mality. The major topic here concerns selecting the correct mean squares for testing. Chapter 10, devoted to post hoc multiple-comparison tests, is not very modern and does not incorporate the developments of the last 10 years. Chapter 11, on nonparametric tests, is also routine and outdated. For example, the author states incorrectly that the Kruskall–Wallis test can be used only with equal sample sizes and mentioning nothing of their use for randomized blocks. Chapter 12, Selection of Statistical Tests, offers a set of flow charts to aid the reader in selecting the appropriate test for the given design. These charts take the reader back to all the previous chapters for discussions of the design and selection of ANOVA tables for identification of appropriate mean squares for testing and for aid in selection of post hoc tests.

Section IV is the highlight of the book, presenting the 100 examples or, as the author calls them, the research design problems. A "learning by correction" strategy is used. Chapter 13 is a brief introduction to the problems; Chapter 14 contains 148 pages of them. Each research design problem contains a statement of the research problem, a description of the study design, a summary of the statistical analysis performed, the investigator's conclusion, and a critique by the author. When mistakes are identified, references to the first part of the book are given to allow the reader to review the correct design or analysis procedure that should have been used. No literature references are given for the problems, and the author is silent on their origin. They may have been generated by the author.

There are abundant number of errors identified in these research design problems. A small selection includes (numbers in parentheses are research design problem numbers) analyzing correlated within-subject data as if it were from independent subjects (4, 9, 10, 13, 47, 89), not taking the effects of dropouts into account (5), using nonrandomized controls to make causal inferences (7), confounding treatment effects with the placebo effect (14), having selection biases (7, 19, 52), having lack of balance in the treatment groups (28, 33), performing multiple t tests without correction for multiple testing (29), not having an appropriate control group (33), not recognizing problems of floor or ceiling effects (31), confusing statistical and biological significance (30), inappropriately assuming that the inability to reject the null hypothesis is equivalent to taking the null hypothesis as true (12, 28, 78), not having a long enough study run for acclimatization of subjects with study conditions (32), using the wrong number of degrees of freedom (46), using the wrong F test (100), and using the subjects as their own control and administering the treatment to all subjects at the second time point (53). For balance, occasionally an example is correct, and Zolman says the conclusions are appropriate (64, 90, 98).

This book has some good features. There is much to be learned from the research design problems. Also, the avoidance of computational formulas allows the reader to focus on concepts and avoid irrelevant computations, which are best done by the computer. Further, the figures, tables, and flow charts in the first part of the book are helpful and great for reference. Even the restriction to no more than three-factor experiments is useful. Beyond this, the experimental results are hard to interpret, especially the interaction effects. The intended audience of experimental biologists will gain much from this book.

Nonetheless a number of problems mar the book. The research design problems have no structure to them. They are not organized according to types of errors and are not summarized by error types. Thus the reader or teacher attempting to find certain types of errors must read the full collection. Next, the appropriate role of nonparametric techniques in experimental design and analysis is lacking. The basic nonexistence of these for designs with more than two treatments and more than two factors is never clearly presented. The treatment of nonparametric procedures is neither complete nor modern. Similarly, a modern treatment of multiple comparisons is missing. Also, random effects models and multivariate analysis techniques are not discussed. Finally, techniques such as logistic regression for dichotomous dependent variables are not mentioned. Unfortunately, because of these omissions, the statistically sophisticated student or researcher will find *Statistics: Experimental Design and Statistical Inference* incomplete and will be disappointed.

Ralph B. D'AGOSTINO
Boston University

Essentials of Biostatistics (2nd ed.)

Robert C. ELSTON and William D. JOHNSON. Philadelphia: F. A. Davis, 1994. xii + 328 pp. $31.95.

The teacher of biostatistics in the health sciences is often faced with a major challenge, mainly related to student dispositions rather than to sound pedagogy. Some students enter the course with a deep fear of mathematics, others with the opinion that there are more important things to do, and still others with the belief that they will never use this material. All of these students beg for the diminution or elimination of rigor, and often the teacher's response is to avoid the sophisticated concepts and computational components and focus on a survey of "basic and fundamental concepts, without getting bogged down in computational details." *Essentials of Biostatistics* is in this latter tradition. Its intended audience is medical and nursing students, and its goal is to explain basic concepts. Its mathematical prerequisite is high school mathematics, to be read as first-year algebra. The book is in its second edition following a successful first edition, demonstrating that there is an audience for it.

The book has 12 chapters:

1. Introduction;
2. Populations, Samples, and Study Design;
3. Descriptive Statistics;
4. Probability;
5. Random Variables;
6. Estimation and Confidence Intervals;
7. Significance Tests and Tests of Hypotheses;
8. Chi-Square Tests;
9. Correlation and Regression;
10. Analysis of Variance;
11. Specialized Techniques (such as multivariate analysis, logistic regression and survival analysis) and
12. Guides to a Critical Evaluation of Published Reports.

Each chapter begins with a list of "Key Concepts," followed by definitions, discussion, and simple numerical examples. Each chapter ends with a detailed summary, a list of further readings, and a set of multiple choice problems in the National Board examination format. Answers to these are given at the end of the book. Also included are three Appendixes: 1, Computational Formulas; 2, Statistical Tables; and 3, Glossary of Symbols and Abbreviations. Finally, there is an extensive, useful index.

On the positive side, the book covers a broad expanse of material. Each chapter is loaded with vocabulary and definitions and contains pertinent numerical examples. All of these come fast on the reader and often without substantial motivation. But the writing is usually clear, and the basic concepts are there and correctly presented. The addition of more examples and discussion in the context of a course should successfully acquaint the health scientist with essential statistical concepts. Further, the book does not talk down to the reader; neither does it make mistakes in trying to be popular. Many students will find it to their liking.

Nonetheless, the book is not without faults. Its development of the mathematical (axiomatic) definition of probability seems useless, and its discussion of Bayes's theorem is confused. Its avoidance of the central limit theorem and lack of a detailed discussion of sampling distributions leave the reader with a substantial deficit in the knowledge of major statistical concepts. Some chapters, such as Chapter 8 on the chi-squared tests, are almost fully computational and fail to avoid getting "bogged down in computational details." Chapter 11, on some specialized topics, has the potential of being incomprehensible. The logit transformation, life tables, survival curves, and the idea of a hazard function need more than a couple of pages.

The existence of a market for this type of book is undeniable. I argue that we should ask more of the students. The book by Pagano and Gauvreau (1993), which has the same topics as this book but asks the student to face computations to better understand the concepts and methods and even supplies a data diskette, is more to my liking. But if the softer survey approach is desired, then Elston and Johnson have supplied a reasonable alternative.

Ralph B. D'AGOSTINO
Boston University

REFERENCE

Pagano, M., and Gauvreau, K. (1993), *Principles of Biostatistics,* Belmont, CA: Duxbury Press.

Efficient and Adaptive Estimation for Semiparametric Models

P. J. BICKEL, C. A. J. KLAASSEN, Y. RITOV, and J. A. WELLNER. Baltimore: The Johns Hopkins University Press, 1993. xiv + 560 pp. $95.

Semiparametric models have recently become a popular choice in statistical modeling. They allow the statistician to model some features of the data parametrically, while treating other features nonparametrically. Consequently, semiparametric models are more realistic and flexible than parametric models and provide more structure than purely nonparametric models. The increased interest in semiparametric models has coincided with the advent of ever-faster computers that now make it possible to carry out the elaborate calculations needed to fit and compare these more complicated models. On the theoretical front, an elegant asymptotic theory has emerged that aids understanding of these models. This asymptotic theory is the work of many researchers. It paints a clear picture of how well smooth characteristics of semiparametric models can be estimated. The key results are abstract convolution theorems and lower bounds on the local asymptotic minimax risk based on the Hájek–Le Cam theory for locally asymptotically normal families.

This monograph treats aspects of this asymptotic estimation theory in the simplest case: models with independent and identically distributed observations. Though some generality is lost by this restriction, much is gained by the clarity of the mathematical development and the resulting simple geometric interpretation of the results. Moreover, a thorough grasp of the theory in this context should be helpful in understanding extensions to more complicated models.

The monograph considers the class of regular estimators and uses various forms of the convolution theorem to characterize least-dispersed regular estimators. It does not pursue the approach via the lower bounds on the local asymptotic minimax risk, which leads to the efficiency concept of a locally asymptotically minimax estimator. The latter concept is a true optimality criterion, as it considers the class of *all* estimators. The notion of a least-dispersed regular estimator is an optimality criterion only within the class of regular estimators. But least-dispersed regular estimators are locally asymptotically minimax for bounded loss functions.

The classical convolution theorem for estimating differentiable functionals in parametric models states that (a) the limiting distribution of a regular estimator is a convolution of a mean zero normal distribution with some other distribution, and (b) a regular estimator with limiting distribution equal to this normal distribution possesses a prescribed influence function, the so-called efficient influence function. These conclusions carry over to estimating smooth (finite- and infinite-dimensional) functionals in infinite-dimensional models. Thus the characterization of least-dispersed regular estimators amounts to describing the efficient influence function and its variance, the so-called information bound.

The efficient influence function has a simple geometric interpretation in terms of certain projections in Hilbert spaces. But explicit calculations depend on the ease with which these projections can be calculated. Thus the monograph places great emphasis on the implementation of the theory in concrete cases. Numerous examples clarify the various aspects of the theory, illuminate the difficulties in applying this theory, and give guidance to tools needed to overcome some of the technical difficulties. Two full chapters, totaling 170 pages, are entirely dedicated to examples, including group models, regression models, biased sampling models, mixture models, missing and incomplete data models, and transformation models. Throughout the monograph, the authors do an excellent job of illustrating the theory with interesting and informative examples.

The latter part of the monograph is dedicated to the construction of estimators in infinite-dimensional models. Although the characterization of efficient estimators is well understood, many open questions remain regarding their construction. The existence of the efficient influence function alone does not guarantee their existence. There is, however, a general theory for the construction of efficient estimators of the finite-dimensional parameters in semiparametric models. In this case, efficient estimators can be constructed if and only if a root-n-consistent estimator of the finite-dimensional parameter and an appropriate estimator of the efficient influence function are available. The underlying construction is not quite satisfactory, as it relies on a sample-splitting trick. But under stronger conditions on the estimate of the score function, this trick can be avoided. This theory and some simple applications are treated at the end of Chapter 7. I would have liked to see these results earlier, maybe in Chapter 3.

The book is organized as follows. Chapter 1 gives an overview. Chapter 2 describes the asymptotic theory for estimation of differentiable functions of the parameter in regular parametric models; these are continuously Hellinger differentiable models with nonsingular information matrices. These strong smoothness assumptions on the model allow the study of what the monograph calls uniformly regular estimators. It characterizes least-dispersed uniformly regular estimators via the convolution theorem and shows that such estimators can be constructed under mild additional assumptions. This chapter also prepares the reader for the geometry of estimating a parameter in the presence of a nuisance parameter. It introduces the question of adaptation: When can a parameter be estimated as well, asymptotically, when the nuisance parameter is unknown as when it is known?) and gives necessary conditions for adaptation.

Chapter 3 discusses estimation of finite-dimensional parameters in infinite-dimensional models. It introduces the important notion of a tangent space for infinite-dimensional models, then presents two approaches for obtaining information bounds for regular estimates and characterizing least-dispersed regular estimates in such models. The first approach, nonparametric in nature, shows that the influence function of a least-dispersed regular estimator of a Hadamard differentiable functional is the canonical gradient of this functional. The second approach, via the scores and score operators, obtains the efficient influence function as the product of the inverse of the efficient information matrix and the efficient score function. The efficient score function is the componentwise projection of the score for the parameter of interest onto the orthogonal complement of the tangent space for the nuisance parameter; the efficient information is its covariance matrix.

Chapter 4 illustrates this theory with a wide selection of examples. Here explicit formulas for tangent spaces and efficient influence functions are obtained.

Chapter 5 discusses information bounds for infinite-dimensional parameters such as distribution functions and cumulative hazard functions. To this end, convolution theorems for Hadamard differentiable Banach space-valued functionals are formulated. To treat nonseparable Banach spaces, regularity of estimators is defined via the modern Hoffmann-Jørgensen and Dudley weak convergence theory. This modern weak convergence theory parallels the classical weak convergence theory for Polish spaces, but works with outer expectations. Thus estimates no longer have to be Borel measurable. This approach is new. Previously, results have been formulated for separable Banach spaces utilizing the more familiar weak convergence theory for Polish spaces. The chapter also addresses results on the connection between regularity and differentiability and gives sufficient conditions for differentiable functionals of the parameters to define differentiable functionals of the models. The latter is very important, as many characteristics of models are defined as functionals of the underlying parameters.

Chapter 6 treats examples. It parallels Chapter 4, but with emphasis on estimating the infinite-dimensional parameter.

Chapter 7 deals with the construction of (efficient and other) estimators of finite- and infinite-dimensional parameters in semiparametric models. The early focus is on generalized minimum contrast estimators, generalized M estimators, and related estimators such as nonparametric maximum likelihood estimators. Uniqueness, consistency, and rates of convergence of these estimators are discussed. These estimators are of interest in themselves and can serve as preliminary estimates in the construction of efficient estimates. The chapter concludes with the aforementioned theory of constructing efficient estimators of the finite-dimensional parameter in semiparametric models.

The monograph closes with an extensive and very helpful appendix, in which the authors collect the necessary background material from functional analysis and operator theory, provide important projection formulas, survey the classical and modern weak convergence theory, and review contiguity and M-estimation theory. Noteworthy is a very nice treatment of projection theory for sum spaces.

The monograph requires familiarity with the basic concepts of functional analysis. A potential reader is well advised to browse through the appendix before reading the text. A solid knowledge of the relevant tools from functional analysis and of the modern weak convergence theory are required to tackle the more abstract Chapters 5 and 6. The first four chapters and parts of Chapter 7 are accessible without this knowledge, however.

Efficient and Adaptive Estimation for Semiparametric Models makes an important contribution to modern mathematical statistics. The authors have done a fine job making this difficult area accessible for a broader audience. They have achieved this by including an abundance and variety of illustrative and important examples. This monograph is required reading for anyone interested in semiparametric models.

Anton SCHICK
SUNY—Binghamton

Measure Theory

J. L. DOOB. New York: Springer-Verlag, 1994. xii + 210 pp. $49.

This splendid work, which offers a general and unified treatment of the theory of measure and integration over abstract spaces from the author's own perspective, is suitable as a text for a graduate course or as a reference. Courses in topology and real analysis would provide the necessary background material. However, exercises are not included. Also excluded is a list of references, which could be helpful to a reader interested in delving deeper into a topic or wishing to be exposed to alternative approaches for the development of the subject.

This book is a treatise on measure theory rather than on probability per se. Nevertheless, significant probability concepts and results are presented and at the rightful time, rather than collected for inclusion in a separate chapter (as was done in Halmos 1974, for example). It is an exceptionally unfortunate and sad historical reality that measure theory was developed