Elements of Large-Sample Theory.

Erich L. LEHMANN. New York: Springer, 1999. ISBN 0-387-98595-6. xii + 631 pp. $79.95.

This introductory book on the most useful parts of large-sample theory is designed to be accessible to scientists outside statistics and certainly to master's-level statistics students who ignore most of measure theory. According to the author, "the subject of this book, first-order large-sample theory, constitutes a coherent body of concepts and results that are central to both theoretical and applied statistics." All of the other existing books published on the subject over the last 20 years, from Ibragimov and Has'minskii in 1979 to the most recent by Van der Waart in 1998 have a common prerequisite in mathematical sophistication (measure theory in particular) that do not make the concepts available to a wide audience.

The idea of a much more elementary book is a brilliant one (which, in the Preface, the author attributes to his wife, Juliet Schaffer). There is a real need for scientists from other fields to understand and use such notions as efficiency, consistency, robustness, nonparametric estimation and goodness of fit. Many examples of reinventions of these notions may be found by browsing biological, computer science, physics, economics, and chemistry journals. The fault for ignorance of the more complex features of statistics is not to be lightly assigned to the creative *rediscoverers,* but rather to statisticians. Most classical mathematical statisticians' articles and books treating these subjects are unreadable for the lay public. Unfortunately, these successive reincarnations build communication gaps between the different communities and statisticians themselves. This book should help curb this trend, and I am happy to recommend it to my non-mathematically trained colleagues.

The table of contents gives a good idea of the book's structure: Mathematical Background, Convergence in Probability and in Law, Performance of Statistical Tests, Estimation, Multivariate Extensions, Nonparametric Estimation, and Efficient Estimators and Tests.

The book starts with all of the essentials that a nonmathematically trained scientist often lacks. I have already successfully sent Figure 2.61 in Chapter 2, of the difference between uniform and nonuniform convergence, to several master students in statistics and Ph.D students from other disciplines.

This book's self-explanatory style makes it ideal for isolated scholars from other disciplines. The professorial voice is present throughout the text in a most friendly way. Many exercises are given to illustrate each chapter, with useful hints provided when necessary.

The author is well known for his scholarship, but the reader will not be overwhelmed by too many extensions or research articles. In most cases, a theorem is credited in the text to its author, and each chapter ends in a complete bibliographical review of relevant material.

This book clearly reveals what a "gentle man" the author really is in the literal sense—each notion is gently introduced, first in its simplest form, then generalized. For instance, on page 488 an example showing the efficiency of a Bayes estimator for the binomial with a beta prior is developed, then the reader is led to the general asymptotic efficiency result for Bayes estimators. This introduction *by example* is most helpful in leading to the general theory painlessly, providing both motivation and illustration.

I can only find one slight lapse in the author's discernment in his choice of presentation and material. Chapter 2, Section 2.7 includes the presentation of limit theorems in the cases when the random variables are not independent or are not identically distributed.

The book combines a delightfully rich perspective as the author spans so much of modern statistics' development. Take, for instance, Von Mises' "plug-in" (A term, I am told, is from Ned Glick in 1972) principle. It's current incarnation is Efron's bootstrap and the clarity with which the author develops this picture makes exposing it to master level students a real pleasure. Actually the entire Chapter 6, which is dedicated to nonparametric estimation, should be a suggested reading for many biologists/economists/philosophers who assimilate statistics with parametric statistics, believing the other type to be simply *heuristics.* This chapter opens with the simplest plug-in estimates $h(\hat{F}_n)$, which can be written as expectations and that can be estimated by U and V statistics.

The more general case is developed by using linearization techniques, the delta method, its generalizations to statistical functionals, and the influence function. This provides real insight into much of what asymptotics is about. (There are even statisticians who claim that the only really useful theorem in mathematical statistics is Taylor's theorem!)

Finally, the author shows his vision of the bootstrap as a simple extension of the plug-in principle to include the case where the functional depends both on the distribution and the sample size. Again, this way of leading the reader step by step through simple examples all the way to being able to grasp the proof of the bootstrap's consistency properties is a real achievement.

Of course, Chapter 7 on efficient estimation can be prescribed to the many researchers now wedded to the maximum likelihood estimator. It provides the methods of comparing the various choices available for estimators, although the author does overemphasize his opinion that the first property expected of a good estimator in the iid case is that it be consistent. At this stage I would have also included a sharp recall to those who use maximum likelihood estimation that such a conclusion is moot as soon as the assumptions of the model's validity are dropped, and robustness must always be considered concurrently. The brevity with which the concept of identifiability was treated is to be regretted, as there are few books that clearly illustrate and define this notion.

Asymptotic theory demands a growing number of observations. The book uses the classical fiction of triangular arrays to make this rigorous. This embeds $X_1, \ldots, X_n$ into an array $X_{ni}, 1 \leq n \leq \infty, 1 \leq i \leq n$. The distributions in different rows can be arbitrarily dependent, or even unspecified.

This sometimes seems natural; for example, in the classical matching problem, a Poisson limit is proved for $S_n = X_1 + X_2 + \cdots + X_n$ with X_i 1 or 0, depending on whether the ith place of the permutation is fixed or not. These $X_1, \ldots, X_n$ cannot be extended in any symmetric fashion, and triangular arrays seem natural. One recent advance in asymptotic theory replaces the hypothetical triangular array by an approximation theorem with error for the actual variables under consideration. For instance, in the permutation example one proves, for n fixed, $\|\mathcal{L}(S_n) - Po(1)\| \leq 2/n!$.

Such tools as Stein's method (Stein 1986) and coupling inequalities offer close to automated ways of deriving such approximations with explicit errors. It would have been nice to incorporate these ideas into statistical applications, as the methods do not use measure theory or any mathematics beyond the prerequisites of this book.

The book has some epidermic defects that will disappear with future editions with regards to some of the mathematical copy-editing, which does not seem quite on a par with the high quality of all the other aspects of the book. For instance, the characters on the Figure 7.1.1 (p. 454) are set in an unreadable way; after a long decryption, one realizes that the font is too large and that the expressions $\theta - c(\theta)$ and $\theta + c(\theta)$ were supposed to lie on the same line. For a $\TeX$ "æsthete," the parentheses are a little painful to follow; opening left parentheses of a certain size like to be closed with the same-sized right counterpart, and when two quantities are going to become infinitely similar, as in $h(\hat{F}_n) \to h(F)$, why shouldn't the size of their brackets be the same, too?

Equation (6.5.14) has a $\to$ that leads nowhere special, and the brackets for the normal distribution's parameters are left unclosed in the middle of page 389. There are also a few author's names consistently mispelled in the references and author index, (e.g., W. Hoeffding, O. Zeitouni, and H. Jeffrey's). But I will stop quibbling and linger rather on what this book provides.

Elements of Large-Sample Theory will give many of us the opportunity to create the intermediate courses that we lack, between the elementary level and the specialists' courses that we teach from the authors' other books (Lehmann 1983, 1986; Lehmann and Casella 1998). These books stress small-sample results, which are valid in precisely specified parametric models. In a personal communication, Professor Lehmann has suggested that if he had to choose all over again, he would have developed the kind of useful approximations developed in the present book.

We can only be thankful for the time that Lehmann has set aside to write carefully a book containing a simple picture of the intricacies of the best in mathematical statistics, instead of striving, as many of our colleagues,

J. Rojo (ed.), *Selected Works of E. L. Lehmann*, Selected Works in Probability and Statistics, DOI 10.1007/978-1-4614-1412-4_97, © Springer Science+Business Media, LLC 2012

to add a few more technically beautiful but practically useless articles to our poor "unreadable" journals.

Susan P. HOLMES
Stanford University

REFERENCES

Casella, G., and Berger, R. L. (1990), *Statistical Inference*, Pacific Grove, CA: Wadsworth.

Ferguson, T. S. (1996), *A Course in Large-Sample Theory*, London: Chapman and Hall.

Glick, N. (1972), "Sample-Based Classification Procedures Derived From Density Estimators," *Journal of the American Statistical Association*, 67, 116–122.

Ibragimov, I. A., and Has'minskii, R. Z. (1981), *Statistical Estimation*, New York: Springer-Verlag.

Lehmann, E. L. (1975), *Nonparametrics*, San Francisco: Holden-Day.

——— (1983), *Theory of Point Estimation*, London: Chapman and Hall.

——— (1986), *Testing Statistical Hypotheses* (2nd ed.), London: Chapman and Hall.

Lehmann, E. L., and Casella, G. (1998), *Theory of Point Estimation* (2nd ed.), New York: Springer-Verlag.

Van der Vaart, A. (1998), *Asymptotic Statistics*, Cambridge, U.K.: Cambridge University Press.

Theory of Point Estimation (2nd ed.).

E. L. LEHMANN and George CASELLA. New York: Springer-Verlag, 1998. ISBN 0-387-98502-6. xxvi + 589 pp. $74.95.

This second edition of Erich Lehmann's 1983 classic text is a most welcome addition to the statistical literature. George Casella has joined in the writing of the new edition and both the (more) senior and (more) junior author have clearly left the stamps of their individual interests and styles on this excellent book.

The new edition is approximately 80 pages thicker and includes roughly 1,000 references, of which about 250 are post-1983. The original edition is clearly visible in the revision and many sections have been carried over with minimal changes.

The first edition was reviewed in JASA by Berger (1984) and Rao (1984). The reader is encouraged to read these reviews, as much of what they say (especially Berger) remains relevant to the second edition.

The major changes are splitting Chapter 4 in the first edition on "Global Properties" into two chapters, one (Chapter 4) on "Average Risk Optimality" (Bayesian Estimation) and one (Chapter 5) on "Minimality and Admissibility"; and concatenating (and reducing) the material in Chapter 5 (Large-Sample Theory) and Chapter 6 (Asymptotic Optimality) into a new Chapter 6 (Asymptotic Optimality).

As a result of these changes, the new edition includes quite a bit more material on Bayesian estimation and quite a bit less material on robust estimators (e.g., medians, trimmed means, and M, L, and R estimates). No doubt space considerations played a role in the decision to omit much of the discussion of robust alternatives to the sample mean. I would have preferred that at least some of this discussion had been retained; it added substantially to the utility and balance of the first edition as a graduate text.

In a somewhat minor change in organization, the discussion of "Convergence in Probability and in Law" is moved to Chapter 1.

Typographically, the new edition is not as attractive as the first. This seems largely due to the increased bulk of material included, which has resulted in a smaller type size and smaller section headings. I find the latter significantly more annoying than the former. A return to larger section headings would be most welcome and helpful without adding greatly to the page count.

One major improvement over the first edition is the placement of the bibliography at the end of the text as opposed to a separate bibliography at the end of each chapter. This excellent 44-page bibliography is extensive and up to date. It is a major addition and should be useful to and welcomed by purchasers of the text. It is in itself almost worth the price of the book.

The "Reference" section at the end of each chapter has been replaced by a "Notes" section. This section contains extensions of the discussions of selected topics, historical comments, and indications of other possible references. These "Notes" sections are typically expansions on mostly historical comments at the beginning of the "Reference" section of the first edition. They are interesting, informative, and well done.

The problem sets are an excellent source of additional information and practice, and they give good opportunities for testing and honing of skills (for students and teachers!). The problem sets are extensive (15–30 pages per chapter), well chosen, and of varying difficulty. The error rate seems to be quite low. One amusing lapse is in Problem 6.24 in Chapter 1, where one is to show that a particular matrix is inevitable (instead of invertible).

The new edition continues the tradition of the first in presenting the material at a non–measure-theoretic level. The careful scholarship and readability of the first edition is admirably continued. The density of typos seems quite small and that of technical errors extremely small. In the spirit of Berger's review, here are some minor errors and one slightly more major one to indicate that I have indeed read the book:

- Example 1.3 on page 229 gives a Bayes estimator for a Poisson sample (of size n) and gamma prior, but the expression given is for the case $n = 1$.
- In Example 3.2 on page 169, the scale-equivariant estimator should be $B|X|^r$ if $X < 0$.
- In Example 4.5 on page 116, $calF_0'$ should be $\mathcal{F}_0'$.
- In Problem 6.9 on page 295, the definition of the n-fold convolution is incorrect (convoluted?).
- In Equation 1.11 on page 88, the exponent of q should be $n - t$, not $n - 1$.
- The slightly more serious error is in the statement of Corollary 1.6 on page 311. The corollary claims that δ_Λ is minimax if and only if $\Lambda(\omega_\Lambda) = 1$, where ω_Λ is the set of θ for which the risk of the Bayes estimator δ_Λ takes on its supremum. The condition is sufficient but not necessary. A class of counterexample is provided by any proper Bayes minimax estimator of a p-variate normal mean for $p \geq 5$ with respect to a prior whose support is all of R^p. For all such procedures, the Bayes risk is strictly less than the supremum (minimax) risk.

This revised and updated edition of *Theory of Point Estimation* is an excellent addition to the literature. It is well written, up to date, well organized, and at an appropriate level. It should remain a standard graduate-level text and reference work for many years. The first edition has been my favorite statistics book for some time, and the new edition is a worthy successor.

William E. STRAWDERMAN
Rutgers University

REFERENCES

Berger, J. O. (1984), Review of *Theory of Point Estimation*, by E. Lehmann *Journal of the American Statistical Association*, 79, 941–942.

Rao, C. R. (1984), Review of *Theory of Point Estimation*, by E. Lehmann *Journal of the American Statistical Association*, 79, 942.

Mathematical Statistics.

Jun SHAO. New York: Springer Verlag, 1999. ISBN 0-387-98674-X. xiv + 529 pp. $74.95.

This text is a nice addition to the collection of graduate textbooks in theoretical statistics at the level of Lehmann's (1986) *Testing Statistical Hypothesis* and Lehmann and Casella's (1998) *Theory of Point Estimation.* It is intended as a first- or second-year graduate text for a two-semester course in mathematical statistics. Students should have a good knowledge of advanced calculus. In fact some knowledge of probability and, to a lesser extent, measure theory is a practical necessity, although the Chapter 1 gives a nice review/overview of measure theory and probability, albeit largely without proofs but with some nice examples. The book is generally quite rigorous and well written, with numerous good examples. It should be a good main or supplementary text for its intended audience.

Two aspects of the book separate it from its principal competitors. The first is the treatment of asymptotic aspects of statistics, which is integrated into the discussion of each topic as opposed to being separated out in a separate chapter. The second distinguishing feature is the treatment in a