

Book Reviews

Books for review are to be sent to the Book Review Editor Jaki S. McCarthy, USDA/NASS, Research and Development Division, Room 305, 3251 Old Lee Highway, Fairfax, VA 22030, U.S.A.
Email: jaki_mccarthy@nass.usda.gov

Handbook of Statistics 29: Volume 29B, Sample Surveys: Inference and Analysis <i>Robert Baskin</i>	577
A Quantitative Tour of the Social Sciences <i>Matt Jans</i>	578
Picturing the Uncertain World: How to Understand, Communicate, and Control Uncertainty through Graphical Display <i>Ana F. Militino</i>	580
Bayesian Analysis for the Social Sciences <i>Yves Thibaudeau</i>	581

Danny Pfeffermann and C.R. Rao (eds). *Handbook of Statistics 29: Volume 29B, Sample Surveys: Inference and Analysis*. New York: Elsevier, 2009. ISBN 978-0-444-53438-5, 666 pp, \$265USD.

My first impression of this book is that it will be extremely useful to me for the purposes of doing statistical consulting with a large number of end users of health conditions, expenditure and insurance data from a family of government-sponsored complex surveys. Often these users are interested in somewhat advanced modeling and inferences that may push the limits of design-based analysis. Part B of this comprehensive volume provides up-to-date information on almost all approaches to building not purely design-based models or doing some form of inference based on data from complex survey designs. There are a total of nineteen chapters as well as introductions to each of the three parts of the book that provide a wide array of technical material relating to inference on survey data.

There are six parts in the total volume with the first three parts in Volume A and parts four, five, and six in Volume B. Part four contains an introduction and eight chapters on different approaches to inference from survey data. Part five is on special estimation and inference problems and has seven chapters together with an introduction. Part six is on informative sampling and theoretical aspects with the information given in an introduction and four chapters. These parts and chapters are written independently and can be read in almost any order. This leads to more than one way of using the volume. If a reader has a specific problem at hand such as analyzing a survey repeated over time or doing a population-based case-control study, then that reader can go directly to the relevant chapters in the book and get up to date information as well as the current references. If a reader wants to see if there are new techniques then that reader might skim parts five or six. On the other hand, a reader may be interested in getting a broad introduction to different approaches to inference based on survey data as mentioned by the editors in the preface. If the reader goes to part four on different approaches they will find eight chapters

on inference. At this point I am reminded of the old saying, “Don’t go to sea with two compasses, if they don’t agree you will be lost.” I wonder if a reader using this volume to gain insight on inference from surveys would not be going to sea with more than one compass. Part four does give an array of approaches to analyzing survey data and it does so in an even-handed way that avoids the controversies of the past. However, I would recommend that a person using this volume to gain insight have access to some well-informed advisor who has already heard the years of debate over these approaches. For the record, the chapters in part four are as follows. The first chapter is on model-based prediction of finite population totals, the second is a discussion of design-based and model-based inference of model parameters, the third is on calibration weighting, the fourth on estimating equations, the fifth on nonparametric and semiparametric estimation, the sixth on resampling methods, the seventh on Bayesian methods, and the eighth on empirical likelihood methods.

Part five is about special estimation and inference problems. The first two chapters are on small area estimation. There are then two chapters on analysis of surveys repeated over time and analysis of longitudinal surveys. The last three chapters of this section are on categorical data analysis, inference on distribution functions and quantiles, and scatterplots with survey data. The final part contains four chapters on case control studies, inference under informative sampling, asymptotic in finite population sampling, and decision-theoretic aspects of finite population sampling.

Be aware that there is almost no mention of software in this volume. Perhaps the editors should consider a volume C: “Here’s How You Do This in SAS, STATA, or R.”

Overall, this is an excellent and useable reference on analysis and inference in sample surveys. It is especially useful to consulting statisticians who must deal with complex survey data, it can be useful to analysts who must deal with complex survey data, and I hope, as the editors mentioned in the preview, that it can be useful to graduate students to learn about the area of complex survey data.

Robert Baskin

Email: rmbaskin2@comcast.net

Andrew Gelman and Jeronimo Cortina (eds). *A Quantitative Tour of the Social Sciences*. New York: Cambridge University Press, 2009. ISBN 978-0-521-86198-4 (hardback). \$94.99USD.

Believe it or not, there are still areas in the statistical world where social science is equated with qualitative research. As a quantitative social science researcher, I periodically have to explain to my more statistical colleagues that even though I study the psychology of survey response, I actually use numbers and statistical tests. Likewise, there are still social scientists whose statistical training has left large gaps in their quantitative ability. Pair these observations with the reality that statisticians are increasingly expected to apply their skills to social problems, and social scientists are expected to have statistical skill far beyond what was available to the most experienced statisticians a generation ago. *A Quantitative Tour of the Social Sciences* may not solve the residual divide between social and statistical sciences, but certainly makes a large step forward, making the contemporary interdisciplinary scientific milieu explicit.

The 350-page (including references and index) hard-cover book (also available in paperback) is divided into seven parts with three or four chapters each. Chapters average about fourteen pages, making each a fairly quick read. The book comes out of a graduate course taught as part of Columbia University's Quantitative Methods in Social Sciences Program. The chapters are edited transcripts of lectures given by each of the instructors. Those who know the authors personally or have heard them lecture might be able to imagine their voices lifting from the pages. As edited transcripts, the chapters retain a much less formal tone than one usually encounters in text books, which was welcome to this reader. I found that this more conversational style made the book quite easy to read and very accessible. Exercises are given at the end of each chapter, so the book could easily be used as a text for a course similar to the one generating it.

The social sciences covered include economics, sociology, political science, and psychology, with overview chapters on methods and models within the social sciences, the history of quantitative social science, and a more statistical chapter on causal inference. The breadth of topics covered certainly fulfills the title of the book.

The chapters I found most enjoyable were those covering the social and quantitative topics I knew the least about going in (e.g., political science, propensity score matching). I came away from this book knowing more about classical problems, research, and theory in fields that I've never studied before, and new insights into areas I have studied. I particularly enjoyed the chapters on causal inference, which in my mind is a topic that is at the core of interactions between social and statistical sciences. Social scientists want to know why people and groups behave the way they do, and statistical scientists can provide the tools to explore the causal relationships underlying these "why" questions. A chapter in this unit titled "Some Statistical Tools for Causal Inference with Observational Data" (Chapter 22) explains propensity score matching in a very clear way using an example based on real data. After reading this chapter I have a grounding in this technique that I did not have before, and enough knowledge about it to at least think about how I can apply it to my own work. This is the sign of a good book to me.

My reservations are mostly about the chapters covering psychology, which happens to be the social science field I know the best. I suspect most readers with deep knowledge of any one of the fields covered in this book might feel the same way about their own field. One weakness I felt in the psychology chapters concerns the number of citations. The rate of citation varies widely across all chapters in the book, but it seemed relatively low in the psychology chapters. The authors recommend further reading at the end of each chapter, which helps make up for the scarcity of in-text citations. In terms of content, the chapter on cognitive and social psychology (Chapter 19) orients itself around what I know to be classical theory and research in social cognitive psychology. However, it also spends a lot of time discussing psychodynamic (e.g., Freudian) theory, which from my understanding does not play a large role in modern psychology. Similarly the chapters covering decision science do not cover some of the theories and research I have come across in my own literature reviews on the topic (e.g., fuzzy trace theory, theory of reasoned action and planned behavior, and research linking affect and decision making). There was also no coverage of biological psychology, neuropsychology, or perception research. Keeping in mind that these chapters are taken from lecture notes, and that the authors are two highly distinguished and accomplished researchers (and one graduate student, now faculty)

my comments are more about the scope of the psychology section, and reflect some daydreaming about “how I would teach the topic”, but I do not question the accuracy or broader relevance of any of the content per se. It is natural to be most critical about the topics we know the best, and I think the chapters would be very informative and enjoyable for someone new to psychology.

The largest missing piece for me was any coverage of psychometrics, including factor analysis and latent variable modeling (e.g., structural equation modeling). Factor analysis is such a cornerstone of quantitative psychological research and theory that it seems like a major oversight to not include this topic in the tour.

Overall I think this book would be helpful for anyone who works at the intersection between social and statistical sciences (e.g., education, public health), or teaches courses preparing students for such work. The book is accessible enough to be used in an advanced undergraduate course, and is ideal for a graduate-level course. Instructors would likely want to supplement the text with their own perspective and expertise (or that of guest lecturers) to liven and extend the thought-provoking discussions that begin in the book. This book is accessible enough that it could even be used as part of a discussion group or book club in an academic or professional setting. The chapters are short, and the conversational format leaves many open questions to discuss. In summary, this book helps bring integration to an often fast-paced and intellectually demanding research culture. I commend the authors and editors for turning the ephemeral experience of such a fascinating course into material that can be shared more broadly.

Matt Jans

*U.S. Census Bureau
4600 Silver Hill Road
Washington, DC 20233*

U.S.A.

Phone: 301-763-6724

Email: matthew.e.jans@census.gov

H. Wainer. *Picturing the Uncertain World: How to Understand, Communicate, and Control Uncertainty through Graphical Display.* Princeton, U.S.A: Princeton University Press, 2009. ISBN 978-0-691-13759-9, xv +244 pp, \$29.95.

The main goal of this book is to explain basic statistical concepts in an enjoyable way. It is the result of the large professional experience of the author who wants to illustrate how, when, and why statistical concepts can help to interpret uncertainty. This aim is achieved within the book. The focus is on the graphics, but some other tools are also considered.

The book has twenty-one chapters classified into five parts: Introduction and Overview, Political Issues, Educational Testing, Mostly Methodological, and History. The Introduction is mainly devoted to the De Moivre’s description of the expected variation of the arithmetic mean. This is the chapter I liked the most, because the sampling distribution of the mean is one of the more complicated concepts to be explained to beginners. This is why the author provides curious examples and counterexamples to clarify this concept. In Political Issues, the author shows different examples that have appeared in newspapers to explain the interpretation of graphs. The examples, given in five chapters, are

related to politics, medicine, and education. Specifically, the explanation of the relationship between student's test scores and money spent on education is very appealing. Educational Testing has three chapters. The first one is about a problem close to disabled people and how to use statistics to navigate between the standards of testing under nonstandard conditions and the court decisions that affect them. The second is about the conclusions regarding differences in SAT test scores between ethnic groups. The author looks into the validity of the conclusions. The third is also about the SAT and discusses the implications of misunderstanding the concept of statistical error. Mostly Methodological is the more technical part. It has five chapters. The first is about the validity of linear extrapolation for prediction purposes. The second demonstrates how a mixture of statistical tools, statistical thinking, and various graphic forms combine to provide us with knowledge and information. The last two chapters show uncertainty, graphically using examples from the National Assessment of Educational Progress and the Mendel Effect. Part five is about History and is split into seven chapters. Here, the author uses modern tools to investigate ancient puzzles in two chapters, explores the origins of some modern tools, defends the wisdom of the ancients against contemporary misuses, and finally by a heart-rending use of graphics illustrates the World War II tragedy of the Nazis' Kovno Ghetto.

Above all the book is very well written. It is easy to read, and the presentation is clear and nice. The examples come from the real world and reference to historical information is widely made. Graphical display with regard to the presented examples is given, but not all the graphics are of general interest. So a more general focus in this area would be welcomed. All in all, this book can be a useful tool for practitioners of statistics and for teachers, who can find good material, examples, and exercises for their students.

Ana F. Militino
Universidad Pública de Navarra
Campus de Arrosadia
31006 Pamplona
Spain
Email: militino@unavarra.es

Simon Jackman. *Bayesian Analysis for the Social Sciences.* West Sussex, UK: Wiley and Sons, 2009. ISBN 978-0-470-01154-6. 564 pp, \$90USD.

What I think are the two most important features of this book are: First the impressive breadth of its contents, second the judicious level of mathematical skills that is needed to seriously study the book.

I do not recall reading a book with such a comprehensive presentation of Bayesian statistical topics since *Probabilities and Statistics* (DeGroot, 1986; 1975), which was published long before the advent of computationally intensive simulation methods. The author chronicles advances in Bayesian statistics from its very beginnings, in the writings of Laplace in eighteenth-century France to modern simulations on high-power computers in the best academic institutions of our times. At the same time, the information in the book is packaged in a sensible order so that a lecturer will need to spend only minimal efforts retrieving material from the book to prepare his/her lectures.

The second important feature of the book is its level of technical difficulty. This level is appropriate, as it is advanced enough to enable the author to introduce and discuss most pertinent topics in Bayesian statistics. But it is not so advanced as to limit the number of prospective readers who stand to benefit from reading the book. As such, much of the book is within the reach of undergraduate students who major in mathematics, as early as their sophomore year.

Now, the reference to “Social Sciences” in the title of the book needs to be put in perspective here. The social scientists best capable of digesting this book are those with advanced calculus and formal algebraic training. The book will not be a good fit for students working on a social sciences curriculum strictly in the context of a liberal arts education. But, given they possess the basic mathematical skills required, social scientists, and students and professors in the field, will find the examples and applications presented in the book very relevant indeed, and instructive.

The book is divided in three parts in a way that establishes logical boundaries for its use as a textbook. The first and second parts hold just about the right amount of material for a two-semester undergraduate introductory course for students in mathematics major programs, and comparably trained social scientists. The material in the third part is at a level of complexity suitable for graduate training. In that respect, the book could serve as a central reference for a one-year seminar-style course on special topics in Bayesian statistics.

In addition to enjoying the comprehensive presentation of the subject of Bayesian statistics as a whole, I found myself smiling at discussions on idiosyncrasies of Bayesian theory that most authors sidestep. I particularly enjoyed the discussion of hierarchical models as stemming from simple ANOVA schemes (Section 7.2). This is the way I learned hierarchical models 30 years ago, and it frustrates me greatly how this topic routinely gets pulled out of nowhere at levels of complexity much greater than needs be.

I enthusiastically recommend this book as a comprehensive reference which one should find in any competent statistician’s bookcase, and as a must for academicians who are serious about teaching Bayesian statistics, with or without the social sciences bent.

Reference

DeGroot, M.H. (1975) (2nd ed. 1986). *Probability and Statistics*. Addison-Wesley Ed.

Yves Thibaudeau
Mathematical Statistician and Principal Researcher
Statistical Research Division
U.S. Census Bureau
Washington, DC 20233
U.S.A.
Email: Yves.thibaudeau@census.gov