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 Indeed, the color plates of the first edition depicting trivariate contour surfaces have been replaced by a black and white picture.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