Classification Algorithms, by Mike James, New York: Wiley-Interscience, 1985, xii + 211 pp., $34.95.

The author of this book uses the term "classification analysis" to refer to what is more commonly known as discriminant analysis, the assignment of new observations or objects to existing, well-defined categories. This is an unfortunate choice of terminology, because it is also widely used to describe cluster analysis, the identification of natural groupings in a currently unstructured collection of objects. Cluster analysis is not discussed in this book, however, except to distinguish it from classification analysis.

This book provides a nice overview of classification analysis pitched at a reasonable mathematical level for the reader familiar with the basics of probability, statistics, and matrix algebra. (The appendix contains a light review of the necessary matrix algebra.) The author presents the material in a highly readable fashion with intuitive explanations of concepts, often appealing to geometry, used throughout the text. Sections involving more mathematical detail are marked appropriately and can be skipped without loss by the uninterested reader. Two short pages of mathematical notation used are provided near the front of the book for easy reference.

A useful feature is the collection of BASIC programs integrated into the text for implementing the classification methods described. After each technique or extension is discussed, a BASIC program or a modification of a previous program is given. Complete details of the programs are provided, which may add to the reader's understanding of the computational aspects of the techniques. The appendix contains a BASIC program for generating test data as practice input to the classification programs.

A few deficiencies keep this book from being a more effective introduction to classification techniques. The author fails to provide sufficient references to other sources for the reader to pursue topics of interest. The book would also benefit from the inclusion of more interesting examples using real data from different areas of application. The three examples given are used primarily to demonstrate the BASIC programs and provide the user of the programs with a check for coding errors. Finally, although the author chooses not to include clustering under classification analysis, the role of cluster analysis in classification certainly deserves more attention than the one or two paragraphs and the single reference given in the first chapter.

A brief description of the topics covered in each chapter follows. Chapter 1 provides an overview of the book and discusses the differences among the terms "classification analysis," "discriminant analysis," "cluster analysis," and "analysis of variance" used to describe statistical techniques related to classification problems. As stated earlier, this book deals exclusively with classification analysis, also known as discriminant analysis.

Chapter 2 discusses the total error of classification (TEC) criterion and Bayes's rule. The Bayes classification rule under a multivariate normal assumption is developed in Chapter 3 for both equal and unequal covariance matrices. Chapter 4 provides the initial set of BASIC programs for linear and quadratic discriminant analysis, along with a discussion of the workings of the programs. As extensions of these techniques are discussed, the necessary modifications to the programs are given.

Chapter 5 addresses practical issues associated with the use of discriminant analysis, including bias caused by using sample estimates for population parameters in the classification rule, robustness of discriminant functions to nonnormality, the use of criteria other than TEC (including minimum cost, minimax error, and fixed error rate), allowing for nonclassification of borderline observations, and missing or incomplete information. Chapter 6 discusses various methods of estimating the error rates of classification for evaluating the usefulness of classification rules.

Chapter 7 explains the use of canonical analysis to reduce the number of variables considered in a classification problem to a more reasonable number. Chapter 8 is a more useful chapter that briefly describes variable-selection techniques for choosing from a large number of possible variables. BASIC programs are provided for canonical analysis and stepwise discriminant analysis.

Chapter 9 contains brief sketches of classification with categorical data and of nonparametric techniques, such as nearest neighbor, for classifying data into defined categories. Finally, Chapter 10 is a broad discussion of classification problems in the areas of artificial intelligence and pattern recognition.

J. Brian Gray
Texas Christian University

Empirical Processes With Applications to Statistics, by Galen R. Shorack and Jon A. Wellner, New York: John Wiley, 1986, xxxvii + 938 pp., $59.95.

Readers of *Technometrics* know that general concepts can be applied successfully to solve practical problems if concept and problem are somewhat equivalent, and that uncertainty about the degree and extent of this equivalence is an integral part of the excitement and challenge of practical applications. Readers also know that theoretical concepts can have exciting applications to other theoretical concepts and that the equivalence of the concepts involved, once recognized, is often complex but well defined.

This long and clearly written book is about the latter, theoretical kind of application. The preface states that the book "is intended for graduate students and research workers in statistics and probability. The prerequisite is a standard graduate course in probability and some exposure to nonparametric statistics" (p. x).

The uniform empirical process $U_n(t)$ is defined as $\sqrt{n}[G_n(t) - t]$, where $G_n(t)$ is the familiar empirical distribution function for a sample of n independent random variables uniformly distributed on the closed interval [0, 1]. The book treats $U_n(t)$ as a stochastic process on [0, 1], relating it to Brownian bridges and other processes. The applications to statistics are theoretical and oriented to areas of most interest to probabilists—for example, Kolmogorov–Smirnov tests and L statistics. Results for the uniform empirical process can easily be modified for sampling from any distribution; in some cases no modification is needed. It is interesting, therefore, to speculate that empirical processes tend to be informative about various robust or distribution-free statistical procedures; they tend not to provide useful results about traditional, generally less robust procedures.

The authors have worked hard to organize the book for easy reference. The detailed table of contents is 18 pages long; Chapter 1 is both introductory and a survey of some results, but not a comprehensive summary; the first of two appendixes contains " many of the most useful inequalities in probability theory" (p. 842). The publisher has made it difficult for the reader to follow cross-references by omitting chapter numbers and section numbers from running headings.

Vincent Hodgson
Amoco Corporation

Potential Pattern Recognition in Chemical and Medical Decision Making, by D. Coomans and I. Broechaert, Letchworth, U.K.: Research Studies Press, 1986, xiii + 256 pp., $59.95.

The potential-function method of discriminant analysis is a nonparametric technique based on kernel-density estimation that has its roots in the work of Rosenblatt (1956) and Parzen (1962). This