



American Society for Quality

Ranked Set Sampling: Theory and Applications by Zehua Chen; Zhidong Bai; Bimal K. Sinha

Review by: Chunming Zhang

Technometrics, Vol. 47, No. 1 (Feb., 2005), pp. 100-101

Published by: [American Statistical Association](#) and [American Society for Quality](#)

Stable URL: <http://www.jstor.org/stable/25470950>

Accessed: 15/07/2013 02:49

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Statistical Association and American Society for Quality are collaborating with JSTOR to digitize, preserve and extend access to *Technometrics*.

<http://www.jstor.org>

To summarize, *Basic Statistics and Data Analysis* is not suitable for students majoring in math, science, and engineering, because of the sparse coverage of statistical topics applicable to these fields. As a general audience text, the book is well written in terms of style and readability. However, the instructor would have to be predisposed to include a heavy dose of nonparametric statistics in an introductory course and plan on presenting clearer guidelines on when such tests are appropriate.

Robert V. BRILL
Astaris

Statistics for Research (3rd ed.), by Shirley DOWDY, Stanley WEARDEN, and Daniel CHILKO, Hoboken, NJ: Wiley, 2004, ISBN 0-471-26735-X, xvi + 627 pp., \$94.95.

Statistics for Research is written to accommodate a two-semester introductory statistical methods course for graduate students coming from various research disciplines in the natural and social sciences. A strong mathematical background is not assumed, and throughout the text mathematical rigor is substituted with a conceptual approach to teaching statistical concepts. However, to stay away from what the authors call a "black box" approach, they do include examples in which the details of the calculations are presented.

There are several additions to the third edition that are worth noting. First, besides discussing SAS software, the authors also discuss the use of JMP, a Windows-based statistical software package developed by the SAS Institute. Second, the authors have added a few sections early in the text that discuss probability concepts. In the previous editions, probability was discussed as needed throughout the text. Third, a section on logistic regression has been added to the last chapter that briefly introduces the idea of maximum likelihood estimation and goes through one detailed example. Finally, other additions include discussions of retrospective studies, risks, odds ratios and repeated-measures studies, the Bonferroni correction, and ratio and difference estimation. There are also additions to the exercises that appear at the end of each section.

The authors do a very nice job by introducing topics with many examples and illustrations. Shortly after a definition is given, an example is sure to follow that makes the idea hit home in a very concrete manner. The exercises at the end of each section are very well thought through and provide the student with a chance to test his or her understanding. I especially like the true–false review exercises at the end of each chapter. One excellent use of the true–false questions would be to have the students write an explanation about why the question is true or false.

The text comprises 14 chapters:

1. The Role of Statistics. Provides a motivational introduction and basic ideas of inferential statistics.
2. Populations, Samples, and Probability Distributions. Covers random variables, probability distributions and expected value; provides many examples.
3. Binomial Distributions. Incorporates the ideas of type I and type II error nicely within testing a hypothesis using a binomial distribution.
4. Poisson Distributions. Motivates the Poisson distribution and shows its approximation to the binomial distribution.
5. Chi-Square Distributions. Focuses on goodness-of-fit tests, tests of homogeneity and independence; discusses relative risks and odds ratios and briefly discusses the nonparametric median test.
6. Sampling Distribution of Averages. Provides a nice discussion of population mean and variance versus the estimated mean and variance.
7. Normal Distributions. Covers the central limit theorem, confidence interval for the mean, approximations to binomial and Poisson distributions, and the nonparametric rank test.
8. Student's t Distribution. Contains a very nice flowchart for different situations depending on whether the interest lies with testing the mean or variance with continuous data or a proportion with discrete data.
9. Distributions of Two Variables. Covers regression and correlation, with a nice section on computer usage using JMP.
10. Techniques for One-Way Analysis of Variance. Provides a very clear discussion on why ANOVA is concerned with variances when testing means and a nice section on multiple comparison techniques.
11. The Analysis-of-Variance Model. Covers the model, assumptions, and transformations.
12. Other Analysis-of-Variance Designs. Discusses nested, randomized complete block, Latin square, factorial, split-plot, and repeated-measures designs.

13. Analysis of Covariance. Provides very nice examples and explanation of the ANCOVA concept, along with an interpretation of SAS output.
14. Multiple Regression and Correlation. Contains a section on logistic regression.

A selected reading section at the end of each chapter contains a fairly good list of reference material. In addition, answers to odd numbered exercises are given in the back of the text, along with many look-up tables (of, e.g., t and F critical values).

In summary, *Statistics for Research* provides material for a solid two-semester introductory statistics course for graduate students in the natural and social sciences. I could see, however, the need for the text material to be supplemented with lab exercises that clearly show how to perform statistical analysis using statistical software beyond what is done in the text with SAS and JMP. The text is easy to read, and students will enjoy the wide range of examples and illustrations. The additions made to the third edition make it a nice improvement over the first and second editions.

William A. BRENNEMAN
Procter & Gamble Company

Diagnostic Checks in Time Series, by Wai Kung LI, Boca Raton, FL: Chapman & Hall/CRC, 2004, ISBN 1-58488-337-5, xiii + 196 pp., \$69.95.

A researcher expecting to learn about diagnostic checks in time series will be sorely disappointed by this book on two counts, for it is much more narrowly focused than the title suggests. First, it refers only to stationary time series, thus omitting diagnostic tests for cointegrating regressions, for example. Second, it concentrates almost exclusively on portmanteau tests, which constitute only a small subset of all diagnostic tests.

Chapter One (4 pp.) is the introduction. Chapter Two (18 pp.), "Diagnostic Checks for Univariate Linear Models," covers (partial) autocorrelation tests of residuals, with extensions to seasonal autoregressive moving average (ARMA) models and periodic autoregressions. Chapter Three (19 pp.), "The Multivariate Linear Case," addresses vector ARMA models, Granger causality tests, and transfer function noise modeling. Chapter Four (17 pp.), "Robust Modeling and Diagnostic Checking," presents robust varieties of portmanteau and cross-correlation tests. Chapter Five (20 pp.), "Nonlinear Models," considers tests for nonlinear model structure; tests for linear versus specific nonlinear alternatives, goodness-of-fit tests for nonlinear time series, and choosing between different families of nonlinear models. Chapter Six (31 pp.), "Conditional Heteroscedasticity Models," presents autoregressive conditional heteroscedasticity (ARCH) model, tests for the presence of ARCH errors, diagnostic checks for ARCH models both univariate and multivariate, and a test for causality in the variance. Chapter Seven (16 pp.), "Fractionally Differenced Process," considers exact and approximate methods for estimating the parameter of interest, gives a portmanteau test statistic and an example based on tree-ring widths. Chapter Eight (15 pp.), "Miscellaneous Models and Topics," discusses ARMA models with nonnormal errors and other nonnormal time series, the autoregressive conditional duration model, and a power transformation to induce normality.

The text contains several typos, which suggests that the equations are similarly marred by inadequate proofreading. (I did not check the equations for errors.) A researcher who desires a resource for portmanteau tests applied to stationary time series will find the book useful.

B. D. McCULLOUGH
Drexel University

Ranked Set Sampling: Theory and Applications, by Zehua CHEN, Zhidong BAI, and Bimal K. SINHA, New York: Springer-Verlag, 2004, ISBN 0-387-40263-2, xii + 224 pp., \$59.95.

This comprehensive and up-to-date monograph covers, systematically and in simple language, the theory and applications of ranked set sampling (RSS), an improved technique related to traditional simple random sampling (SRS). Strong emphasis is placed on theoretical developments in RSS. In the meanwhile, the practical orientation and broad coverage will appeal to researchers and scientists working in sampling techniques, experimental designs, nonparametric statistics, and related fields.

The book comprises eight chapters of varying length. The inclusion of sections at the end of the main chapters to collect more technical arguments and to give bibliographic notes works well and helps the readers explore in more depth aspects of RSS in which they are interested.

Chapter 1 introduces the notion and general procedure of RSS. This very useful chapter will enable readers to quickly enter into the realm of RSS, learn about its historical developments, and identify applications of particular interest.

Chapters 2 and 3 discuss balanced RSS. In particular, Chapter 2 focuses on nonparametric RSS, in which no assumption on the underlying distribution of the variable of interest is made. This chapter studies in detail the relative efficiency of RSS with respect to SRS in the estimation of a population mean, a smooth function of means, and population quantiles. The authors also consider the inference procedures, such as the construction of confidence intervals and hypothesis testing. To facilitate the inference procedures based on RSS sample quantiles, they also discuss the kernel method of density estimation. This section is quite interesting. The chapter also presents some robust procedure based on M -estimates with RSS data.

Chapter 3 addresses parametric RSS, where the underlying distribution of the variable of interest is assumed to belong to some parametric family (e.g., location-scale family and shape-scale family) of distributions. The authors nicely lay out the theoretical foundation for the parametric RSS via Fisher information. The maximum likelihood estimate (MLE) based on RSS and its relative efficiency with respect to MLE based on SRS are studied, and the best linear unbiased estimate for location family of distributions is dealt with.

Chapter 4 studies unbalanced RSS. This chapter first develops the methodology of analyzing RSS data for the inferences on distribution functions and quantiles, as well as general statistical functionals. The optimal designs for the parametric location-scale family and for nonparametric estimation of quantiles are discussed in detail. This chapter also contains methods of Bayes design and adaptive design.

Chapter 5 explores classical distribution-free tests in the context of RSS. The authors consider the sign test, signed rank test, and Mann–Whitney–Wilcoxon tests and revisit the issue of the optimal design for distribution-free tests. Readers with a prior knowledge of nonparametric tests at the level of Gibbons and Chakraborti (2003) will find this chapter informative and easy to understand. For readers not familiar with these standard topics, some brief additional explanation and references might have been beneficial for the wider accessibility.

Chapter 6 describes RSS with concomitant variables. A multilayer RSS scheme and an adaptive RSS scheme using multiple concomitant variables are developed; the general regression analysis using RSS is discussed; and the design of optimal RSS schemes for regression analysis, on the basis of the concomitant variables, is explored.

Chapter 7 illustrates RSS as a data reduction tool for data mining, whereas Chapter 8 exemplifies the practical features of RSS via case studies. The inclusion of this last chapter on case studies with RSS further enhances the value of this monograph for practitioners and applied statisticians.

In the development of RSS, the choices of ranked set size k and cycle number m are directly pertinent to practical problems. This book would have been more useful had some more detailed discussion on the choices been added. However, overall I would highly recommend this well-written and reasonably priced book to researchers and practitioners, all of whom are likely to use one or more of the methods it discusses.

Chunming ZHANG
University of Wisconsin

REFERENCE

Gibbons, J. D., and Chakraborti, S. (2003). *Nonparametric Statistical Inference* (4th ed.). New York: Marcel Dekker.

Highly Structured Stochastic Systems, edited by Peter J. GREEN, Nils Lid HJORT, and Sylvia RICHARDSON. New York: Oxford University Press, 2003. ISBN 0-19-851055-1, xxii + 509 pp., \$69.50.

This book is a compilation of 10 years study regarding the development of highly structured stochastic systems (HSSS). This topic has undergone many

major developments during this period, and the contributors in this text are many of the leading developers in this area. The text points out that most of the origins of this field were developed in many diverse fields, and it was the editors's intent to develop and enhance understanding by building a common structure. As such, the ESF initiated and funded a program to develop a cross-disciplinary collaboration. This text is a summary of this work and the multiple workshops held over the past 10 years.

The topic of HSSS itself is beginning to emerge with significant practical implications due to the innovations discussed in this text. The book is written by multiple content experts and contains further discussions using addendums written by experts in the respective topics. This allows the reader to delve into other developments on a specific topic with instant access to the major theme. It was apparent that the editors spent a lot of time ensuring that these multiple discussions were complementary to the work being presented.

As a whole, the book is well laid out and the topics are discussed with great detail without being cumbersome. Moreover, the reading is continuous, and great care was taken to keep major themes consistent—an extremely difficult task given the number of contributors to the text. With this in mind, I now provide a detailed discussion of the book.

The first sections of the book are devoted to the introduction of HSSS models and their methodological foundations. The authors assume a basic understanding of the definitions and terms used in these models, but the topics are covered such that an interested reader can pick up the book without being overwhelmed. Although a more formal treatment of DAG models in statistics was given in *Probabilistic Networks and Expert Systems* (Cowell, Dawid, Lauritzen, and Spiegelhalter 1999) two of the authors of that book are contributors to this text. The first chapters discuss graphical models and their interplay with causal inference, specifically DAG models, Bayesian networks, chain graphs and graphical time series models are discussed with applications from areas such as DNA analysis and social science.

The second part of the book deals with algorithms for calculating HSSS models. The major developments regarding Markov chain Monte Carlo (MCMC) theory are presented, and multiple MCMC methodologies are discussed along with their relevance to solving particular problems. These chapters provide a comprehensive discussion of MCMC techniques from theoretical understanding to practical implementation discussions. Although the discussions are well developed, a casual reader can grasp and build from these discussions.

The third set of chapters deals with HSSS applications in epidemiologic models, spatial point processes, spatiotemporal processes, ecologic applications, binary analysis, image analysis, and genetic modeling. The final section discusses HSSS model criticism and Bayesian nonparametric smoothing, model criticism being the HSSS approach to model selection using such methods as sensitivity studies and predictive capability analysis. The nonparametric chapters cover standard material from the Bayesian perspective.

Overall, *Highly Structured Stochastic Systems* provides an extensive summary of the work done over the past 10 years on HSSS models, with numerous examples to demonstrate their strength. In the coming years, the use of HSSS models will become widespread with the rise of data information and the need to specify complexity in applicable terms. This book will be an oft-cited reference when these issues become necessary to analyzing complicated relational data.

Nicholas ROSE
Frequency Marketing, Inc.

REFERENCE

Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J. (1999). *Probabilistic Networks and Expert Systems*. New York: Springer-Verlag.

Bayesian Artificial Intelligence, by Kevin B. KORB and Ann E. NICHOLSON. Boca Raton, FL: Chapman & Hall/CRC, 2004. ISBN 1-58488-387-1, xxviii + 364 pp., \$79.95.

Bayesian networks (BNs) are located in the confluence of statistics and computer science. They are increasingly being applied in medicine, biology,

TECHNOMETRICS, FEBRUARY 2005, VOL. 47, NO. 1