This section will review those books whose content and level reflect the general editorial policy of *Technometrics*. Publishers should send books for review to Eric R. Ziegel, BP Naperville Complex, Mail Code C-7, 150 West Warrenville Road, Naperville, IL 60563-8460. (ziegeler@bp.com)

The opinions expressed in this section are those of the reviewers. These opinions do not represent positions of the reviewer’s organization and may not reflect those of the sponsoring societies. Listed prices reflect information provided by the publisher and may not be current.

The book purchase programs of the American Society for Quality can provide some of these books at reduced prices for members. For information, contact the American Society for Quality, 1-800-248-1946.

| Title                                                                 | Authors                                                                 | Page |
|----------------------------------------------------------------------|------------------------------------------------------------------------|------|
| Robust Diagnostic Regression Analysis                                | Anthony Atkinson and Marco Riani                                       | 81   |
| Applied Logistic Regression                                          | David W. Hosmer and Stanley Lemeshow                                   | 81   |
| Finite Mixture Models                                                | Geoffrey McLachlan and David Peel                                      | 82   |
| Variance Components Estimation                                       | Poduri S. R. S. Rao                                                    | 82   |
| Fundamentals of Modern Statistical Methods                           | Rand R. Wilcox                                                        | 83   |
| Practical Geostatistics                                              | Simon W. Houlding                                                     | 83   |
| Calibration and Reliability in Groundwater Modelling                 | F. Stauffer, W. Kinzelbach, K. Kolvar, and E. Hoehn (editors)          | 84   |
| Modeling Survival Data                                               | Terry M. Therneau and Patricia M. Grambsch                             | 85   |
| Statistical Analysis of Microstructures in Materials Science         | Joachin Ohser and Frank Mücklich                                       | 86   |
| Analysis of Multivariate Survival Data                               | Philip Hougaard                                                       | 86   |
| Pattern Classification, Second Edition                               | Richard O. Duda, Peter E. Hart, and David G. Stork                    | 87   |
| Wavelet Transforms and Time-Frequency Signal Analysis                | Lokenath Debnath (editor)                                             | 87   |
| Bayesian Inference in Wavelet-Based Models                            | Peter Müller and Brian Vidakovic (editors)                            | 87   |
| Nonlinear and Nonstationary Signal Processing                        | W. J. Fitzgerald, R. L. Smith, A. T. Walden, and P. C. Young (editors) | 88   |
| Statistical Consulting                                              | Janice Derr                                                            | 89   |
| Introduction to Statistics and Data Analysis                         | Roxy Peck, Chris Olsen, and Jay Devore                                 | 90   |
| Running Your Machines with SPC                                        | James C. Abbott                                                       | 90   |
| The Handbook of Applied Acceptance Sampling                          | Kenneth S. Stephens                                                    | 90   |
| Applied Stochastic Modelling                                         | Byron J. T. Morgan                                                     | 91   |
| Numerical Methods of Statistics                                      | John F. Monahan                                                       | 91   |
| Editor Reports on New Editions, Proceedings, Collections, and Other Books |                                                                 |     |
| Statistical Thinking                                                 | Roger Hoerl and Ronald Snee                                           | 92   |
| Trusting Measurement Results in the Chemical and Process Industries  | ASQ Chemical and Process Industries Division Chemical Interest Committee | 93   |
| Applied Statistics and Probability for Engineers                     | Douglas C. Montgomery and George C. Runger                            | 93   |
| Reliability Improvement with Design of Experiments                   | Lloyd W. Condra                                                       | 93   |
| Multivariate Statistical Modelling Based on Generalized Linear Models | Ludwig Fahrmeir and Gerhard Tutz                                      | 94   |
| Data Mining Cookbook                                                 | Olivia Parr Rud                                                        | 94   |
| Genetic Algorithms and Engineering Optimization                       | Mitsuo Gen and Runwei Cheng                                            | 95   |
| Risk Analysis, Second Edition                                        | David Vose                                                             | 95   |
| Biostatistical Methods: The Assessment of Relative Risks             | John M. Lachin                                                        | 95   |

© 2002 American Statistical Association and the American Society for Quality

*Technometrics*, February 2002, Vol. 44, No. 1
Robust Diagnostic Regression Analysis, by Anthony Atkinson and Marco Riani, New York: Springer-Verlag, 2000, ISBN 0-387-95017-6, xiv + 327 pp., $79.95. As a consultant, I am always on the lookout for new books that help me do my job better. I would recommend practitioners of regression, that is, probably most of us, to read and use this book.

Anthony Atkinson and Marco Riani develop a novel methodology for examining the effect each observation has on the fitted regression model. Robust fitting procedures are combined with regression diagnostics, graphics, and a “forward” processing through the observations to provide a new way of identifying influential and/or outlier observations while simultaneously determining the best fitting model. The method is initially introduced for simple linear regression, but individual chapters are devoted to applying the methodology to nonlinear models in general and generalized linear models in particular.

The role of data transformations to normality is also explored. Throughout the book, a large number of fully worked examples provide the reader real insight to the power of the methodology. Theory is kept to a minimum and matrix notation is used throughout.

Chapter 1 presents some regression examples that illustrate the need for a methodology to identify and account for outliers in simple and multiple regression. Chapter 2 starts with an appropriately short introduction to least squares estimation and associated hypothesis testing. The authors next derive many of the more common influence diagnostics and introduce a “mean shift outlier” model wherein a dummy variable is used in the linear model to assess the impact of a specific observation on parameter estimation and model fit. This tool for examining the effect of a single deletion is then placed in the context of a “forward search” algorithm. The forward search algorithm is made up of three steps. The first step addresses the choice of an initial subset of the observations that will be used with a robust estimation procedure (least medians of squares or LMS) to provide an initial model fit and a “good” estimate of the residual error. If the model contains $p$ parameters, the initial subset will be the one subset out of the $n$ potential subsets that minimizes the sum of squared residuals from the LMS fit. If there are too many potential subsets, the choice will be made after examining a large sample of subsets of size $p$. In the second step, the size of the subset is incrementally increased; at each increase the subset with the smallest sum of squared residuals is kept. Typically this only requires one new observation to enter the fitting set, but there may be cases where some observations drop out and are replaced by others not originally in the subset. Note that as the subset size increases, those observations outside the fitting subset look less and less like outliers. What is important about this approach is that it starts with a subset that is assumed to be outlier free, or it contains unmasked outliers that will be replaced as the subset size increases. This second step is repeated until all observations are included in the fitting subset. The third step of the algorithm is monitoring changes in fit statistics, specifically the residual mean square, parameter estimates, associated $t$-statistics, and even diagnostics, such as Cook’s distance, as observations are incrementally added to the fitting subset. Typically the residual mean square will increase as subset size increases, but the increase will be smooth. Outliers, being added at the end of the Step 2 process, will tend to produce dramatic changes in these fit statistics, hence making identification easy. For smaller samples it is even possible to monitor the value of all residuals from the fit of the current model. Large residual values for observations not in the current model are indicative of potential outliers.

Applied Logistic Regression (2nd ed.), by David W. Hosmer and Stanley Lemeshow, New York: Wiley, 2000, ISBN 0-471-35632-8, xii + 373 pp., $84.95. Applied Logistic Regression is an excellent book that balances many objectives well. While of most appeal to readers with a biology or biostatistics slant, it should also satisfy the statistical practitioner who specializes in reliability modeling. All statistical practitioners with an interest in logistic regression can benefit from this book.

While not recommended as the main text for an upper level undergraduate or graduate level course in its subject, it is ideal as a supplemental text or as a stand alone source for one seeking a broad introduction to logistic regression.

The sequence of topics in the book is a natural one for the reader with a background in multiple linear regression: parameter estimation, significance testing, variable selection and transformation, goodness of fit testing, and regression diagnostics.

The authors follow these topics with an example in Chapter 5 that concisely and completely describes how to present the results of a logistic regression to clients at practically any level of statistical sophistication.

The mathematics is at the level accessible to one who has successfully completed a beginning course in linear inference. The authors have a good sense of which topics require a more detailed derivation and which ones to leave to other references. References are liberally and appropriately cited in the main text, and a generous list 13 pages long can be found at the end of the book. Very recent and more historic articles are equally well represented.

The book covers how some major statistical packages, SAS and STATA being the most mentioned, analyze logistic regression data. In some cases where useful algorithms are not a standard part of the common packages, the authors show programming strategies for obtaining the algorithms’ results approximately. Equally important, the book also points out some of the background assumptions that need to be verified before accepting the results of a computer output. The novice is also warned which outputs are especially easy to misinterpret.

The numerous well placed derivations do not overwhelm the book. The statistical layman has much to gain from reading it. In the introduction, the authors explain clearly the difference between logistic and normal theory linear regression. Each step of the model building process is clearly laid out with well chosen examples and smooth transitions to the next step.

The authors surround the analysis with a coherent philosophy of model building: the iterative nature of models, the importance of choosing a reasonable number of variables, an appreciation of the limits of asymptotic diagnostics, the necessity of respecting the knowledge of the client, and the necessity of a good model to meet the rigor of sound statistics and sound theory of the client’s subject.

In the context of logistic regression, the authors point out where the class of $m$-asymptotics is more appropriate for diagnostics than the asymptotics introduced in a beginning inference course. The importance of the number of distinct covariate patterns, as opposed to the number of variables, in assessing potential overfit is also clearly explained.

In the closing part of the book, the authors supplement the basic logistic model with a variety of special purpose models that have appeared in recent decades: one to one matched study, one to $M$ matched study, multinomial logistic, ordinal logistic, and logistic analysis of correlated data. The treatment is concise but rich enough for the newcomer to logistic regression to clearly perceive where they are best used.