Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: http://www.tandfonline.com/loi/uasa20

Book Reviews

To cite this article: (2016) Book Reviews, Journal of the American Statistical Association, 111:516, 1840-1851, DOI: 10.1080/01621459.2016.1257826

To link to this article: http://dx.doi.org/10.1080/01621459.2016.1257826

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION
2016, VOL. 111, NO. 516, 1840–1851, Book Reviews
http://dx.doi.org/10.1080/01621459.2016.1257826

Analyzing Sensory Data with R. Sébastien Lê and Thierry Worch. Boca Raton, FL: Chapman & Hall/CRC Press, 2014, xvii + 356 pp., $89.95 (H), ISBN: 978-1-46-656572-2.

What adjectives would a panel of human subjects use to describe the flavor of a glass of wine? The answer to this question, and to others relating to the human senses, is examples of sensory data. Such data are usually complex and their analysis requires non-traditional statistical tools. Analyzing Sensory Data with R is a comprehensive and practical book written for sensory scientists. The book introduces techniques for the analysis of sensory data with a strong emphasis on multivariate analysis. Throughout the book, real data from a variety of sensory experiments are used to illustrate how to analyze sensory data appropriately using R.

The field of sensory research is rapidly developing, often requiring the development of new or modified statistical methods. The open-source nature of the R language, with its active community of contributors, has enabled it to keep pace with this methodological development, making it the tool of choice for the analysis of sensory data. The book uses the dedicated SensoMineR package developed by the authors, along with other R packages such as FactoMineR.

This book aims to provide sensory scientists with a textbook for analyzing sensory data. Although the conceptual basis for the statistical models is provided, the book is written from a very practical perspective to provide instruction on how to analyze real data. The examples used in the book emphasize a good understanding of the objective of the experiment, the nature of the data collected and its statistical notation, how to implement the appropriate method of analysis using R, and how to read the output and carefully interpret the results.

The book is organized into three parts: (1) quantitative descriptive approaches for products that are evaluated according to different sensory attributes (this part mainly covers the analysis of variance, principal component analysis, and multiple factor analysis); (2) qualitative descriptive approaches for products that are evaluated as a whole (this part covers more complex statistical methods including the Bradley–Terry and Thurstonian models, multivariate methods with correspondence analysis, and multiple correspondence analysis); and (3) affective descriptive approaches for products that are evaluated with respect to a hedonic measure. The third part describes more complicated analyses that are built upon the statistical methods introduced in the first two parts.

Within each part of the book, several independent chapters are presented. Each chapter focuses on only one main sensory topic and consists of three sections. The first, *Data, Sensory Issues, and Notation*, presents the nature of the sensory evaluation and its objective. The second, *In Practice*, describes the statistical analysis using R, the key output and interpretation. The third section, *For Experienced Users*, describes the variants or extension of the methods to the other sensory tasks or statistical methodology. Each chapter ends with many exercises and recommended readings, which makes it ideal for use as a course textbook.

The book does not presume the reader has a good knowledge of R. Thus, some details on R functions are provided when they are first introduced. Detailed comments on R commands for running the statistical analysis are also provided. An R Survival Guide in the appendix will also be helpful to R beginners.

The book has a dedicated website at *http://www.sensorywithr.org*. The website contains the datasets used for examples and exercises, as well as some additional useful information about the book and R software. Solutions to exercises using R are also posted on the website.

Sensometrics is a dynamic and fast-moving field, yet books on how to analyze complicated sensory data are rare. This practical book is a valuable resource for scientists learning how to analyze sensory data appropriately. It can be used as a course textbook to teach new sensory scientists the full suite of existing statistical techniques, or as a reference handbook for the experienced sensory scientist seeking information on how to analyze a specific type of sensory data. Each chapter is self-contained, so it is easy for readers to jump from one chapter to another.

The book has room for improvement, most readily seen in certain figures that could be improved for clarity. It would also benefit readers if there were a more thorough treatment of experimental design preceding the introduction of the tools of analysis. Nonetheless, the book is a welcome addition to the offerings in this field. Analyzing Sensory Data with R, is well-organized, easy to read, and provides practical instruction in statistical techniques for analyzing sensory data using R.

Gui-Shuang Ying
University of Pennsylvania

Bayesian Inference for Partially Identified Models: Exploring the Limits of Limited Data. Paul Gustafson. Boca Raton, FL: Chapman & Hall/CRC Press, 2015, xxi + 174 pp., $92.95 (H), ISBN: 978-1-43-986939-0.

Loosely speaking, a model is identified if different values of the parameters cannot produce the same joint probability distribution of the observables. If the parameters are identified, then as data accumulate the true values of the parameters can be learned, that is, they can be consistently estimated. Unidentified parameters cannot be consistently estimated.

In this little gem of a monograph, Paul Gustafson (p. 2) asks:

> "Given the appeal of identification and the woe associated with its absence, it is legitimate to ask: why ever bother with statistical models that lack identification? Why not just stick to the nice identified models? To some extent, the statistical community does just that. The overwhelming majority of statistical models proposed in the methodological literature and used in the wider scientific literature are indeed identified models."

Gustafson challenges this accepted wisdom and argues that partially identified models should not be so quickly dismissed. After all, whether the appeal of consistency of procedures (estimators) in arbitrarily large samples need be transferred to procedures based on samples that are decidedly not arbitrarily large is unclear. Rather than an "unidentified model," Gustafson prefers the more positive phrase "partially identified model" (PIM).

Gustafson considers an observational study comprising a suite of variables (Y, X, C), where Y is a "response" variable, X

is a "treatment" variable, and C is a collection of "covariates" or "potential confounders." Based on a random sample from this suite, interests turn to the conditional distribution $Y|X, C$. Given a parametric distribution for this suite involving an identified parameter, analysis may proceed in the usual fashion.

But Gustafson discusses three possible threats to this validity: unobserved confounders, measurement errors, and selection bias. All three amount to a corruption of the data in which the parameters of interest are no longer identified. The implications for the frequentist researcher are inconsistent estimators and a lack of valid confidence intervals (even asymptotically). The temptation is to secure identification by literally untenable exact parametric restrictions, and hence, the model is admittedly misspecified.

This situation is in stark contrast to the situation seen through Bayesian eyes. The often cited quote (p. 157) of Dennis Lindley reminds us (assuming a proper prior):

> "In passing it might be noted that unidentifiability causes no real difficulties in the Bayesian approach."

Bayes theorem yields a proper posterior distribution for all parameters. Minimizing expected posterior loss yields the usual Bayes point estimates, credibility regions, hypothesis tests, and predictions. If there is no ambiguity over the prior, then credibility regions have the correct posterior size. However, there is no Bayesian free lunch: unidentifiability implies there are some parameters for which the prior and posterior are the same (no learning occurs). Also the posterior conditional distribution of unidentified parameters given the identified parameters always equals its prior counterpart.

Chapter 2 lays out the basic statistical theory and properties surrounding Bayesian inference with a PIM. The key idea leveraged is that of reparameterization. Particularly, at least some PIMs are amenable to separating out parameters "inside" the likelihood function from those "outside." Gustafson (p. 16) calls such a reparameterization a transparent reparameterization. Those parameters inside the likelihood are directly informed by the observed data, while those outside are unidentified and at best indirectly informed by data. Gustafson (p. 18) distinguishes between three types of transparent reparameterizations: (i) *factorable* (unidentified and identified parameters are independent under the prior), (ii) *loose* (unidentified and identified parameters are *dependent* under the prior, but the *support* of the prior conditional distribution of the unidentified parameters given the identified parameters does not depend on the identified parameters), and (iii) *sticky* (unidentified and identified parameters are *dependent* under the prior, and the *support* of the prior conditional distribution of the unidentified parameters given the identified parameters does depend on the identified parameters).

Gustafson (p. 19) defines a target as a scalar invertible function of interest depending on all the parameters. For a factorable transparent reparameterization, there is *no indirect learning* about the target. For a loose transparent reparameterization, there is *weak indirect learning* about the target, but the support of the target given the identified parameters is the same as the support of the prior marginal distribution of the target. Finally, for a sticky transparent reparameterization, there is *strong indirect learning* about the target and the support of the target given the identified parameters is a subset of the support of the prior marginal distribution of the target. In this last case, asymptotically, the support of the target corresponds to the identification region in non-Bayesian PIMs.

Chapter 2 also describes how reparameterization can assist in computing posterior quantities, and how reparameterization gives insight into properties of Bayesian estimators. The marginal posterior distribution of the unidentified parameters may or may not equal the marginal prior distribution, that is, the "heresy" of learning about unidentified parameters. If the prior exhibits dependency between identified and unidentified parameters, then as the researcher learns about the identified parameters, this leads to revision of beliefs about the unidentified parameters.

Chapter 3 addresses the earlier noted temptation when working with a PIM to make enough background assumptions to achieve model identification, even though these assumptions may be untenable. But inferences from an identified, but possibly misspecified model, generally have poor properties.

Throughout the monograph there are many examples of PIMS. Chapter 1 looks at misclassification of a binary variable. Chapter 2 considers an example of trying to infer properties of a joint distribution from marginal distributions. Chapter 3 considers a missing data problem and a binary misclassification in a case–control study. Chapters 4–6 consider PIM examples in more depth, generalizing some of the earlier examples. Chapter 4 looks at more models for misclassified data. Chapter 5 considers examples involving instrumental variables. Chapter 6 looks at further examples involving prevalence estimation with a hidden population, and gene–environment interaction and disjoint data sources. A common theme in these examples is examination of the ramifications of partial identification, particularly in terms of how inferences change, and particularly the extent to which they sharpen or not, as more data accumulate.

Chapter 7 discusses some further topics and returns to some of the earlier examples. There are some brief remarks on computational issues, followed by a characterization of the value of information obtained from data in a PIM context. This raises the question of how much data is worth collecting. Chapter 7 also reviews some of the real-data applications of PIMs that have recently appeared in the literature.

Chapter 8 closes with thought-full comments on the present state of affairs with PIMs and some historical comments on their development.

In summary, Gustafson has drawn together many discussions of identifiability from previous Bayesian analyses (including his own), which are not widely known in non-Bayesian circles. The writing is concise. The examples are simple and insightful. The reader need not be a Bayesian to appreciate this fine monograph.

Dale J. Poirier
University of California, Irvine

Clinical Trial Biostatistics and Biopharmaceutical Applications. Walter R. Young, and Ding-Geng (Din) Chen, eds. Boca Raton, FL: Chapman & Hall/CRC Press, 2014, xxxv + 544 pp., $119.95 (H), ISBN: 978-1-48-221218-1.

Currently in its 72nd year, the Deming Conference was originally organized in close collaboration with the former Statistics Department of Princeton University and has more recently expanded to involve experts with a broad range of statistical experience throughout industry and academia. The annual conference aims to provide a learning experience on recent developments in statistical methodologies. *Clinical Trial Biostatistics and Biopharmaceutical Applications* is a collection of chapters contributed by prominent Deming Conference speakers highlighting novel methodological developments applicable in clinical trials and biopharmaceutical science. The two editors are experts in this field. Walter Young has been active in the biopharmaceutical industry for over five decades, with career positions at Lederle Laboratories and Pfizer, and as the long-time Chairman of the Deming Conference on Applied Statistics. Young's coeditor, Ding-Geng (Din) Chen, is a Clinical Professor in Biostatistics at the Gillings School of Global Health at the University of North Carolina at Chapel Hill and has more than 20 years of experience in academics and the pharmaceutical and biotech industries.

Two historical notes presenting an overview of the Deming Conference itself precede, and set the stage for the origins of, the technical content. The first, by Princeton University's Stuart Hunter, describes the evolution of the focus, environment, and sponsorship of the conference. Then Walter Young volunteers "Some Nonstatistical Reminiscences" of his tenure as the Deming Conference Chairman, honoring those individuals and organizations that kept the conference running through the inevitable growing pains.

The rest of the book is divided into five sections, each containing several chapters that are broadly applicable in clinical and biopharmaceutical research. Thirty-seven field experts from academic and governmental institutions, as well as prominent pharmaceutical companies, contributed sections on a range of topics that are imperative in the proper conduct of modern clinical studies.

Section I, Emerging Issues in Clinical Trial Design and Analysis, addresses several important issues that may arise in the planning and design of a trial, including regulatory considerations for various trial designs, appropriateness of randomization methods, first dose ranging in drug development, QT/QTc clinical trials for nonantiarrhythmic drugs, and controversial issues in noninferiority trials. Section II, Adaptive Clinical Trials, is divided into three chapters: a discussion of the roles of adaptive designs in drug development trials, then a description of how to use adaptability to optimize group sequential design in the context of survival trials with nonproportional hazards, and finally a demonstration of group sequential design in the R package gsDesignR. Issues of design, analysis, dose-finding, and data-censoring, in the specific context of oncology clinical trials are addressed in Section III. Section IV, Multiple Comparisons in Clinical Trials, begins with an introduction to the general concept of multiple test problems and then presents techniques to correct for multiplicity in adaptive trial designs, trials requiring multiple simultaneous pairwise comparisons based on the risk difference, and trials with multiple correlated endpoints. This section also includes a chapter on graphical approaches to multiple testing that yield intuition-building visualizations of both standard multiple test procedures and more modern "gatekeeping procedures" to control the familywise error rate. Section V, Clinical Trials in a Genomic Era, addresses some of the increasingly common "big data" challenges faced by statisticians and bioinformaticians. The first chapter describes classification, validation, and survival prediction methodologies for *-omics* data. The second chapter outlines several uses of biomarkers in clinical studies, with specific emphasis on the role of biomarkers in understanding therapeutic pathways. The section closes with a chapter on the statistical evaluation of surrogate endpoints in clinical studies.

Clinical Trial Biostatistics and Biopharmaceutical Applications is presented as a tool for investigators and practitioners. The reader's familiarity with the fundamentals of clinical trial design and analysis is assumed. Indeed, an appreciation of the coincident statistical, logistical, and ethical issues inherent in reputable, reproducible clinical research is valuable to understand the motivation for the methodological developments described in this book. Beyond the broad groupings into thematically related sections, each chapter is self-contained with a detailed table of contents and collection of selected references. There is no sequential development or cross-referencing across chapters or sections. Instead, each chapter briefly introduces a topic, establishes context through definitions or practical examples, then describes current methodology or extensions to deal with the highlighted study design feature or data type. The requisite statistical and mathematical knowledge varies by chapter, from simple probabilities and percents in Chapter 2, Review of Randomization Methods in Clinical Trials, to sophisticated model-building and model development, probability theory, and matrix algebra in Chapter 19, Statistical Evaluation of Surrogate Endpoints in Clinical Trials. Illustrations of the methodologies in each chapter similarly run the gamut from basic toy examples and tables, to simulation studies, to real clinical trial data.

The episodic structure of this text makes it unsuitable as a single course textbook. However, some individual chapters could be useful complementary reading for a graduate level biostatistics or epidemiology course. For example, Chapter 5, Controversial (Unresolved) Issues in Noninferiority Trials, could nicely supplement a course in the design of clinical studies; it presents a clear explanation of noninferiority trials and discusses several trial components (missing data, choice of analysis set, multiple comparisons, and adaptive designs) for which standard approaches used in superiority trials are inappropriate or require further development. Chapter 10, Competing Risks and their Applications in Cancer Clinical Trials, carefully distinguishes cause-specific-hazards and cumulative-incidence analytic approaches in the presence of competing risks, and might be a worthwhile addendum in a survival analysis methods course.

Overall, the contributed chapters in this collection are clearly written, easily digestible, and well-referenced. The book is a useful resource for the clinical trialist interested in obtaining a quick overview of standard practices and current methodological development for a specific biostatistical application, or a worthwhile read for the researcher seeking to familiarize him or herself with a diversity of emerging topics in clinical trials.

Megan T. Smith
University of California, Irvine

Design and Analysis of Experiments, Volume 3: Special Designs and Applications.

Klaus Hinkelmann (ed.). New York: Wiley, 2012, xxvii + 555 pp., $145.00 (H), ISBN: 978-0-47-053068-9.

This book is the third in a three-volume series on the design and analysis of experiments. The first two books in the series are *Design and Analysis of Experiments, Volume 1: Introduction to Experimental Design, Second Edition* and *Design and Analysis of Experiments, Volume 2: Advanced Experimental Design*, both by Klaus Hinkelman and Oscar Kempthorne. The motivation for the third volume, as described in the preface, is that the ideas discussed in the first two volumes "have, over the years, been expanded for and adapted to special situations and applications by many researchers. The stimulus has come very often from scientists and practitioners working in applied fields, such as genetics, medicine, marketing, manufacturing, industrial production, agriculture, forestry, pharmacy, engineering, defense, national security, and others. These areas may require special adaptation or implementations of existing designs and/or special methodologies for analyzing data from such experiments. The reason for writing *Volume 3: Special Designs and Applications* is to acquaint readers with these types of problems. Each of the 15 chapters gives an introduction, often with historical background, to the topic under consideration and then discusses solutions to the particular problems with references to most recent results."

The chapter titles, which also describe the topics discussed, are the following: Genetic Crosses Experiments (Chapter 1), Design of Gene Expression Microarray Experiments (Chapter 2), Spatial Analysis of Agricultural Field Experiments (Chapter 3), Optimal Designs for Generalized Linear Models (Chapter 4), Design and Analysis of Randomized Clinical Trials (Chapter 5), Monitoring Randomized Clinical Trials (Chapter 6), Adaptive Randomization in Clinical Trials (Chapter 7), Search Linear Model for Identification and Discrimination (Chapter 8), Minimum Aberration and Related Criteria for Fractional Factorial Designs (Chapter 9), Designs for Choice Experiments for the Multinomial Logit Model (Chapter 10), Computer Experiments (Chapter 11), Designs for Large-Scale Simulation Experiments with Applications to Defense and Homeland Security (Chapter 12), Robust Parameter Design (Chapter 13), Split-Plot Response Surface Designs (Chapter 14), and Design and Analysis of Experiments for Directional Data (Chapter 15). As this list indicates, the topics and areas of application are quite diverse.

I found the overall quality of the chapters to be quite good. The authors of the chapters are well-known experts and have been active in research on the topics they discuss. The chapters range in length from 15 to 70 pages, but most are about 30 pages. This is not long enough for a comprehensive treatment of a topic, but it is sufficient to provide a readable introduction and overview, not unlike survey articles that appear in journals. Each chapter includes extensive references and the book contains both a name and topic index. Some chapters include numerical examples, as well as data and SAS, S-plus, JMP, or Minitab output from the examples. The interested reader can then try to replicate the analysis. In a few cases, SAS and S-plus code for carrying out analyses is provided.

Most books on the design and analysis of experiments, including many advanced books, do not provide a picture of the breadth of research issues in experimental design and analysis. This book takes a step toward rectifying this. Many of the topics discussed appear in other textbooks and there are entire books devoted to a few of the topics (e.g., fractional factorial designs, computer experiments, and split-plot designs). However, I am not aware of any book that matches the topical breadth of this book. I believe this book will be useful to researchers (including graduate students) who are looking for a readable introduction to new research topics in experimental design. These are not always easy to find. The book would also make a good reference for an advanced course in special topics in the design and analysis of experiments and is a nice addition to the Volumes 1 and 2 of this series.

William I. Notz
The Ohio State University

Geometry Driven Statistics.

Ian L. Dryden and John T. Kent (eds.). New York: Wiley, 2015, xviii + 394 pp., $120.00 (H), ISBN: 978-1-11-886657-3.

This is a book dedicated to Professor Kanti Mardia and his long-lasting and impressive contribution to many areas of statistics including multivariate analysis, directional data analysis, spatio-temporal modeling, and shape analysis. The book also reviews Mardia's contribution to many application areas such as geophysics, medicine, and bioinformatics. As the title suggests, the common theme among these diverse topics is geometry and its application in statistics. As someone who firmly believes that geometry provides a natural language and a robust framework for dealing with statistical problems, I found the book quite fascinating. It provides a collection of articles from high-profile researchers, many of whom have collaborated with Mardia in the past. Each article describes a geometrically motivated statistical method and presents its applications for solving scientific problems. Although the statistical methods discussed in these articles seem quite diverse, the editors emphasize that they all fall under the category of "Geometry Driven Statistics."

The book starts by providing historical background on Mardia's life and work. The first two chapters are based on two separate interviews conducted by Nitis Mukhopadhyay, one in 1999 (prior to Mardia's retirement in 2000) and another one in 2014. In these two chapters, Mardia explains how he became interested in different topics such as directional data analysis and spatial statistics. He also reviews some of his influential publications.

Although these two chapters focus on Mardia's work in different areas of statistics, they also provide a historical background (albeit from Mardia's point of view) on how these areas evolved throughout his academic life. Chapter 3 simply includes a list of selected publications by Mardia and his co-authors.

The remaining chapters are divided into four parts focusing on different areas of Mardia's research: (1) directional data analysis, (2) shape analysis, (3) spatial, image, and multivariate analysis, and (4) bioinformatics. However, the editors point out again (and show this with the cartoon illustration on the book cover) that these are all related topics, with geometry as their unifying theme.

*Directional data analysis.*This part of the book is devoted to statistical methods involving parameters that are located on a circle or more generally are constrained to the surface of a sphere. As discussed in Chapter 5, such parameters are typically the result of an oscillatory system, for example, sales of a seasonal product or biological rhythms. Chapter 6 discusses circular-circular regression models for investigating the relationship between a circular response variable and a circular explanatory variable (e.g., spawning time and the time of low tide in marine biology). The remaining two chapters in this section deal with two-sample tests for directional data and the use of Riemannian barycenters in data analysis (more specifically, hurricane trajectories).

Shape Analysis. Articles presented in this section are related to the study of more sophisticated manifolds (i.e., beyond the sphere). Chapter 8 focuses on geometric morphometrics using the formalism of the bending energy of the thin-plate spline. Chapter 9 discusses nonparametric methods for shape-based statistical analysis of medical images. Chapters 10 and 11 are related to distributions on shape spaces and methods that solve the registration problem as a part of shape analysis.

Spatial, CAO Image, and Multivariate Analysis. This section focuses on evaluating model diagnostics (Chapter 12), forecasting (Chapter 13), and visualization (Chapter 14) in the context of spatial models and multivariate statistical methods. The last chapter in this section is devoted to fingerprint image analysis.

Bioinformatics. The remaining part of this book presents several chapters related to geometric methods in bioinformatics, which has been a main interest of Mardia over the past decade. The topics include: analysis of protein structures (Chapter 16 and Chapter 18), divergence measures (Chapter 17), and methods for protein matching and alignment.

Overall, I think this book provides a nice overview of some geometry-driven statistical methods and their applications in directional data analysis, shape analysis, spatial statistics, and bioinformatics. It could be an excellent starting point for students and researchers who are interested in these areas. Also, I encourage new researchers to carefully read Section 2.10, where Mardia discusses his views on the future of statistical science and outlines some of the most important challenges that statisticians would face over the next few decades.

Babak Shahbaba
University of California, Irvine

Handbook of Mixed Membership Models and Their Applications. Edoardo M. Airoldi, David M. Blei, Elena A. Erosheva, and Stephen E. Fienberg (eds.). Boca Raton, FL: Chapman & Hall/CRC Press, 2014, xxxiii + 586 pp., $119.95 (H), ISBN: 978-1-46-650408-0.

Traditional clustering places data points into mutually exclusive categories. Mixture models allow a model-based approach to clustering; each component captures the distribution of data points within a single cluster, and the mixing proportions form a distribution over clusters since each data point must belong to exactly one cluster. But, as the *Handbook of Mixed Membership Models and Their Applications* makes clear with a number of compelling applications, it is often the case that data points can more realistically be said to belong to more than one group at a time. For instance, a document may cover a number of different topics; a voter's political attitudes may derive from multiple ideologies; and an individual in a social network may interact with others based on a number of different personal identities (work, family, friends, and so on). In this case, we say that the data point exhibits *mixed membership* across multiple groups or classes. *Mixed membership models* provide a modeling framework for this broader setting. Whereas a number of excellent texts have, over many decades, addressed approaches to mixture modeling, there seems to have been no comprehensive overview of mixed membership models before the *Handbook of Mixed Membership Models and Their Applications*.

The editors of this volume—Edoardo Airoldi, David Blei, Elena Erosheva, and Stephen Fienberg—have notably worked at the forefront of research in various subfields of mixed membership modeling since the field's inception. Indeed, the preface, which outlines their initial meetings and interactions over the span of two decades, can also be read as a succinct history of some of the major developments within mixed membership modeling. The strength of their collaboration in editing this volume can be seen in the book's organization. The volume is divided into six major modules: an introduction to mixed membership models, the grade of membership model, topic models, semi-supervised models, sequence and rank data models, and finally network models. These modules mostly represent a natural partition of current research within mixed membership models and seem to roughly correspond to the historical research interests of the editors. The one module with less cohesion than the others is the module on sequence and rank data, which appears to be a catch-all for models and applications that did not fit elsewhere. For the reader, the generally strong organization yields a convenient introduction to each research community in turn.

Each chapter within a module generally introduces an application area, describes a single model or a few closely related models, and provides an experimental evaluation. This style of presentation makes the book ideally suited for a reader with some prior experience in mixture modeling or mixed membership modeling. This volume will provide such a reader with a broad perspective on recent developments—including a diversity of models and applications—within the field. In particular, an individual actively working within one of the subcommunities of mixed membership research would benefit from consulting this handbook. Compared to choosing articles by following citations, this handbook would give such a reader a much better sense of the very broad range of work that exists in the field as a whole. If the reader is interested in a particular modeling or application area, the organization into modules makes finding the relevant material very easy. Thus, the book serves its purpose well as an overview or snapshot of a field of research. It is not intended to be a teaching textbook or cover-to-cover read.

One of the main strengths of the book, fulfilling the promise of its title, is the wealth of applications described therein: the book provides mixed membership models for political ideology from surveys, medical diagnosis, risk factors for chagas disease, progression of dementia, Irish election data, document corpora, and ancestral genetics, to name just a few. The majority of these chapters are accessible to the statistician or data scientist, though there is an occasional chapter that would be better read with more specialized knowledge of an application (e.g., basketball or genetics). The book chapters cover both modeling and inference for these applications in extensive detail, often with full derivations of Monte Carlo samplers or variational approximations. The reader will not find more traditionally theoretical concerns such as asymptotics, consistency, and related results in this volume.

We would like to highlight one unusually strong chapter: "Care and Feeding of Topic Models" by J. Boyd-Graber, D. Mimno, and D. Newman. Not only is this chapter a pleasure to read, but it is an exemplar of material one would hope to find in a traditional "handbook." The chapter meticulously guides the reader through the details of approaching a particular mixed membership application (topic models for text) in practice: how to preprocess data, how to initialize and track inference, how to update hyperparameters, how to interpret and diagnose results, and perhaps most importantly, common failure modes and how to identify them. These are exactly the details that are often swept under the rug in conference and journal publications but are incredibly important to the end user of these models. We especially recommend this chapter to any practitioner, at any level, with an interest in applying topic models.

While the high-level organization of the present volume is strong, the editing at the individual chapter level occasionally leaves something to be desired. Multiple chapters within some modules repeat the introduction of the same model (e.g., latent Dirichlet allocation) using different notation. Section 23.5.2 appears to be missing from Chapter 23; there is a section title, but no text. Finally, Larry Wasserman contributed a compelling critique of mixture models—many of the salient points of which carry over to mixed membership modeling—to the introduction from his blog (with permission to the authors). We would have nonetheless preferred if it were more clear that this section was essentially an extended quotation.

The *Handbook of Mixed Membership Models and Their Applications* draws from the experience and research strength of its editors and authors to deliver both an overview of the historical development of mixed membership modeling and a snapshot of the current state of the art. The diverse set of chapters covering cutting-edge research suggests that mixed membership, like mixture modeling before it, is becoming an essential tool for the modern statistician. We believe this book sets the stage for a rich body of future work.

Trevor Campbell and Tamara Broderick
Massachusetts Institute of Technology

Inferential Models: Reasoning with Uncertainty. Ryan Martin, and Chuanhai Liu. Boca Raton, FL: Chapman & Hall/CRC Press, 2015, xix + 254 pp., $89.95 (H), ISBN: 978-1-43-988648-9.

This book promises a lot, and it certainly delivers in terms of new insight and improved understanding of old ideas. It demonstrates an approach to the development of inferential methods that arose from its authors' attempts to understand and extend the foundations of statistical inference. The inferential models (IM) of the title encompass more than the distributional and sampling rules that make up what is often called a statistical model. They form a framework for development of methods of inference that might be understood as extensions and replacements for Fisher's ill-fated fiducial argument and Dempster-Shafer belief functions. The major claim is that the IM approach yields optimal inference based on prior-free probabilistic inference about the parameter of interest; the approach allows "conversion of frequency probabilities, which come from the posited model for the observable data, to belief probabilities about the parameter of interest" (p. 6) without a prior probability distribution for that parameter. It seems to offer the proverbial Bayesian omeletes without breaking the Bayesian eggs, an exciting prospect that will, no doubt, be met with some resistance and skepticism.

Why is it that we cannot use existing methods that eschew Bayesian priors to obtain prior-free probabilistic inference? Say we wish to make probabilistic statements about θ based on an observed value $X = x$ that we assume is a random deviate from a known distribution with parameter vector $\theta \in \Theta$. The observation x allows us to derive a likelihood function because we know the relevant sampling distribution, and that function allows ranking of various potential values of θ with regard to their plausibility given the observation of x. However, as Fisher pointed out, likelihoods are not "full-blown" probabilities and the likelihood "communicates evidence of a type too weak to supply true probability statements" (Fisher 1973, p. 75). A Bayesian combination of the likelihood function with a prior probability distribution for θ yields probabilistic inference for θ, but this is obviously prior-dependent and it is not always possible or desirable to use a prior. Unfortunately, without a prior there is no usable probabilistic aspect of θ, so the combination of x with a statistical model does not yield probabilities for the parameter of interest. Frequentist testing methods circumvent this difficulty by externalizing the probabilities into calibrated probabilities of erroneous decisions, but the trade-off is that they ignore much of the evidential content of the data.

The promise of the inferential model approach is to provide a prior-free probabilistic predictions of θ that contain the full evidential content of the observations along with full frequentist calibration of error probabilities. A key procedure of IM is the exploitation of auxiliary variables about which probabilistic statements can be made even though they are unobservable. For example, suppose θ is a location parameter and consider a variable $u \in U$ where $u = \theta - x$. While the distribution of X can only be known as a function of the unknown θ, the distribution of U can be fully obtained from the statistical model even though u is unobservable. Given that knowledge of x and u would give an exact value for the parameter of interest, θ, predictions about u can be used to derive prior-free probabilistic statements about θ.

The book is based fairly closely on a small collection of its authors' articles, but its collation, introductory, and linking sections, and extensions make the book more than worthwhile even for those who have read the articles. The first chapter is a brief presentation of necessary background material such as probability theory, standard approaches to statistical, and scientific inference, along with a discussion of how prediction can be used for inference, an idea that is important for understanding the IM framework. The final chapter is devoted to discussion of areas for future research, so the book should serve the authors' intention well that it be the beginning of the development of tools for optimal inference.

The book does have some aspects that might diminish its usefulness to some readers. The index is remarkably short at less than 2 pages and its coverage is therefore limited. That weakness is shared by many books of similar nature, but it is lamentable nonetheless. The exposition of the IM procedures is quite mathematical, and this reader, only modestly skilled in mathematics, would have been more comfortable with textual explanations less reliant on formulas. While there are worked examples throughout the book, some of them included explanations that were just as formula-based as the surrounding material, and thus were less useful to me than they would be to a more mathematically trained reader. I similarly found the exercises at the end of each chapter to be challenging. However, I note that the level of mathematics required to make sense of the book would not be higher than that attained by the average statistician, despite being a stretch for the average pharmacologist! The book surmounts these problems easily and largely delivers on its promise. It should be read by all statisticians with an interest in the foundations and development of the statistical methods for inference.

Reference

Fisher, R. A. (1973), *Statistical Methods and Scientific Inference* (3rd ed.), New York: Hafner Press. [1846]

Michael J. Lew
University of Melbourne

Mathematical Statistics: Basic Ideas and Selected Topics, Volume II.

Peter J. Bickel and Kjell A. Doksum. Boca Raton, FL: Chapman & Hall/CRC Press, 2015, xix + 465 pp., $99.95 (H), ISBN: 978-1-49-872268-1.

The first edition of Bickel and Doksum's *Mathematical Statistics* appeared in 1977, almost 40 years ago. During the years that have passed statistics has advanced and changed in dramatic ways. Responding to these changes, the second edition of *Mathematical Statistics* (2015), now in two volumes, is a major and thorough revision of the original text. These volumes are intended as texts for Ph.D. level courses in mathematical statistics. Volume I of the revision presents classical statistical concepts at an advanced level, and Volume II considers more modern material, covering developments mainly from the years that have passed since the first edition. The presentation in both volumes avoids use of measure theory, but the general level of presentation is quite advanced, and considerable mathematical maturity is needed to appreciate many of the derivations and proofs. I would not suggest trying to use either volume as the text for a masters level course, and Volume II would be most suitable for the second or third semester of a Ph.D. level sequence.

Volume II has six major chapters, numbered 7–12, following an introduction to the second volume, labeled Chapter I. Each chapter includes a sizeable collection of problems in its last section. The authors note that the chapters cover "selected topics," and are not intended to include everything of potential interest to current Ph.D. students and instructors. But although other areas of interest can be easily suggested, the authors have done a superb job of selecting topics comprising most of the essential knowledge needed for modern research. Furthermore, these modern topics are considered with greater depth and sophistication than is usual in a general purpose text. And throughout its pages the book does a good job of linking the mathematical developments to major examples. The choice of topics and examples, along with the depth of coverage are the most attractive features of this volume. These strengths generally outweigh my concerns about the book, detailed near the end of this review.

Following the introduction, Chapter 7 presents tools for asymptotic analysis that will be used in the sequel. It begins with a discussion of separable stochastic processes and weak convergence for random functions in $l_\infty(T)$, followed by results from empirical process theory, including a discussion of maximal inequalities based on bracketing. After this there is a discussion of the delta method in function spaces, including a discussion of Gâteaux (directional) and Fréchet (total) derivatives, and a final section with a discussion of von Mises and Hoeffding expansions. Examples in this chapter give approximations for Kolmogorov statistics, Cramér-von Mises statistics, multivariate medians, quantile processes, linear combinations of order statistics, and von Mises and U-statistics.

The material covered in Chapter 8 is more classical. Similarity and completeness are discussed first, with the emphasis more on testing than unbiased estimation, followed by results on equivariant and minimax procedures, with some discussion of admissibility and the James–Stein estimator.

Chapter 9 gives an extensive treatment of semiparametric inference. The first section gives motivating examples; approaches to estimation based on numerous variants of the likelihood function (empirical/modified likelihoods, delta method likelihoods, and more); and methods for estimation based on the sieve method and regularization. Next, a section on asymptotic theory opens with a general consistency criteria, followed by an example-driven discussion of asymptotic normality. These examples include biased sampling models, with the Kaplan–Meier and Nelson–Aalen estimates as special cases, Cox proportional hazards models, and partial linear models. The final three sections concern efficiency in semiparametric models; testing and goodness of fit based on empirical process theory; and Le Cam style asymptotic theory for likelihoods, including local asymptotic normality and contiguity.

Monte Carlo methods are the main topic of Chapter 10. After sections about examples and basic methods, there is a section on bias correction and confidence intervals using the bootstrap, including asymptotic theory justifying its use and some discussion of the m out of n bootstrap for situations where the standard approach fails. The final section is devoted to Markov chain Monte Carlo, covering the Metropolis–Hastings algorithm and Gibbs samplers with applications to Bayesian and frequentist inference. However, details are provided mainly for finite chains.

Chapter 11 considers nonparametric estimation of density functions on the real line and regression functions of a single covariate. They start with kernel estimates for densities and derive uniform local approximations for the estimation bias and variance, and global approximations for integrated mean squared error, followed by a discussion of minimum contrast estimates to reduce bias near a boundary. Next they consider approaches to density estimation based on sieves or regularization, with some discussion of splines. These sections also give a few results on tuning and bandwidth selection, and results on confidence intervals and bands. The final section explores nonparametric regression in a similar fashion.

The book concludes with an extensive chapter on high-dimensional prediction and classification. Statistical modeling approaches based on kernel methods, sieves, and Bayesian and penalized least squares, are presented in tandem with ideas from machine learning, including neural networks, support vector machines, boosting, and methods based on tree structures. There are sections devoted to: classification and prediction; asymptotic risk, including material on optimal prediction, optimal rates of convergence, minimax rates for integrated mean squared error for subsets of the Gaussian white noise model, and sparse submodels; oracle inequalities for classification and shrinkage estimators (based on Stein's unbiased estimate of risk); cross-validation for tuning and measuring performance; model selection and dimension reduction, including inference after model selection; and current frontiers.

As much as I appreciate the content of this book, at times I found it more difficult to read than necessary. The book is very self-referential. In a way this is good as it helps link common ideas from different sections of text. But I felt that many passages had too few details to understand the intent without jumping elsewhere. In particular, the authors refer back to Volume I often enough that Volume II is difficult to read and use without a copy of the first volume at hand. Also, examples discussed in opening sections of chapters, related to developments in the chapter or previous discussions, can be confusing without hunting for details only stated in other passages. For a good understanding of the empirical process theory used in the book, readers may need to refer often to van der Vaart and Wellner (1996). The authors' writing style is also rather informal. At times this is appealing as it gives a sense of how the authors are thinking about the material presented, but at other times more precision would make derivations clearer and easier to understand. Similarly, the authors' decision to avoid measure theory seems too casual at this level. Although they do manage to convey the main ideas effectively without it, the choice limits generality and makes it difficult to express certain ideas accurately and succinctly. Finally, I should note that the book seems to have quite a few typographical errors.

While other general statistics texts at a similar level (Schervish 1996; Lehmann and Casella 2003; Lehmann and Romano 2008; Keener 2010; Shao 2010) touch on some of the topics covered in this book, none of them cover the modern material in this book with comparable depth. As such it is certainly a valuable contribution to our advanced literature on theoretical statistics.

References

Bickel, P. J., and Doksum, K. A. (2015), *Mathematical Statistics: Basic Ideas and Selected Topics* (Vol. I, 2nd ed.), Boca Raton, FL: Chapman and Hall/CRC. [1847]

Keener, R. (2010), *Theoretical Statistics: Topics for a Core Course*, New York: Springer. [1848]

Lehmann, E. L., and Casella, G. (2003), *Theory of Point Estimation* (2nd ed.), New York: Springer. [1848]

Lehmann, E. L., and Romano, J. P. (2008), *Testing Statistical Hypotheses* (3rd ed.), New York: Springer. [1848]

Schervish, M. J. (1996), *Theory of Statistics*, New York: Springer. [1848]

Shao, J. (2010), *Mathematical Statistics*, New York: Springer. [1848]

van der Vaart, A. W., and Wellner, J. A. (1996), *Weak Convergence and Empirical Processes: With Applications to Statistics*, New York: Springer. [1848]

Robert W. Keener
University of Michigan

Multivariate Density Estimation: Theory, Practice, and Visualization (2nd ed.). David W. Scott. New York: Wiley, 2015, xviii + 350 pp., $115.00 (H), ISBN: 978-0-47-169755-8.

The first edition of David Scott's book *Multivariate Density Estimation: Theory, Practice, and Visualization* was published in 1992, and presented a modern approach reflecting the vast research effort in this field. Its thorough examination of multivariate estimation was ahead of its time, with competitors such as Silverman (1986) and Wand and Jones (1994) focusing predominantly on the univariate case. Scott combined methodology and asymptotic theory with relevant data analysis, up-to-date computer graphics, problems for each chapter and available software, a format appealing to both practitioners and researchers. In the years since its publication, many of us have enjoyed dipping into the book for particular calculations, theorems, or sections.

After more than 20 years the second edition of Scott's book was published by Wiley in 2015. The immediate question is *What has changed?*

At a first cursory glance the reader will notice that the second edition is about 30 pages longer, and the presentation and display of most figures have changed, though mostly they convey the same data, analysis, and information. Unfortunately, the change did not include more color in the illustrations, as we are used to seeing these days in scatter or surface plots and which often aid the interpretation. Indeed, the color plates of the first edition depicting trivariate contour surfaces have been replaced by a black and white picture. This change could have been a consequence of "reengineering the figures from S-plus into R" (author's words) or Wiley's choice.

The second edition has the same nine chapters as the first, but some sections or parts of sections are new, and it is these I will focus on. Keeping in line with the first edition, the new parts contain motivation, valuable calculations, and theorems as well as data analysis illustrating the theory or method. In my opinion this has always been one of the strengths of Scott's book.

There are three definite highlights for me among the new parts:

- Polynomial histograms
- Zero-bias bandwidths and adaptive kernel estimators
- Clustering via mixture models and modes

These parts probably reflect most of the 20+ new references dated 2000 or later. The relatively small number of new references is an indication in itself that current research is no longer as focused on density estimation as it had been at the time of the publication of the first edition.

Polynomial histograms are described in a new nine page section that appears at the end of Chapter 4, "Frequency Polygons." Polynomial histograms circumvent the discontinuity issues of the common (flat) histograms, yet can be computed efficiently for dimensions where kernel estimators begin to become inefficient. In his last application section of the chapter, Scott illustrates some of these ideas on "spline-like" histograms, and refers the reader to the relevant literature.

The last nine pages of Chapter 6, "Kernel Density Estimators," acquaint the reader with zero-bias bandwidths, describe how to estimate these bandwidths in practice, and culminate in illustrating the superiority of adaptive kernel estimators in a bivariate example. The last section considers computational aspects of kernel methods and how to make these more efficient in the univariate as well as the multivariate setting.

Finally, Chapter 9, "Other Applications" contains the new Section 9.2.4, "Clustering via Mixture Models and Modes," which starts with Gaussian mixtures and mode trees with an emphasis on the number of modes when $d > 1$, and testing ideas for the existence of modes. The section refers to Sizer, the excess mass approach, and the iterative pairwise replacement algorithm (IPRA). Rather than providing details of these methods the author shows a pertinent illustration for each approach that may inspire the reader to examine one (or all) of the methods further.

These new parts were all enjoyable to read. Other changes include new pages or sections in

- Chapter 1: The distribution of pairs of points as the dimension grows
- Chapter 2: Criteria for measuring the difference between densities
- Chapter 3: Error criteria for bin width selection and bandwidth choices for derivatives of a density
- Chapter 6: Extensions of biased and unbiased cross-validation to a multivariate setting and an introduction to data sharpening

The three appendices on Computer Graphics in R^3, DataSets, and Notation and Abbreviations are unchanged from the first edition.

I found it a little disappointing that no new datasets were introduced in the second edition, and although Scott refers to massive datasets in the new introduction, I could not find evidence of any such being used in the book. Also, while the datasets used in the book can be downloaded from the publisher's webpage, I could not find any reference in the book or the webpage regarding availability of code.

Minor disappointments aside, Scott's second edition is a useful update to the original and well suited for teaching an upper-level undergraduate course or a beginners graduate course, and presents a valuable resource for statisticians.

References

Silverman, B. W. (1986), *Density Estimation for Statistics and Data Analysis*, Boca Raton, FL: Chapman and Hall/CRC. [1849]

Wand, M., and Jones, M. (1994), *Kernel Smoothing*, Boca Raton, FL: Chapman and Hall/CRC. [1849]

Inge Koch
Australian Mathematical Sciences Institute

Perfect Simulation. Mark L. Huber. Boca Raton, FL: Chapman & Hall/CRC Press, 2015, xxii + 228 pp., $89.95 (H), ISBN: 978-1-48-223244-8.

Among randomized algorithms that are designed to sample from a given target distribution, perfect sampling is king. Most randomized samplers, including the widely popular Markov chain Monte Carlo (MCMC) algorithms, produce dependent samples whose asymptotic distribution is the target distribution of interest. A central issue is understanding how long the simulation must proceed before the distribution of the chain's realizations is close enough to the target. Perfect simulation, on the other hand, generates *independent samples* with distribution *exactly equal to the target* in finite time, at least when it can be applied. Imagine the massive excitement that went through the statistical computation community when Propp and Wilson (1996) introduced the coupling from the past (CFTP) sampler in 1996. CFTP is based on a simple yet revolutionary idea that enables an MCMC sampling scheme to be redesigned to guarantee an iid sample from the target. Two important caveats were soon revealed: CFTP cannot be implemented with many MCMC algorithms, and when implementation is possible the running time may be prohibitively long.

As stated in the preface, "then around 2000 something interesting happened. Protocols that were very different from CFTP, such as the Randomness Recycler, were developed. CFTP was now not the only way to draw from the stationary distribution of a Markov chain!" (p. xxii). The author himself was at the center of this second wave of creativity, being the author of a number of important articles on perfect simulation. This places him ideally to write the first book fully dedicated to the development of the field. Moreover, he is known as a speaker who has the ability to convey complex mathematical ideas in a clear manner. The writing here follows suit; I found it crisp and concise. There were only a few places in the book where I could have benefited from a more extensive discussion.

The first three chapters review some of the fundamental concepts related to perfect simulation. Included here are some of the essential MCMC algorithms, a general definition of the Acceptance/Rejection method and the CFTP sampler. The latter two are singled out as the "two most important protocols for creating perfect simulation algorithms." I particularly like the elegant "Fundamental Theorem of Perfect Simulation" and how its use permeates the whole book.

Chapters 4, 5, and 6 (Bounding Chains, Advanced Techniques using Coalescence & Coalescence on Continuous, and Unbounded State Spaces) present different design techniques for integrating Markov chains within perfect simulation. Unlike CFTP some of these methods are interruptible in that an attempt that fails to produce a sample can be abandoned and be replaced by a new one. The read-once CFTP variant also avoids the need of running "from the past," and this modification simplifies implementation and can significantly reduce computation time. These methods are applied to well-known physics models (e.g., Ising model and the hard core gas models), the antivoter model and infinite graph models, among others. Although the list of applications of perfect sampling is growing, few fall within Bayesian computation, especially when compared with the huge impact MCMC has had on the analysis of complex Bayesian models. A possible companion to these chapters is the review article of Craiu and Meng (2005) that would complement the book with more emphasis on the statistical applications of CFTP, read-once CFTP and Fill's algorithm along with a particular focus on coupling techniques and variance reduction tricks that can be easily implemented for these procedures.

Chapter 7 is devoted to perfect simulation for Spatial Point Processes and the subsequent chapters move beyond the use of Markov chains for perfect simulation. Chapters 8 and 9 review the randomness recycler and advanced Accept/Reject, respectively. I found intriguing the concept of a Bernoulli factory and its use for generating solutions for general stochastic differential equations discussed in Chapter 10. Chapter 11 builds an interesting connection with the sampling methods for doubly intractable distributions, which is a domain that has recently seen much development. The last four chapters are areas of active ongoing research. They could be used by a researcher who wants to get up to speed with many of the most recent and promising ideas in perfect sampling.

The book has no conventional exercises. Going through its theoretical derivations, however, can be a useful, yet nontrivial exercise. The notions are clearly defined and illustrated by examples that start simple and grow more complex as the readers' understanding grows. The book is self-contained since there is "a bit of measure theory" in Chapter 1 and the proofs throughout the book are complete. The algorithms are presented in pseudocode. The whole book is probably too rich to be covered in a single semester graduate course in statistics or probability, but selected topics, especially the early chapters can be used. The students in such a course must have a solid background in probability.

It is not clear if perfect simulation will be fully embraced for statistical computation in the near future, but I like the optimistic tone at the end of the book: "Coupling from the past opened up a multitude of problems for simulation [...]. The set of problems addressable by perfect simulation protocols continue to grow, albeit more slowly than after the CFTP jumpstarted the field. [...] The most interesting open question about perfect simulation is this: how far can these ideas be taken?" (p. 218). In this reviewer's opinion, that is a question worth exploring.

References

Craiu, R. V., and Meng, X.-L. (2005), "Perfection Within Reach: Exact MCMC Sampling," in *Handbook of Monte Carlo*, eds. S. Brooks, A. Gelman, G. L. Jones, and X.-L. Meng, Boca Raton, FL: Chapman & Hall/CRC, pp. 199–226. [1850]

Propp, J. G., and Wilson, D. B. (1996), "Exact Sampling With Coupled Markov Chains and Applications to Statistical Mechanics," *Random Structures and Algorithms*, 9, 223–252. [1850]

Radu V. Craiu
University of Toronto

Statistical Learning with Sparsity: The Lasso and Generalizations. Trevor Hastie, Robert Tibshirani, and Martin Wainwright. Boca Raton, FL: Chapman & Hall/CRC Press, 2015, xv + 351 pp., $98.95 (H), ISBN: 978-1-49-871216-3.

The authors of some of the seminal articles on exploiting sparsity have written a book that describes the many applications of sparse methods and recent research results in the area. Exploiting sparsity in real-world problems has proven to be very effective in a wide range of application areas such as recommendation systems and signal reconstruction (compressed sensing), to name just two. This book introduces concepts with minimal jargon and is written assuming that the reader has only a basic background in optimization. However, understanding the advanced material toward the end of each chapter requires more knowledge of optimization techniques. The book includes all the major branches of statistical learning. For each topic, the authors first give a concise introduction of the basic problem, evaluate conventional methods, pointing out their deficiencies, and then introduce a method based on sparsity. Thus, the book has the potential to be the standard textbook on the topic.

Each chapter includes comprehensive references and notes describing the evolution of its methods. For many the authors have developed toy problems to illustrate how the lasso-based solution depends on tuning parameters. The corresponding plots and other figures are helpful in developing the intuition behind why the lasso methods work. Additional detail is provided as exercises. Thus, the exercises complete the description of the methods in each chapter and are not just instruments for assessing how well the reader has understood the material. The complexity of the exercises makes the book suitable to be used as a textbook for an advanced graduate course. However, the book does not include detailed case studies of applications or code that can be directly applied to datasets. In the context of a formal course, the instructor could provide additional material to cover applications and thus provide a mix of both theoretical foundations and practice.

Chapter 2 describes the core principles behind the lasso and its properties. The authors present the advantages of exploiting sparsity while ensuring that the solutions can be efficiently computed, a recurring theme in the book. The authors draw parallels with other statistical approaches and in particular show how sparse constraints can be interpreted from a Bayesian perspective. This chapter includes an overview of the theory and its historical development. In Chapter 3, the authors give the formulation for different generalized linear models, including for special cases that arise in practice such as when observations are sequences of different lengths. A real-world but easy-to-follow example is provided for each generalized linear model. In Chapter 4, the authors generalize the lasso principle to different penalty models. Again, the authors present real-world problems where such generalization is useful. Chapter 5 is an overview of optimization methods that are applicable to the lasso and its variants and includes a runtime comparison of two methods. This overview will be useful to the practitioner who needs guidance in choosing an appropriate optimization algorithm. Chapter 6 presents techniques to compute the statistical strength of the variables from the lasso and includes both traditional methods such as the bootstrap and newer research directions.

The remaining chapters present how the sparsity property can be applied to several standard problems with matrices, multivariate analysis, and graphical models. Matrix completion methods based on exploiting the sparsity principle are described in Chapter 7. The chapter includes a description of the popular Netflix challenge and the winning solution. Thus, this chapter can also be used independently to augment reading on recommender systems. Chapter 8 covers sparse variants of principal components analysis, correlation analysis, linear discriminant analysis, and clustering. Multiple ways of using sparsity for graphical model structure learning are presented in Chapter 9, with a particular emphasis on Gaussian graphical models.

The parallel development of compressed sensing techniques is covered in Chapter 10 along with sparse signal approximation methods. The book concludes with a set of theoretical results on bounds on the magnitude of error in the lasso solution and proofs for these. Figures are provided to illustrate these results in small problems and this helps in developing an appreciation of the theory even without an in-depth understanding of the theorems.

Anand Panangadan
California State University, Fullerton