and occasional discrepancy from becoming anything more than a minor hindrance.

Chapter 1, "Getting Started," presents the use of data records, run charts, cause-and-effect diagrams, histograms, stratification, and scatter diagrams. These methods are introduced and illustrated through a door-slab-effort example. This method of presentation is not only clear, but it also defines a process by which to apply the various tools and would enable beginners to apply the methods to their work.

Chapter 2, "Making Sense of Data," introduces frequency tables, histograms, measures of location and spread, the normal distribution, sample versus population data, and the meaning of significance tests and confidence intervals. Because references to population parameters occur before the section on sample versus population data, those with no statistical background may find the concepts a bit confusing. The author provides intuitive definitions for many of the concepts however, which should compensate for any confusion over notation. A unique and useful feature about this chapter is the helpful hints and words of warning for interpreting the graphical presentations and statistical methods.

Chapter 3, "Picture Your Numbers," presents more graphical methods such as dot plots, boxplots, quartile plots, histograms, normal plots, run charts, multi-var charts, and scatterplots. Similar to previous chapters, the methods are demonstrated through examples, and valuable information to assist the user with interpretation is included.

Chapter 4, "Why Design Experiments," effectively illustrates the disadvantages of testing one factor at a time and provides brief discussions on the importance of replication, reliable measurement, randomization, and blocking. The remainder of the chapter focuses on the analysis of one- and two-factor designs and may be helpful as a reference for those already familiar with the analysis of these designs. Because the authors have chosen to make the discussion of the underlying analysis theory brief, however, those lacking previous knowledge may find it difficult to perform a complete analysis without consulting further references.

Chapter 5, "Understanding SPC Charts," includes information on constructing and interpreting control charts for means, ranges, individual readings, percentage of nonconforming items, cumulative sum charts, and process capability indexes. The presentation, however, is of British-style control charts, which contain both action and warning lines. Even though the action lines roughly correspond to the control lines given on American-style control charts, the differences will present difficulties for those attempting to construct American-style charts.

Overall, the book provides concise and useful summaries for many statistical tools. Through useful examples, it provides an ample background for a beginning to apply many of the elementary tools. Partly due to the nature of the topic, however, it may be difficult for those without prior statistical knowledge to directly apply the experimental design and SPC techniques in latter chapters. Consequently, this book is best as a quick pocket reference for the beginner who needs to implement some simple graphical and statistical methods.

Marecy Abate
Sandia National Laboratories


The book is well written and covers a wide range of statistical methodologies including pattern-recognition techniques such as the principal-component analysis (PCA), cluster, classification and discrimination analyses, neural networks, and projection pursuit, geostatistical method of kriging, and numerous other univariate/multivariate parametric and nonparametric methods. In the preface, the authors stated that the book was aimed at professionals and graduate students from each of the two fields, statistics and astronomy, who are interested, but not especially trained, in the other field. The authors failed, however, to include examples using real astronomical datasets and missed greatly on showing the practical merit and applications of the statistical methods cited in the book. This limits the usage of the book to only research-oriented graduate students. A typical user in astronomy will find little help from the book without consulting a statistician.

Chemometrists have found many useful applications of regression, principal component, discriminant, and cluster analyses in their various calibration and pattern-recognition applications. No doubt, like chemometrics, the area of astrostatistics can also benefit from routine use of the various multivariate statistical methods. The book would have been more useful for both audiences, astronomers and statisticians, had the authors included some case studies using the various methods covered in the book. For example, the kriging methodology is briefly described in Chapter 4. As a practicing statistician and a frequent user of kriging in environmental site-characterization applications, I was looking forward to seeing an example showing the applicability of kriging on a real astronomical dataset. Its mere mention in the book does not convince me enough that kriging can be successfully used in astronomical applications because kriging does depend on certain assumptions (Cressie 1993) that may not be easily satisfied by the datasets obtained from the astronomical applications.

The authors have put a lot of emphasis on including their own research work in the area of mathematical statistics (e.g., Chap. 5) at the expense of some important and relevant topics such as the outlier identification and robust procedures. Just like the various chemometric and environmental applications, outliers are inevitable in datasets originated from astronomical applications, a fact well recognized by the authors because the words "outlier" and "robust" are mentioned several times in the book. No proper treatment is included for the identification of outliers, however.

Most practitioners do understand what the outlying observations are; what they do not know is how to identify them and subsequently treat them. Outliers when present distort all classical estimates such as the maximum likelihood estimates (MLEs) including the means, variances, covariances, the principal components, and the ordinary least squares (OLS) estimates. In Chapter 4, the authors briefly compare the OLS regression and the minimum absolute deviation regression. Some real datasets would have been very useful to shed light on these statistical issues. One simple example from astronomy widely used in the statistical literature (Rousseauw and Leroy 1987; Singh and Nocerino 1995) is the dataset consisting of 47 stars in the direction of Cygnus (Hertzsprung–Russell Diagram of Star Cluster CYG OB1). This dataset is known to consist of 43 stars following the main sequence and 4 giant stars (outliers) from a different population.

The objective is to obtain a calibration model that describes the main sequence of stars adequately. This example shows the users how the outliers can distort the classical estimates, MLE and OLS, which have a "zero" breakdown point. It is important to let the users know that there are robust and resistant procedures that are not that sensitive to the outlying observations. The authors also mention that the PCA (and, of course, the discriminant functions) are sensitive to outliers. A typical user would have appreciated the book more if some of the widely available robust regression and PCA techniques (Devlin, Gnanadesikan, and Kettenring 1981; Rousseauw and Leroy 1987; Rousseauw and van Zomeren 1990; Singh and Nocerino 1995) were also included in the book.

The entire Chapter 5 is devoted to the various resampling techniques. This is a very good chapter for graduate students in mathematical statistics pursuing careers in the areas described. In this chapter, the researchers areas of the authors have been overemphasized (e.g., Secs. 5.2, 5.3 and 5.4). For example, the authors start talking about outliers and breakdown point and robust estimator. The bootstrap estimator of the variance of the median and its breakdown point are discussed in some details, whereas nothing is mentioned about the robustness and breakdown points of the more commonly used estimates based on the maximum likelihood procedure and the OLS regression. The information described in this chapter is incomplete and could be misleading to a typical user with limited statistical background. In Chapter 7, several linear regression methods (Fig. 7.2) are discussed. It would have been very useful if the authors briefly discussed the influence of outliers on regression models. This actually may be annoying to statisticians because they know that all classical MLE and OLS regression estimates are very sensitive to outliers and have a "zero" breakdown point, which should have been clearly stated in the book. A chapter devoted to these issues of breakdown point, identification of outliers, and robust estimation would have made the book more useful and complete.

In conclusion, Babu and Feigelson have written a good book for graduate students pursuing research-oriented careers in mathematical statistics. An astronomer with little statistical background (as mentioned in the pref-
ace), however, will not be able to use the statistical methods cited in the book without consulting expert statisticians. The book does not provide guidelines to a nonstatistically oriented practitioner in the area of astronomy. Moreover, in the absence of case studies and real datasets, there is not much in the book for a statistically oriented professional because most statistical professionals and academicians would know about the statistical methods listed in the book. Personally, I would have appreciated the book more if, instead of just listing the various statistical methods, the authors had included a few case studies with real data showing how statistical methods can be used in various astronomical applications.

There seem to be a few typing errors; for instance, "nonlinear" on page 136 second paragraph last sentence should be "linear."

Anita SINGH
Lockheed-Martin

REFERENCES


This book will probably hold little interest for readers of *Technometrics*. In the preface the authors indicate that the book grew out of Stat 1: Statistical Thinking, one of a series of courses taught at Swarthmore College to fulfill the mission of a liberal arts college. This book obviously is intended for a very nontechnical audience. Having said that, however, it is also obvious that this book does a very good job of presenting statistics to a very nontechnical audience.

The book has several unique characteristics that make it particularly valuable in a course that is intended to teach statistical literacy. For example, formulas are all collected in special sections in each chapter; exercises at the end of each chapter are extensive and broken down into three categories: Review, interpretation, and analysis and some exercises direct the reader to a website for data; the text material is interspersed with pictures, cartoons (there are several of these), and highlighted boxes called "Stop and Ponder" that instruct the reader to think about the topic in terms of the reader’s environment; and highlighted summaries are placed strategically in the margins.

The book consists of 14 chapters covering the basic topics found in an introductory statistics textbook. The first four briefly cover sampling and descriptive statistics. The fifth chapter covers probability and mentions the basic probability distributions. As an example of how the book presents topics, the Poisson distribution is illustrated in Chapter 5 using the number of baseball no-hitters occurring in one day as the random variable. Included in the discussion of the illustration is a classic photograph of Cy Young. The next eight chapters cover estimation and hypothesis testing, including correlation, analysis of variance (ANOVA), contingency tables, and regression. The last chapter is entitled "Statistics for Everyday Life" and includes some misuses of statistics. The book ends with the line "Stat’s all folks."

The first 10 chapters of this book cover material appropriate for a single-term introductory course in statistics aimed at a nontechnical audience. After Chapter 10 the text seems to run out of steam. The topics of ANOVA, rank statistics, multivariate regression, and two-way ANOVA are covered in a perfunctory manner. The authors try to cram a wide range of rather complex topics into the last hundred pages of the book.

The book is well written, contains much anecdotal information, has a wide variety of exercises, and is cleverly illustrated. It is very easy to read and would be an excellent textbook for use in an introductory statistics course for students majoring in the social sciences or humanities.

William J. Wilson
University of North Florida


This book could easily have been titled *Everything You Always Wanted to Compute With Trimmed Means*, but more on this later. The two biggest strengths of *Introduction to Robust Estimation and Hypothesis Testing*, by Rand R. Wilcox, are its practical usability and its inherent structure. The former strength is due to the availability of S-Plus functions on the publisher’s website. Having access to the functions described in the book is an obvious asset. The other side of the coin, however, is that the book is not of much use without S-Plus. The latter strength of the book, its structure, is due to the discussion of standard-error estimation immediately after the presentation of a particular statistic. This structure presents point and interval estimation almost simultaneously—a nice format.

The introductory first chapter starts off by describing some of the pitfalls in assuming an underlying Normal model. Immediately thereafter the subject of data transformation, a popular subject in the literature (e.g., Hoaglin, Mosteller, and Tukey 1983), is dismissed because of problems with potential outliers. Following this, an example of the mean versus the median is given to show the adverse effects of outliers. The chapter finishes with instructions for downloading the functions described in the rest of the book. These functions are made available on the publisher’s website. The instructions did not work as smoothly as indicated in the book, so I will detail the steps that worked for me. I found that the file "allfun," which contains the entire collection of functions, should be saved as a text file in the "c:\spluswin\home" directory, which is the default directory for S-Plus (if your default for S-Plus is something different, then this step needs to be modified). The S-Plus command to access these functions would then be source("allfun.txt"). This worked in a system consisting of Netscape 3.0, Windows 95, and S-Plus 3.3 for Windows.

Chapter 2 is the most mathematically technical and attempts to lay the foundation for robust theory. Because the book is intended as an "introduction" to robust methods and is intended for those with some basic knowledge of analysis of variance and regression, this chapter is weak because it omits the simplest and strongest argument for using robust methods. Although the author does a reasonably good job of explaining equivariance, qualitative and quantitative robustness, the influence function, and breakdown point, he does not talk specifically about robustness of validity and omits robustness of efficiency. Basically, the argument that robust methods resist outliers but still achieve good performance when the model is true is never explicitly stated. The chapter ends with a bivariate Winsorization that appears unrelated to the rest of the chapter.

Chapter 3 serves as a nice bridge between the theoretical material of Chapter 2 and the heart of the book, Chapters 4 through 9. Statistics are presented along with their standard-error estimates, as in the rest of the book, but theoretical concepts like the influence functions and asymptotics are given as well. It starts out with trimmed means, and the author shows his preference for the 20% trimmed mean. It is interesting that the reason given for this preference is that the trimming guards against "complete disaster but sacrifices relatively little" in situations in which the mean or 10% trimmed mean are better, if not optimal. (This argument should have been generalized in the previous chapter!) One could argue that the 20% trimmed mean has only 20% breakdown and that various M-estimators could easily improve on that. The treatment of M-estimators is adequate at best. The connection to maximum likelihood (after all, the "M" in M-estimation) is never made, and redescribing M-estimators of location, like the biweight, are dismissed because of possible estimation problems. Nevertheless, it is the author’s book and he is entitled to his