Journal of Official Statistics
Vol. 7. No. 3, 1991, pp. 377–383

Book Reviews

Books for review are to be sent to the Book Review Editor Gösta Forsman, Department of Mathematics, University of Linköping, S-581 83 Linköping, Sweden.

Casella, G. and Berger, R.L., Statistical Inference. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, California, 1990. ISBN 0-534-11958-1. xviii + 650 pp., (paperback) £19.95.

According to the preface the purpose of this book is to build theoretical statistics from the basic principles of probability theory. The book is intended for first-year graduate students majoring in statistics or in a field where a statistics concentration is desirable. The prerequisite is one year of calculus.

There are twelve chapters and two appendices as outlined below:

1. Probability Theory
2. Transformations and Expectations
3. Common Families of Distributions
4. Multiple Random Variables
5. Properties of a Random Sample
6. Principles of Data Reduction
7. Point Estimation
8. Hypothesis Testing
9. Interval Estimation
10. Decision Theory
11. The Analysis of Variance
12. Linear Regression

There are tables for normal, t-, chi-square- and F-distributions and a catalog of common distributions.

There are many exercises following each chapter, a total of 570 exercises, but no answers or solutions are provided. An author index and a subject index are also included. There are nine pages of references.

The authors suggest that two distinct types of courses can be taught from this book. The "more mathematical" course is for students majoring in statistics and requires a solid mathematics background; Chapters 1–9 are recommended in their entirity together with selected topics from Chapters 10–12. The "more practical" course is aimed at students not majoring in statistics, and all twelve chapters are recommended with the exclusion of the more theoretical and mathematical sections.

As is clear from the chapter headings above, the book deals not only with inference but also with probability. The first four chapters cover the basics of probability theory while Chapter 5 mixes probability and statistics.

Chapter 1 introduces probability theory up to the level of defining a Borel field. In Chapter 2 the moment generating function is used while the characteristic function is

mentioned under the heading "Miscellanea" at the end of the chapter together with other generating functions.

In the probability chapters some of the mathematical content is presented in separate sections. Chapter 2 has a section about differentiating under an integral sign and Chapter 4 has a section about inequalities and identities.

Besides elementary probability theory the first four chapters also include more advanced sections about exponential families, location and scale families in Chapter 3 and a section about hierarchical models and mixture distributions in Chapter 4.

Chapter 5 is a mixture of probability and statistics and can be the starting point for a course in statistical theory for students with some probability background. Random samples, sums of random variables from a random sample, convergence concepts, sampling from the normal distribution and order statistics are dealt with in this chapter.

In Chapter 6 three principles of data reduction are treated. These are the sufficiency principle, the likelihood principle, and the invariance principle.

Chapters 7–9, representing the central core of statistical inference, estimation, and hypothesis testing, are divided into methods of finding appropriate statistical techniques and methods of evaluating these techniques. There is also a section in each of these chapters titled "Other Considerations" where the authors indicate how the rules of statistical inference may be relaxed and still produce meaningful inferences.

Chapter 7, which deals with point estimation, presents the methods of moments, maximum likelihood estimators, Bayes estimators and invariant estimators. In the "Other Considerations" section of this chapter the asymptotic variance of maximum likelihood estimators and Taylor series approximations are treated.

In Chapter 8 likelihood ratio tests, invariant tests, Bayesian tests and the union-intersection method of test construction are discussed. The "Other Considerations" section in this chapter presents the asymptotic distribution of likelihood ratio tests.

Chapter 10 deals with estimation and hypothesis testing from a decision theoretic view. Bayes rules, admissible decision rules and minimax rules are discussed. There is also a presentation of Stein's paradox.

In Chapter 11 one-way analysis of variance and randomized complete block designs are treated. A discussion of simultaneous estimation of contrasts is included.

Chapter 12, about linear regression, is restricted to simple linear regression. On the other hand a discussion of regression with errors in variables is included in this chapter.

I read the book with great pleasure. It is clearly and simply written. The many examples facilitate the introduction of new concepts. The book certainly attains its goal of giving a student with a solid mathematics background knowledge of theoretical statistics. I am, however, uncertain of whether the book is also suitable for a more "practical" course. For instance, in Chapter 12 about regression analysis it seems more natural to continue the section about simple regression with a section about multiple regression instead of treating the problem of errors in both variables. Non-parametric methods are also missing.

As a teacher for students of applied statistics I would have liked more material about estimation and hypothesis testing in the multiparameter case and also for dependant observations, perhaps also something about numerical optimization for non-linear estimation problems. It is somewhat frustrating to realise that even after studying this solid book you do not have the necessary theoretical background to understand how to estimate an ARIMA model! I was surprised to see that the Wald and the Lagrange multiplier tests were not mentioned in connection with the likelihood ratio test since these tests are nowadays presented in elementary econometrics textbooks.

It would also have been nice if the book had included something about bootstrap and jackknife analysis. Here the jackknife is mentioned in an exercise.

Mats Hagnell
University of Lund
Lund
Sweden

Dudewicz, E. J., Chen, P., and Taneja, B. K., Modern Elementary Probability and Statistics, with Statistical Programming in SAS, MINITAB, & BMDP, American Sciences Press, Inc., Syracuse, NY, 1989. ISBN 0-935950-19-2. ix + 375pp., $34.50.

According to the authors, this book has been developed for introductory courses in elementary probability, elementary statistics, and elementary probability and statistics. The only prerequisite is high school algebra. The book consists of three parts with 18 chapters in all. The contents are:

Part I

1. Randomness, Descriptive Statistics
2. Events, Probability, Conditional Probability
3. Bayes' Theorem
4. Combinatorial Probability
5. Random Variables, Probability Distributions, Means
6. Binomial Random Variables
7. Normal Approximation, Central Limit Theorem

Part II

8. Point and Interval Estimation
9. Hypothesis Testing
10. Linear Regression and Correlation
11. Goodness-of-Fit Chi-Square Test
12. Independence in *rxc* Contingency Tables
13. Analysis of Variance (ANOVA) and Design
14. Medians and Nonparametric Statistics
15. Statistical Selection and Decision

Part III

16–18 Statistical Programming in SAS, MINITAB, and BMDP

Each chapter is divided into sections which all end with a number of problems. Selected answers to these problems can be found at the end of the book. The examples are "realistic," easy to understand, and are taken from a large variety of areas. I believe that these examples will serve as a source of inspiration to learn more about statistics. They will probably also convince the reader of the usefulness of statistical methods. Theorems are given without proofs but are often made plausible by intuitive arguments. Unfortunately there are no references to textbooks for those readers who want to go deeper into the theory.

The book contains two areas which are not so common in elementary textbooks, i.e., statistical selection and decision and statistical programming. In Part III short introductions to the program packages SAS, MINITAB, and BMDP are given and examples of how to use them for descriptive statistics, correlation, regression, and analysis of variance. The program codes used to produce the analysis as well as the outputs are presented. Detailed comments to each output are given. In spite of the small number of pages devoted to this part, the reader gets an idea of how the packages work.

I find this book very suitable for introducing statistics to students not having more mathematics than basic algebra. Even for readers with greater mathematical knowledge the book can be recommended as an introduction to probability and statistics.

Staffan Uvell
University of Umeå
Umeå
Sweden

Marriott, F. H. C., A Dictionary of Statistical Terms, 5th Edition. Longman Group UK Limited, Harlow, Essex, 1990. ISBN 0-582-01905-2. viii + 223 pp., £25.

F. H. C. Marriott, in the preface to the fifth edition, states that he attempted ". . . to

maintain the standards set by the original authors." He certainly was successful in this attempt. The form of the dictionary has been maintained and with few exceptions terms have not been deleted. The notable exception to the deletion process is, as pointed out in the preface, the omission of a number of Italian terms, largely deemed to be both ephemeral and obsolete. The current edition continues the practice started in the fourth edition of giving, when possible, the author and publication date where the term was first used.

Where a term is used in different disciplines, the meaning of the term in each discipline is given, e.g., a-error is defined in terms of hypothesis testing and also as a producer's risk in quality control. In some places, a nonrecommended usage is given, but the reader is warned; e.g., area sampling which usually refers to a method of sampling used when no complete frame is available, is sometimes used as meaning the sampling of a domain to determine area (acreage) of a given crop. The latter usage is not recommended. Obsolete terms are identified, as is the case with "external variance," which appeared in 1930 to denote the variance of predictions of the future movements of a particular time series based upon a given form for its trend. Presently, this term is used in quality control to denote variance between primary units in a two-stage sampling process. Some terms are identified as ones that should be avoided, e.g., "harmonic distribution" is identified as a term to be avoided because it could be confused with the distribution of harmonics in spectral analysis.

I found the dictionary easy to read and use. Critical words or phrases are presented in boldface type and the cross-referencing seems adequate. This dictionary will be invaluable to consulting statisticians working in multidisciplinary environments. Students will also find this dictionary a help to their studies.

G. Morris Southward
New Mexico State University
Las Cruces, NM
U.S.A.

Rao, C.R. and Rao, M.M., (eds.), Multivariate Statistics and Probability: Essays in Memory of Paruchuri R. Krishnaiah. Academic Press, Inc., New York, 1989. ISBN 0-12-580205-6. xviii + 567 pp., $59.95.

This book, in memory of Professor P.R. Krishnaiah is a collection of 35 essays, which were originally to be presented on Krishnaiah's 55th birthday. However, due to his untimely death, the book had to be dedicated to honor his memory. There is a short preface by the Editors C.R. Rao and M.M. Rao, followed by a short biography of Krishnaiah by M.M. Rao and a list of publications of Krishnaiah. To get an idea about the types of essays and the names of contributors, a few selected titles are given below:

"Kernal Estimators of Density Function of Directional Data" by Z.D. Bai, C.R. Rao, and L.C. Zhao; "Ergodicity and Central Limit Theorems for a Class of Markov Processes" by Rabi N. Bhattacharya and Oessok Lee; "Conditionally Ordered Distributions" by Henry W. Block and Allan R. Sampson; "Minimal Complete Classes of Invariant Tests for Equality of Normal Covariance Matrices and Sphericity" by Arthur Cohen and John I. Marden; "On the Area of the Circles Covered by a Random Walk" by P. Erdös and P. Révész; "Empirical and Hierarchical Bayes Competitors of Preliminary Test Estimators in Two Sample Problems" by Malay Ghosh and Bimal K. Sinha; "Smoothness Properties of the Conditional Expectation in Finitely Additive White Noise Filtering" by H.P. Hucke, G. Kallianpur, and R.L. Karandikar; "Isotonic M-Estimation of Location: Union-Intersection Principle and Preliminary Test Versions" by Azza R. Karmous and Pranab K. Sen; "Stochastic Integrals of Empirical-Type Processes with Application to Censored Regression" by Tze Leung Lai and Zhiliang Ying; "Paradoxes in Conditional Probability" by M. M. Rao; "Asymptotic Expansions of the Distributions of Some Test Statistics for Gaussian ARMA Processes" by Masanobu Taniguchi; and

"Estimating Multiple Rater Agreement for Rare Diagnosis" by Joseph S. Verducci, Michael E. Mack, and Morris H. DeGroot.

Unfortunately, unlike the *Proceedings of Multivariate Analysis Symposia* published by Krishnaiah, these papers are not grouped according to statistical area. They range from a generalization of the Cauchy-Binet formula to the estimation of bispectral density functions. Also, it does not appear that the choice of the contributors was limited to close friends and co-workers of Krishnaiah, which would have been better. If the choice of topics would have been restricted to multivariate analysis and areas closely related to it, the book would have acquired a more cohesive look. I am sure, the Editors could have found some recent, to-be-published material by Krishnaiah himself in his last days and its publication would have proved very appropriate and fitting. One of the contributors in this book is Dr. C.G. Khatri who was as prolific a researcher as Krishnaiah. He also died at a very young age and this is mentioned in a footnote on page 319. However, a little more on the collaboration between him and Krishnaiah in the area of multivariate analysis and the editing of the *Journal of Multivariate Analysis*, would have made the world of statisticians aware of its loss due to the deaths of these two stalwarts.

There is a wealth of material in this volume. But most of the material is highly mathematical, involved, and theoretical. The chapters include not only areas of multivariate analysis but also aspects of probability and stochastic analysis, finite sampling and asymptotic results, decision theory, Bayesian analysis, time series, Markov processes, central limit theory, and even foundational problems. The only papers that are strictly in the statistical inference area are: "Analysis of Odds Ratios" by Subramanyam and Bhaskara Rao; "Estimating Multiple Rater Agreement for Rare Diagnosis" by Verducci, Mack, and DeGroot; "Interference Properties of a One-Parameter Curved Exponential Family" by Ruiz-Rivas; and "Cuadras and An Improved Estimation Method for Univariate Autoregressive Models" by Pukkila. These papers are probably the only ones which have numerical data and illustrative examples. There is a preponderence of papers that deal with asymptotic expansions. The volume is therefore of little use to an applied statistician, but it is not intended to be so. Even for a highly trained mathematical statistician, the reading will be tough on occasions. The preface mentions discussions of practical applications and computational solutions but one has to dig deep to find them. All papers start tersely with what they want to achieve and none mentions any warm reference to the authors' association with or memories of Dr. Krishnaiah. The average length of papers is about 15 pages. The shortest one is 6 pages and the longest one is 27 pages. The quality of publishing is excellent – as is to be expected from the Academic Press.

Anant M. Kshirsagar
University of Michigan
Ann Arbor, MI
U.S.A.

Tufte, E.R., Envisioning Information. Graphics Press, Cheshire, Connecticut, 1990. ISBN 0-961-3921-1-8. 121 pp., $46.

Envisioning Information is an exceptionally beautiful book. The overall structure and layout of the book superbly illustrate the authors' subject matter. In this book, Tufte expands on the material presented in his prior work, *The Visual Display of Quantitative Information* (Tufte 1983).

The book is divided into six chapters. The first chapter, Escaping Flatland, describes his goal which is to illustrate effective ways of conveying complex information on a two-dimensional surface – the flatland. His basic premise is that visual displays of information should be designed to enhance visual reasoning by finding ways to display complex data so that the focus is on the data and not on the design. He decries the use of what he calls "chartjuck" and "posteriz-

ation" of information that result in visual displays that are cluttered with distractions and that contain little information – that are "data-thin". He feels that it is best to present small amounts of data in tables, and that visual displays are best suited for presenting information that is complex and that the "history of information displays and statistical graphics is entirely a progress of methods for enhancing density, complexity, dimensionality, and even sometimes beauty."

In the remaining five chapters of the book, Tufte explains and illustrates techniques that can be used to construct "information-thick" visual displays. In chapter 2, Micro/Macro Readings, he illustrates how the appropriate use of detail can clarify information. A visual display that embodies this principle allows the viewer, to first obtain an overall understanding of the data conveyed by processing the information as a whole and, second, to then enhance his understanding by narrowing his focus on to the details of the information selecting those parts that are of interest to him. In fact, Tufte's entire book follows this principle. It is possible to read the whole book in a few hours and glean the general principles he posits; whereas, careful reading and review of specific chapters and sections yields more information on how to implement these principles.

In the third chapter, Layering and Separation, Tufte discusses procedures for highlighting the different types of information in a complex visual display without creating surplus visual images that interfere with the viewer's ability to understand the structure and patterns in the data. He illustrates how color, shading, type style, intensity, and so on can either interfere with or enhance the display of information. These techniques are particularly important for the design of small multiple visual displays that enhance the viewer's ability to perceive change and differences in data. Visual displays that make use of small multiples (chapter 4) are composed of multiple similar designs that vary according to some other variable that is not included in each individual element of the display. Tufte notes that comparisons are facilitated by placing the entire display within the scope of a single eyespan.

Tufte devotes the last two chapters of the book to a discussion of color and methods for conveying information on changes over time in three-dimensional space. Four fundamental uses of color in information design – to label, to measure, to represent reality, and to decorate – are discussed. In addition, he illustrates how the poor use of color can clutter a chart so that the data are obscured. In the space and time chapter, he focuses on time tables, route maps, and dance notation. The ideas he presents are useful for displaying statistical information that varies over space and time.

I think that this book is useful for those who are concerned with presenting the results of applied statistical studies. However, one needs to read both his earlier work (Tufte 1983) and this book to fully understand the material. In contrast to the earlier work, I found that it was harder to glean the lessons from the text. In *The Visual Display of Quantitative Information*, the author summarizes the principles of each chapter for the reader; however, he does not use this helpful technique in *Envisioning Information* so that it is more difficult to extract the points that are directly useful for presenting statistical information.

The cost and time associated with creating the information rich visual displays that the author recommends could be high. However, as the use of personal computers and desktop publishing software becomes more widely available, the methods that he explicates will become more accessible for agencies and firms engaged in producing statistical reports. These books should definitely be read by those who are involved in producing statistical graphics software, and the software needs to be designed so that the user has the flexibility to produce visual displays which follow his principles.

In summary, the ideas that Tufte discusses are very important for applied statisticians. Our work is not effective unless we can clearly convey our findings to the users. Thus, I recommend both this book and the earlier one to those who are engaged in preparing statistical reports. In addition,

these ideas should be considered in statistical training programs.

References

Tufte, E.R. (1983). The Visual Display of Quantitative Information. Cheshire, Connecticut: Graphics Press.

Judith T. Lessler
Research Triangle Institute
Research Triangle Park, NC
U.S.A.