when it was written. Statistical computing, the centerpiece of multivariate analysis, is generally simply ignored. By now the book seems like one of those Wiley reprints of a classic textbook. Though published in softcover, it comes with a hardcover price.

REFERENCE

- Ling, R. (1990), Review of *Principles of Multivariate Analysis*, by W. Krzanowski, *Technometrics*, 32, 225.

Visualizing Categorical Data, by Michael Friendly, Cary, NC: SAS Institute, 2000, ISBN 1-58025-660-0, xii + 436 pp., \$69.95.

The author is very familiar to me from Friendly (1991), the first SAS book on statistical graphics. SAS used this quote from my report (Ziegel 1993) in all their advertisements: "If there is any SAS book beyond the manuals that every statistician using needs to own, this is it. This book is fantastic and gets my very highest recommendation."

Still, I never can remember to put something quotable in each of my reviews. Friendly (1991) remains the book in the SAS Books for Users Series that I use most regularly. SAS graphics have always been impossible for me to configure off the top of my head.

This book shares some of the same values as Friendly (1991). Appendix A, "SAS Programs and Macros," gives instructions for 39 different procedures. The code is available in the SAS archives. That adds a lot of capability to the toolkit of any SAS user. A second appendix has the SAS code for the datasets. The data are also in the archives.

This is a big heavy book with its printing all on glossy multicolor pages. A short introductory chapter discusses both categorical data and some new graphical tools, such as mosaic displays, sieve diagrams, and trilinear plots. These are all special tools for graphing categorical data.

In Chapter 2, these graphical tools are applied first to discrete data. Though special plots are developed for displaying binomial and Poisson data, a wide variety of distributions are considered. In Chapter 3, categorical data is introduced with two-way contingency tables. Sieve diagrams are used to show patterns of association. A more colorful display for categorical data is the fourfold plot. The chapter also introduces agreement plots, association plots, and trilinear plots, all novel views of data that create a whole new visualization world. Even more elaborate are the mosaic displays that are used in Chapter 4. These are the graphical display tools for n -way tables, which are the subject for the chapter.

The book returns to two-way tables in Chapter 5, which is about correspondence analysis. SAS has a procedure, PROC CORRESP, that does correspondence analysis, which is used for data analysis, but the author also provides a CORRESP macro that produces plots. The CORRESP macro also is applied for displays with data from n -way tables and multiple correspondence analysis. Being recently divorced, I enjoyed the example about marital status and premarital sex. Biplots are also discussed as a visualization method for two-way and three-way tables.

The book reaches more familiar territory for statistics in science and engineering with Chapter 6, "Logistic Regression." There are some nice macros for plotting log odds and probabilities. Likelihood function plots are provided for quantitative predictors. Some novel line and bar graphs provide displays for qualitative predictors. Bubble plots and scatterplot matrices provide enhanced displays for regression diagnostics. All of these tools are applied across a wide range of logistic regression models.

Graphics for tables and models come together in the visualizations for Chapter 7, "Log-linear and Logit Models." As the book has progressed, the focus has evolved from display of data to display of results from and evaluation of the models. A wide variety of models are described in this chapter. The chapter draws on visualizations developed throughout the course of the book.

Friendly (1991) used SAS to do familiar graphical displays. Here the subject is not SAS; rather it is graphical analysis for data and models. This book is a textbook on categorical data analysis and modeling that emphasizes the use of both familiar and novel graphical displays. For users of categorical data-analysis methods in SAS, particularly those familiar with Stokes, Davis, and Koch (2000), whose second edition was reported by Ziegel (in press), this book is another essential part of one's SAS reference library.

REFERENCES

- Friendly, M. (1991), *SAS System for Statistical Graphics*, Cary, NC: SAS Institute.
- Stokes, M., Davis, C., and Koch, G. (2000), *Categorical Data Analysis Using the SAS® System* (2nd ed.), Cary, NC: SAS Institute.
- Ziegel, E. (1993), Editor's Report on *SAS System for Statistical Graphics*, by M. Friendly, *Technometrics*, 35, 101.
- (in press), Editor's Report for *Categorical Data Analysis Using The SAS® System* (2nd ed.), by M. Stokes, C. Davis, and G. Koch, *Technometrics*, 44.

Geostatistics for Environmental Scientists, by Richard Webster and Margaret A. Oliver, Chichester, U.K.: Wiley, 2001, ISBN 0-471-96553-7, xiii+271 pp., \$105.

This is yet another book in the publisher's Statistics in Practice Series that is produced by their U.K. division. These excellent books have mostly not been focused on statistical practice in engineering or the physical sciences and therefore typically have been relegated to the Editor's Reports. Most recently, see the report on Brown and Prescott (1999), reported by Ziegel (2000a). This book covers a topic that has been very important at BP. Various producing, refining, and manufacturing operations occasionally have resulted in the contamination of soil and groundwater. Mapping the data that have been collected in these situations has required the use of geostatistics.

The book claims that it presumes essentially no background from its readers in either advanced math or statistics (p. 8). Curiously, mathematical equations and statistical concepts such as expectations and integrals seem to abound. Generally readers are treated to much more statistical methodology than the environmental scientists that I know would want to see. In addition, the statistical computing perspective is certainly not acceptable here. Readers are expected to use the formulas to write their own computer programs to implement these methods. Who would want to do that? The book has an appendix that illustrates this procedure for Genstat, which is probably not a useful approach for U.S. audiences.

The book has an excellent focus and an effective organization. The authors take the same perspective that I have as a practitioner. First, someone brings you this set of data. Then it is desirable to use geostatistics to make a representation of the situation. It is hoped that this representation will provide some information that will help in making a decision on how to do remediation. In Chapter 2, the development of the methodology starts with an introduction to basic statistics. Chapter 3 is an interlude that presents the inadequacies of some classical interpolation methods as a way of motivating the use of geostatistical tools. Chapter 4 introduces the use of regionalized variables to describe spatial variability. Chapters 5 and 6, respectively, cover the calculation of the variogram from the data and the fitting of the semivariogram model. Another diversion occurs in Chapter 7, where spectral analysis methods are presented. In Chapter 8, kriging is introduced, including ordinary kriging, simple kriging, universal kriging, and lognormal kriging. The book concludes with two additional chapters that enhance the basic kriging capabilities that one acquires from Chapter 8. Chapter 9 presents methods for several variable measurements simultaneously and then adds cokriging to the kriging capabilities. Chapter 10 presents indicator kriging and disjunctive kriging as methods for the important task of determining the probabilities of exceeding thresholds. There is no discussion of the conditional simulations that are so prevalent in geology for this type of application of geostatistics.

This is a great book for any statistician who works with environmental data for which mapping and inferences from mapping are important for making decisions. Overall it seems too much focused on methodology and too little focused on implementation to be a good primer for environmental scientists and engineers, though it probably is about at the same technical level as Hohn (1999), reported by Ziegel (2000b), a comparable book for earth scientists. The applications in the environmental book come from soil science. The book originated as an extension of the geostatistics material that was included in the book by Webster and Oliver (1990), which was reviewed in *Technometrics* by Cohen (1992).

REFERENCES

- Brown, H., and Prescott, R. (1999), *Applied Mixed Models in Medicine*, West Sussex, U.K.: Wiley.
- Cohen, J. (1992), Review of *Statistical Methods in Soil and Land Resource Survey*, by R. Webster and M. Oliver, *Technometrics*, 34, 497–498.
- Hohn, M. (1999), *Geostatistics and Petroleum Geology* (2nd ed.), Dordrecht: Kluwer.
- Webster, R., and Oliver, M. (1990), *Statistical Methods in Soil and Land Resource Survey*, New York: Oxford University Press.
- Ziegel, E. (2000a), Editor's Report on *Applied Mixed Models in Medicine*, by H. Brown and R. Prescott, *Technometrics*, 42, 442–443.
- (2000b), Editor's Report on *Geostatistics and Petroleum Geology* (2nd ed.), by M. Hohn, *Technometrics*, 42, 444.

Statistical Methods in Spatial Epidemiology, by Andrew B. Lawson, West Sussex, U.K.: Wiley, 2001, ISBN 0-471-97572-9, x + 277 pp., \$89.95.

There have been recent reports for books on statistical methods in epidemiology, such as Woodward (1999), reported by Ziegel (2000). Excellent books on spatial statistics abound, most recently that of Griffith and Layne (1999), reviewed for *Technometrics* by Symanzik (2001). Spatial epidemiology methods bring these two methodologies together for the study of the effect of geography on the incidence of disease. For example, do people who live near petrochemical plants or nuclear facilities show greater prevalence for certain cancers? Spatial epidemiology provides methods for studying the effect of environmental variables and even other kinds of explanatory variables on diseases.

The book is laid out in two sections. The first section contains five chapters organized under the heading, "The Nature of Spatial Epidemiology." The first chapter provides a number of examples. There is a chapter that deals with geographical representation and mapping. Two chapters on methods, one showing the relevant statistical models and the other dealing with data analysis and inference, conclude the first part of the book. This material requires some background in mathematical statistics.

The second part of the book is titled "Important Problems in Spatial Epidemiology." The first couple of chapters present problems relating to small areas, clustering of diseases and sources of environmental pollutants as disease hazards. The next two chapters move on to issues on the larger geographic scale, mapping of the incidence of disease and establishing monitoring methods for health surveillance. The last two chapters present modeling issues for ecological analysis and for infectious diseases. There are several technical appendixes, including one on software.

REFERENCES

- Griffith, D., and Layne, L. (1999), *A Casebook for Spatial Statistical Data Analysis*, New York: Oxford University Press.
- Symanzik, J. (2001), Review of *A Casebook for Spatial Statistical Data Analysis*, by D. Griffith and L. Layne, *Technometrics*, 44, 373–376.
- Woodward, M. (1999), *Epidemiology: Study Design and Data Analysis*, Boca Raton, FL: Chapman & Hall/CRC.
- Ziegel, E. (2000), Editor's Report on *Epidemiology: Study Design and Data Analysis*, by M. Woodward, *Technometrics*, 42, 327.

A Course in Time Series Analysis, edited by Daniel Peña, George C. Tiao, and Ruey S. Tsay, New York: Wiley, 2001, ISBN 0-471-36164-X, xvii + 460 pp., \$115.

This book was derived from lectures that were given for a time series analysis course that was presented by the European Courses in Advanced Statistics for five days in September 1997. The Preface notes (p. xv) that the lectures are focused on both academic teachers or researchers and on statistics professionals. The intent for each topic in the book is "to present an overview of the current status of the area." So, despite the title, this is not really a good textbook for a course in statistics. Each of the 16 chapters has a different author or authors. Some of the participants have written more than one chapter as author or coauthor. The Preface (ibid.) even notes that "for a basic course this book should be complemented with some of the excellent texts available." However, the book would certainly be an excellent complement, particularly for advanced and current topics, for a first graduate course in time series analysis.

Following an introductory chapter full of some great illustrations, the book then unfolds in three parts. The first part is about univariate time series. Its introductory chapter discusses autocorrelation, spectra, and Kalman filters. Then a basic chapter on autoregressive moving average (ARMA) models is followed by several chapters on a sequence of related topics—model fitting and checking, prediction and model selection, outliers, influential and missing observations, automatic modeling, season adjustment, and signal extraction. Most of this material relates directly to ARMA modeling. This body of material also carries the content through about three-fifths of the book. The editors note in the Preface (ibid.) that they tried to keep the notation consistent and create linkages across the different chapters. One advantage of the multiple authors is that each seems to have brought along an example, so the chapters all have a variety of good illustrations.

The second part of the book presents a number of advanced topics for univariate models. These are methods that take time series analysis well beyond ARMA modeling. The first of these chapters, Chapter 9, is about heteroscedastic models, which are useful in modeling situations involving a lot of volatility. These include autoregressive conditional heteroscedasticity (ARCH), generalized ARCH, and several other types of models for the evolution of conditional variance. Chapter 10 is about nonlinear time series models, for which it provides a very nice practical overview. Chapter 11 deals with Bayesian time series methods, Chapter 12 with nonparametric time series analysis, and Chapter 13 with neural network models. All five chapters in this section are excellent introductory presentations of topics that are usually either advanced topics in a time series book or subjects for very advanced books about those topics. The nonparametric methods are various types of regression procedures. The methods for modeling with neural networks include multilayer perceptrons, autoregressive neural networks, and recurrent perceptrons.

The third and final part of the book moves on to multivariate methods. Here the topics for the chapters are vector ARMA models, cointegration in the vector autoregressive model, and linear dynamic multi-input/multi-output systems. These relatively short chapters again feature mostly straightforward introductions to some typically advanced topics. There is never any consideration of how one uses software to do any of the modeling. There are a lot of references. I would have given the book a different title, something like *A Survey of Methods for Time Series Analysis*. As long as the price is not an issue, the book is a nice addition to anyone's time series library.

Seasonal Adjustment With the X-11 Method, by Dominique Ladiray and Benoit Quenneville, New York: Springer-Verlag, 2001, ISBN 0-387-95171-7, xxii + 225 pp., \$59.95.

This methodology was familiar to me as the focus of the third computer program in the Amoco statistics program library of the 1970s. The library also included LINWOOD and NONLINWOOD, two well-known regression programs that were described by Daniel and Wood (1980). The X-11 method actually was developed in the 1960s at the U.S. Bureau of the Census and apparently still survives today. The book also discusses an autoregressive integrated moving average (ARIMA) variant, X-11 ARIMA, from Statistics Canada, and X-12 ARIMA, another product of the Bureau of the Census.

Software such as SAS/ETS, PROC X11, is still around for implementing the procedures.

Most of the book is contained within a single chapter called "The Various Tables," which is Chapter 4. There are, unbelievably, 74 tables organized to carry out the five stages (B, C, D, E, and F) of estimation—(B) preliminary estimation of extreme values and calendar effects, (C) final estimation of same, (D) final estimation of different seasonal components, (E) components modified for large extreme values, and (F) seasonal adjustment quality measures. These tables all provide the means for carrying out the iterative procedures of the X-11 methods.

The book begins with a brief history of seasonal adjustment. The chapter on the tables is preceded by chapters outlining the X-11 method and giving the details of its methodology of using moving averages to estimate the trend-cycle and seasonality for a time series. The tables work with an example with 10 years of the monthly index of industrial production in France. The book concludes with a chapter on modeling the effect of the Easter holiday. Easter apparently introduces a blip in a number of important economic sectors, and the effect occurs irregularly in both monthly and quarterly data.

REFERENCE

Daniel, C., and Wood, F. (1980), *Fitting Equations to Data* (2nd ed.), New York: Wiley.

Optimization Heuristics in Econometrics, by Peter Winker, West Sussex, U.K.: Wiley, 2001, ISBN 0-471-85631-2, xiii + 333 pp., \$79.95.

This book is not so esoteric for industrial statisticians as one might infer from its title. First, it is more about optimization heuristics than about econometrics. Optimization heuristics apply to any situation in which optimization methodology is needed. Second, the applications that are from econometrics mostly have analogous situations in the physical and engineering sciences.

This also is a very fine textbook. It does an excellent job of motivating one's interests in optimization heuristics. It presents some considerable detail on the method for applying its optimization heuristics, called threshold accepting, without wallowing in a large amount of mathematics and methodology. The last chapter in this part of the book, the second part, which focuses on heuristics, is titled "A Practical Guide to the Implementation of Threshold Accepting." More than half of the book is devoted to applying the techniques to five applications, the last five chapters and the last part of the book.

The book begins with four introductory chapters. These intend to "provide the reader with a first intuitive understanding as to why optimization heuristics might be a useful tool in statistics and econometrics and what the main ingredients for a good implementation of such heuristics are" (Preface, p. xii). The author believes that "the standard tools applied to optimization problems are often not appropriate" because "they provide an approximate solution of unknown quality or use heavily simplifying assumptions in order to obtain a model which can be tackled by standard methods" (Introduction, p. 1). The last chapter in the first part of the book contrasts the classical optimization paradigm with optimization heuristics. These include procedures that are already familiar tools in modern applied computing, such as genetic algorithms, simulated annealing, and neural networks. These methods meet the requirements of "nearly optimum or close to optimum solutions," "robust to changes in problem characteristics," and "easier to implement." (ibid.)

Some of the economics and econometrics applications are familiar statistics applications in engineering. There is a chapter on experimental design. This one is somewhat technical and includes some of the more overtly mathematical designs, such as arrays. There is also a chapter on censored quantile regression. The last applications chapter is about continuous global optimization. There are also a couple of econometrics application chapters, one on identifying the multivariate lag structures in building models for vector autoregressive processes, the other on optimal aggregations in macroeconomic models.

Customer Centered Six Sigma, by Earl Naumann and Steven H. Hoisington, Milwaukee: ASQ Quality Press, 2001, ISBN 0-87389-490-1, xx + 340 pp., \$55.

The first author is widely known for his book on customer satisfaction (Naumann and Giel 1995), reported in *Technometrics* by Ziegel (1997), and

his extensive consultancy and teaching concerning customer-satisfaction measurement and improvement. Here he combines that interest with the current preferred methodology for business process improvement, the Six Sigma process. The cover of the book has "Six Sigma" with a font several times larger than the font for "Customer Centered," so it is somewhat more about Six Sigma than customer satisfaction. It carries the subtitle *Linking Customers, Process Improvement, and Financial Results*.

The book has four parts. The first part, "Why Six Sigma," has two chapters that connect Six Sigma with customer satisfaction. Both concepts are argued to be key ingredients to financial performances. The second section, "Capturing the Voice of the Customer," which constitutes about one-third of the book, offers the substance of the first author's acknowledged expertise in teaching businesses how to do customer-satisfaction surveys, evaluate customer loyalty, handle complaints, and build relationships. There is very little in any of these chapters that connects customer-satisfaction improvement directly to the Six Sigma process.

Part 3, "Preparing the Organization for Six Sigma," has three excellent chapters that very effectively describe the actions necessary to implement the Six Sigma process. The first chapter discusses making linkages to mission statements and corporate strategies and using balanced scorecards. The next chapter deals with selecting the right processes for the initial Six Sigma effort, and the last chapter is concerned with forming teams for Six Sigma projects. These authors do not promote the roles of the black belts and green belts who usually lead these efforts. The last part of the book has several chapters on Six Sigma tools. The tools are placed within the context of several process-improvement steps—aligning with customer needs, analyzing the process, solving problems, presenting results, implementing changes, and determining process capability.

The last chapter in the part concerned with Six Sigma tools includes a lot of material about control charts, but generally there is a minimal and low-level use of statistical methods in this book. Each chapter concludes with another installment of a running discussion of the Six Sigma implementation at Johnson Controls.

The authors have created a different type of Six Sigma book. The process has been adapted to meet their philosophies and interests rather than being presented in its conventional form. There are no references to any of the recent Six Sigma books, such as that of Pyzdek (1999), reported in *Technometrics* by Ziegel (2000).

REFERENCES

- Naumann, E., and Giel, K. (1995), *Customer Satisfaction Measurement and Management*, Cincinnati: Thompson Executive Press.
- Pyzdek, T. (1999), *The Complete Guide to Six Sigma*, Tucson, AZ: Quality Publishing.
- Ziegel, E. (1997), Editor's Report on *Customer Satisfaction Measurement and Management*, by E. Naumann and K. Giel, *Technometrics*, 39, 239–240.
- (2000), Editor's Report on *The Complete Guide to Six Sigma*, by T. Pyzdek, *Technometrics*, 42, 441–442.

Forthcoming Reviews

Books listed here have been assigned for review in the past quarter. Publication of their reviews or reports generally would occur within the next four issues of the journal. Persons interested in reviewing specific books must notify the editor by the publication date for the book. Persons interested in being reviewers should contact the editor by electronic mail (ziegeler@bp.com).

- Analysis of Time Series Structure*, by Nina Golyandina, Vladimir Nekrutkin, and Anatoly Zhigljavsky, Chapman & Hall/CRC
- Applied Statistics With Microsoft® Excel*, by Gerald Keller, Duxbury
- Applying and Interpreting Statistics* (2nd ed.), by Glen McPherson, Springer-Verlag

- The Bayesian Choice* (2nd ed.), by Christian P. Robert, Springer-Verlag
- Bayesian Statistical Modelling*, by Peter Congdon, Wiley
- Binomial Distribution Handbook for Scientists and Engineers*, by Elart von Collani and Klaus Dräger, Birkhäuser
- Causation, Prediction, and Search* (2nd ed.), by Peter Spirtes, Clark Glymour, and Richard Scheines, MIT Press
- Collecting Spatial Data: Optimum Design of Experiments for Random Fields* (2nd ed.), by Werner G. Müller, Physica-Verlag
- Curve Ball*, by Jim Albert and Jay Bennett, Copernicus Books
- Data Mining Cookbook*, by Olivia Parr Rud, Wiley
- Fundamentals of Modern Statistical Methods*, by Rand R. Wilcox, Springer-Verlag
- Independent Component Analysis*, edited by Stephen Roberts and Richard Everson, Cambridge University Press
- Introduction to Linear Regression Analysis* (3rd ed.), by Douglas C. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining, Wiley
- Modern Statistics for Engineering and Quality Improvement*, by John Lawson and John Erjavec, Duxbury
- Multivariate Statistical Modelling Based on Generalized Linear Models* (2nd ed.), by Ludwig Fahrmeir and Gerhard Tutz, Springer-Verlag
- Nonlinear and Nonstationary Signal Processing*, edited by W. J. Fitzgerald, R. L. Smith, A. T. Walden, and P. C. Young, Cambridge University Press
- Optimal Reliability Design*, by Way Kuo, V. Rajendra Prasad, Frank A. Tillman, and Ching-Lai Hwang, Cambridge University Press
- Probabilistic Risk Analysis*, by Tim Bedford and Roger Cooke, Cambridge University Press
- Probability: The Science of Uncertainty*, by Michael A. Bean, Brooks/Cole
- Quality Improvement With Design of Experiments*, by Ivan N. Vuchkov and Lidia N. Boyadjieva, Kluwer
- Regularized Radial Basis Function Networks*, by Paul V. Yee and Simon Haykin, Wiley
- Reliability Improvement With Design of Experiments* (2nd ed., revised and expanded), by Lloyd W. Condra, Marcel Dekker
- Soft Computing*, by L. Fortuna, G. Rizzotto, M. Lavorgna, G. Nunnari, M. G. Xibilia, and R. Caponetto, Springer-Verlag
- SPC: Practical Understanding of Capability by Implementing Statistical Process Control*, by James C. Abbott, Robert Houston Smith
- Statistical Consulting: A Guide to Effective Communication*, by Janice Derr, Duxbury
- Statistical Methods in Bioinformatics*, by Warren J. Ewens and Gregory R. Grant, Springer-Verlag
- Statistical Process Control and Quality Improvement*, by Gerald M. Smith, Prentice Hall
- Statistics: The Exploration and Analysis of Data* (4th ed.), by Jay Devore and Roxy Peck, Duxbury
- The Subjectivity of Scientists and the Bayesian Approach*, by S. James Press and Judith M. Tanur, Wiley
- Time-Series Forecasting*, by Chris Chatfield, Chapman & Hall/CRC
- Trusting Measurement Results in the Chemical and Process Industries*, by ASQ Chemical and Process Industries Div. Chemical Interest Committee, ASQ Quality Press
- Workshop Statistics: Discovery With Data, A Bayesian Approach*, by James H. Albert and Allan J. Rossman, Key College Publishing (Springer-Verlag)

Seasonal Adjustment with the X-11 Method

Ladiray, D.; Quenneville, B.

2001, XXII, 256 p., Softcover

ISBN: 978-0-387-95171-3