Book Reviews

This section will review those books whose content and level reflect the general editorial policy of Technometrics. Publishers should send books for review to Ejaz Ahmed, Department of Mathematics and Sciences, Brock University, St. Catharines, ON L2S 3A1 (dean.fms@brocku.ca).

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 editors or 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 at 1-800-248-1946.

Probability with Applications in Engineering, Science, and Technology
Matthew A. Carlton and Jay L. Devore Morteza Marzjarani 148

Climate Time Series Analysis: Classical Statistical and Bootstrap Methods (2nd ed.)
Manfred Mudelsee Tom Burr 149

Editor Reports on New Editions, Proceedings, Collections, and Other Books

Modern Nonparametric, Robust and Multivariate Methods
Klaus Nordhausen and Sara Taskinen (Eds.) S.E. Ahmed 150

Innovative Statistical Methods for Public Health Data
Ding-Geng (Din) Chen and Jeffrey Wilson (Eds.) S.E. Ahmed 150

Integrating Omics Data
George Tseng, Debashis Ghosh, and Xianghong Jasmine Zhou (Eds.) S.E. Ahmed 150

Selected Papers I of William Feller
Rene L. Schilling, Zoran Vondracek, and Wojbor A. Woyczynski (Eds.) S.E. Ahmed 151

Probability with Applications in Engineering, Science, and Technology, by Matthew A. Carlton and Jay L. Devore, New York: Springer, 2014, xxxii + 776 pp., $64.64 (H), ISBN: 978-1-4939-0395-5.

This is a traditional textbook in probability and statistics suitable for an undergraduate course on the subject for majors as well as nonmajors such as engineering. A student with the calculus level background should be able to comprehend the material with no difficulties. No prior exposure to probability and statistics is needed. The book is divided into eight chapters and four appendices. The last two chapters are more challenging to the students with a weak mathematical background. One strength of the book is a very large number of real life exercises (over 1000) appearing throughout the book. The book is also supported by many real life examples related to the topics covered in the book, which appear right after each topic. Examples also cover a wide range of topics suitable for science and engineering students. The computer codes for simulations and examples in two popular software packages R and Matlab add an interesting flavor to the book. Beginning with Chapter 6, the book turns attention to more technical topics useful to engineers and possibly computer science majors. The book emphasizes simulation heavily.

Chapter 1 covers the concepts of probability including sample space, events, definition of probability, marginal and conditional probability, independence, and simulation of random events.

Chapters 2 and 3 are devoted to random variables and distributions. Discrete random variables and distributions are discussed in Chapter 2. Continuous random variables and distributions are presented in Chapter 3. Again, examples are followed by the computer codes in R and Matlab. Throughout the book and including this chapter, the authors present interesting topics that challenge the students intellectually and which one normally is not expected to see in an introductory textbook. Simulation of continuous random variables, the inverse CDF method, and the Accept–Reject method among others that appear in Chapter 3 along with computer codes are presented skillfully.

Chapter 4 extends the topics covered in Chapters 2 and 3 to the case of two or more variables. Independence, covariance, correlation coefficient, conditional distributions are among the topics covered in this chapter. The Law of Large numbers, central limit theorem, and the law of Total Expectation are also presented here. Again, making the book more attractive not only to statistics majors but also to others, topics such as reliability functions, series and parallel designs, mean time to failures, hazard functions, order statistics and related distributions, simulation of joint probability distributions, and system reliability are presented in this chapter in an orderly fashion.

Chapter 5 covers several statistical inferences including point estimation, maximum likelihood estimation (MLE), confidence intervals, and testing hypotheses for population mean and population proportion. The chapter also presents large sample confidence intervals, p-value, and power of a test. The authors include computer codes for calculating confidence intervals and testing hypotheses for population mean and population proportion in R and Matlab. The last topic in this chapter is devoted to Bayesian inferences.

Beginning with Chapter 6, the book gradually gives a flavor more suitable to engineers. This chapter discusses Markov chains including its property, transition matrix, and the Chapman–Kolmogrov equations, regular Markov chains and the steady-state theorem, irreducible and periodic chains, Markov chains with absorbing states, and simulation of Markov chains.

Chapter 7 is geared again toward engineers and presents statistical properties of random processes, classification of processes, random process regarded as random variable, autocorrelation, and autcovariance functions. Definitions are extended to the case of two random processes. Also, the relationship between random process, exponential, and gamma distributions are presented in this chapter. Furthermore, different types of random processes including Gaussian processes, Brownian motion are introduced here. Continuous–time Markov processes plus a few related topics appear at the end of this chapter. As mentioned before, this and the next chapter have more applications in engineering than statistics.

The last chapter focuses on more advanced technical topics suitable for engineers and requires familiarity with nonrandom signals and filters or LTI (linear, time-interval) systems. Additional topics include signal processing, power spectral density, power in a frequency band, white noise processes, random processes, and statistical properties of LTI systems. This chapter is geared toward more advanced readers who have prior exposure to Fourier transform.

Appendix A contains some statistical tables needed for the examples and exercises appearing throughout the book. Appendix B covers some mathematics background including Fourier transforms and Appendix C demonstrates some important probability distribution. Finally, Appendix D lists answers to odd numbered exercises appearing throughout the book.

There is no doubt that this is a well-organized and well-written textbook. Many real life examples and exercises add a great flavor to the book. One area where the book can be improved is the addition of three topics one normally would expect to see in a traditional introductory statistics textbook. The three missing topics are: analysis of variance (ANOVA), regression (linear and multiple), and nonparametric statistics. All three topics are obviously of great
interest to science and engineering students. The book already consists of over 750 pages. Additional topics obviously will make the book heavier. One possible approach is to divide the book into two volumes where the second volume would include Chapters 6, 7, and 8 plus the new three topics. Also, some topics from Chapters 3, 4 (especially those covered at the end of Chapter 4, e.g.), and 5 could also be moved to the second volume. The additions would make this book an excellent source for a two-semester long course. Needless to say that this book in its present format could also be used in a one-semester course with topics selected at the discretion of the instructor. For a one semester course in probability and statistics, Chapters 1 through 6 should be sufficient. However, due to the missing of the three topics mentioned above, the instructor should supplement the course with these topics. Instructors in science and engineering who are looking for a good textbook for their statistics courses should check this book out first. Speaking of the scope and audience of the book the authors mention in the preface and I quote "... is accessible to a wide audience including mathematics and statistics majors (yes, there are a few of the latter, and their numbers are growing.) According to the recent posting of the American Statistical Association, statistics is one of the fastest-growing degrees in the United States, but the growth may not be enough to satisfy the high demand for statisticians in technology, consumer products, health care, government, manufacturing, and other areas of the economy (ASA News 2015). I believe the comment made in the preface of the book needs to be revised as the number of students seeking degrees in statistics has been satisfactory in the past and is growing.

Morteza MARZJARANI
Saginaw Valley State University (retired)

REFERENCE
ASA News (2015, October 1), "More Students Earning Statistics Degree, But not Enough to Meet Surging Demand for Statisticians," available at https://www.amstat.org/newsroom/pressreleases/2015-StatsDegreeGrowth.pdf. [149]

Climate Time Series Analysis: Classical Statistical and Bootstrap Methods (2nd ed.), by Manfred MUDELSEE. New York: Springer, 2014, xxxii + 454 pp., $229.00 (H), ISBN: 978-3-319-04449-1.

This book provides a good and comprehensive review of climate time series analysis. The URL www.manfredmudelsee.com/book has links to data and software. The target audience is applied statisticians and climatologists.

Many of the climate time series have unequal time spacings unlike much of classical time series analysis, and the bootstrap is heavily used for uncertainty assessments, with attention to preserving shape and serial dependence in the bootstrap samples (although open questions remain). A practical simulation approach is used to assess the quality of bootstrap confidence intervals (CIs), comparing nominal to actual CI coverage.

The main equation (introduced in Chapter 1) assumed for a climate variable \( X(t) \) is \( X(t) = X_{\text{out}}(t) + X_{\text{noise}}(t) + S(t) \), where \( t \) is continuous time (there is a similar version for discrete time), \( X_{\text{out}}(t) \) is the trend process, \( X_{\text{out}}(t) \) is the outlier process, and \( S(t) \) is a variability functioning scaling \( X_{\text{noise}}(t) \). The noise process is assumed to be weakly stationary (not defined) with zero mean and zero autocorrelation.

Part I (fundamental concepts in Chapters 1–3) is as follows.

Chapter 1 includes a helpful introduction to climate data archives, variables, and dating methods, plus the notion of persistence (e.g., the auto-regressive lag one model, AR(1), is persistent) and time series spacings in typical climate variables such as ice core data. The use of proxy variables (the variables that are actually measured rather than the variables of direct interest, e.g., lake sediment bed thickness is a proxy for wind speeds) is nicely explained and also partly summarized in Table 1.1. Chapter 8 describes noise introduced by using proxy variables.

Chapter 2 describes the notion of persistence in more detail, using AR(2) and mixed auto-regressive moving average models (ARMA). Persistence is also known as memory in correlated data; persistence refers to positive serial dependence in a time series.

Chapter 3 covers CI construction using a bootstrap technique adapted for serially correlated data. One option is a blocked bootstrap that resamples blocks of data. Another option is a parametric bootstrap that has uses a model with serial dependence such as an AR(1) model. The notion of nominal versus actual bootstrap CI coverage is described.

Part II is univariate time series in Chapters 4–6 as follows.

Chapter 4 is regression I applied to \( X(t) = X_{\text{trend}}(t) + S(t) \). Weighted, ordinary, and generalized least squares (this should not be confused with general linear models; here, generalized least squares is defined as it usually is; it allows an arbitrary covariance matrix for the residuals) are described. Section 4.2, Nonlinear Regression, allows for polynomial terms, so, for example, \( X(t) = \beta_0 + \beta_1 T(t) + \beta_2 T(t)^2 + S(t) \), which is of course a linear model; however, models that are nonlinear in the parameters are also briefly described. Another type of nonlinear regression is a multiple-region regression such as piecewise linear. Speciality topics such as quantile regression, data mining tools such as neutral networks, errors-in-variables models, and smoothing are also briefly described.

Chapter 5 is spectral analysis, admirably presented in a user-friendly manner for the intended audience. Time series analysis in the frequency domain tends to be more technical than time series analysis in the time domain, but the author does a very nice job of presenting the appropriate level of technical detail.

Chapter 6 is extreme value time series. The nonstationary generalized extreme value distribution is clearly presented as a type of "central limit theorem" for extreme values. The bootstrap is used for constructing confidence limits. I am always cautious about not over-interpreting confidence limits in this type of context. The limits are dependent on modeling assumptions (as the author states in Section 6.2.3), so confidence limit width almost never includes all key sources of uncertainty, even so-called nonparametric limits based on a nonparametric bootstrap.

Part II is bivariate time series in Chapters 7–8 as follows.

Chapter 7 covers correlation and in correlation evaluation, the trend process \( X_{\text{trend}}(t) \) in \( X(t) = X_{\text{trend}}(t) + S(t) \) is very important. Any trend in two time series will lead to nonzero correlation unless the trend is removed. Any remaining nonzero correlation is more meaningful after a de-trending; Chapter 7 assumes a constant mean \( \mu_S \) so \( X_{\text{trend}}(t) \) is a slight mismatch here.

Chapter 8 is regression part II to relate two climate variables, \( X(t) \) and \( Y(t) \). I am pleased to report that the author uses the expression "errors in predictors" rather than the commonly used and less informative expression, "errors in variables." Readers familiar with the errors in predictors literature will appreciate the simple but effective treatment.

Part IV is outlook and future directions in Chapter 9.

Chapter 9 on future directions includes: timescale modeling, novel estimation challenges, higher dimensional problems with many variables, climate models, and optimal estimation.

I think this text provides very good coverage of well-selected topics, both from the applied statistician's perspective (myself) and from the layman-in-climate-data (myself) perspectives. I suspect the coverage will please climatologists also (among the intended audience).

Tom BURR
Vienna International Centre

Editor Reports on New Editions, Proceedings, Collections, and Other Books

Modern Nonparametric, Robust and Multivariate Methods, by Klaus NORDHAUSEN and Sara TASKINEN (eds.). New York: Springer, 2015, xx + 506 pp., $149.00 (H), ISBN: 978-3-319-22403-9.

This Festschrift honors Hannu Oja for his important contributions to the statistical community, and the at large scientific community as well. The Festschrift is divided into four parts; the first part (two chapters) is dedicated to Oja’s career and publications. The chapters in Part I essentially showcase the useful contribution of Oja’s respective authors present and assess the importance of the work and shed some light on his life. The remaining parts are:

Part II: Univariate nonparametric and robust methods (six chapters);
Part III: Nonparametric and robust methods for multivariate and functional data (13 chapters); and
Part IV: Invariant coordinates selection and related methods (six chapters).

BOOK REVIEWS 149
The remaining chapters of the book are written by experts in their respective fields on a multitude of research topics, however they are related to the work of Hannu Oja. The chapters in part II concentrate on univariate methods whereas Part III is devoted to multivariate models. Part IV presents some work on mixed topics.

A bias selection of chapters is

- Permutation Tests in Linear Regression
- Optimal Rank Tests for Symmetry Against Edgeworth-Type Alternatives
- M-Estimators of the Correlation Coefficient for Bivariate Independent Component Distributions
- Algorithms for the Spatial Median
- L1-Regression for Multivariate Clustered Data
- Tyler’s M-Estimator in High-Dimensional Financial-Data Analysis
- Robust High-Dimensional Precision Matrix Estimation
- On ANOVA-Like Matrix Decompositions
- Robust Simultaneous Sparse Approximation
- Nonparametric Detection of Complex-Valued Cyclostationary Signals

The book is well organized and properly structured. The chapters are self-contained. Evidently, the Festschrift comprises a host of useful topics and techniques for researchers in the field and to a degree for practitioners. The chapters are written in intuitive and appealing style, which makes the book accessible to a wider readership. However, the book offers advanced methods in the field, some useful practical examples, and a few figures, including in color. The book has a nice collection of nonparametric robust and multivariate procedures, and some related works on statistical signal processing. It also offers some historical perspectives on some of the work introduced by Hannu Oja. Undoubtedly, this book will be of interest to researchers in statistical science yet has substantial overlap with some related works on statistical signal processing.

In summary, Innovative Statistical Methods for Public Health Data has a good collection of research articles on useful and interesting topics about biomedical and public health data. Researchers and professionals looking to learn more in this field of study could benefit from articles published in this volume.

**S.E. AHMED**
_Brock University_

**Integrating Omics Data**, by George Tseng, Debashis Ghosh, and Xianghong Jasmine Zhou (eds.). New York: Cambridge Press, 2015, x + 461 pp., $119.00 (H), ISBN: 978-1-107-06911-4.

This edited volume showcases in depth important topics in biomedical research, including high-throughput genomic, proteomic, and transcriptomic. The articles published in this text are organized in chapters. There are 19 research and applications-oriented articles in this edited volume. Each chapter follows the theme of the volume, that is, focusing on methods to handle information integration, experimental data, and database problems of omics data. It is anticipated that this collection will be useful for the researchers working in the omics data integration and related biomedical fields. This collection is divided into three groups accordingly:

- Part A—Horizontal Meta-Analysis (five chapters);
- Part B—Vertical Integrative Analysis (general methods; six chapters); and
- Part C—Vertical Integrative Analysis (Methods specialize to particular data types; eight chapters).

The collection contains a number of real data examples, figures, and tables. While it does not provide a list of these, the information of this feature proves to be very useful. The topics of this book are helpful, interesting, and related with the important areas of research in a host arena linked to the omics data. One of the prominent features is that the chapters of this volume focus on the same subject matter making it useful for the readers. I find chapter one especially appealing due to my own research interest, as it blends theory and applications together nicely.

Below is a biased selection of topics of the respective articles in the edited volume:

- MetaOmnics: Transcriptomic meta-analysis Methods for Biomarker Detection, Pathway Analysis and Other Exploratory Purposes
- Network Integration of Genetically Regulated Gene Expression to Study Complex Diseases
- Identify Multi-Dimensional Modules from Diverse Cancer Genomics Data
- Penalized Integrative Analysis of High-Dimensional Omics Data
- Bayesian Models for Flexible Integrative Analysis of Multi-Platform Genomics Data
- Exploratory Methods to Integrate Multisource Data
- Integration of Cancer Omics Data into a Whole-Cell Pathway Model of Patient-Specific Interpretation
- Analyzing Combinations of Somatic Mutations in Cancer Genomes
- Data Integration of Noncoding RNA Studies
- Drug-Pathway Association Analysis: Integration of High-dimensional Transcriptional and Drug Sensitivity Profile

The style and structure of the volume is appealing and unified. Each article begins with an abstract and finishes with a summary and then list of references. The subject index is also provided. Relatively speaking, some chapters are more methodological in nature than others. Most of the chapters seem to be very practical, presenting the numerical work in the form of simulation and data analysis. It is important to keep in mind that this is a collection of wide and
diverse topics, which gives room for a variety of different views both simplified and intended for well-informed audience in the field. In summary, this collection of important articles proves to be an appropriate and timely contribution to a wider host community working in the bioinformatics sector. Researchers and professionals looking to learn more in this field of study could benefit from the articles published in this volume. The content in most of these selected articles has proved to be an enjoyable read. The book contains a wealth of information, a one-stop shop, and can be served as a research reference book. In closing, the book presents interesting and stimulating knowledge for the benefit of a host of research community in bioinformatics field and validates the importance and application of the wonderful discipline of statistical science and other related fields.

S.E. AHMED
Brock University

Selected Papers I of William Feller, by Rene L. SCHILLING, Zoran VONDRAČEK, and Wojbor A. WOYCZYNSKI (eds.). New York: Springer, 2015, xxxvi + 820 pp., $127.00 (H), ISBN: 978-3-319-16858-6.

I could not agree more with the first sentence borrowed from paragraph 2 of the Preface: “Feller is widely known for his brilliant two-volume text books.” The scientific contributions of Feller to the mathematical science community are foundational and unparalleled. He was a prolific researcher and published path-breaking work on probability theory, mathematical biology, measure theory, and geometry. He is an accomplished and distinguished scholar and left a deep and solid footprint on the research field of mathematical science.

The main purpose of this volume is to provide an overview of William Feller’s contributions to mathematical science and beyond. This volume presents selected and seminal work of Feller. One of the important features of the volume is that it contains commentaries and essays written by the experts providing a nice summary and discussions on the respective topics and giving a modern perspective. Volume 1 covers the years 1928–1950; and showcases Feller’s contributions in chronological order and contains 36 reprints of Feller’s original articles.

The collection includes a biographical sketch and some pictures of William Feller, as well as offering a bibliography, curriculum vita of Feller 1906–1970; also provided is a list of Ph.D. students supervised/co-supervised by William Feller. It was a pleasure re-reading some of the articles, especially his 1937 article “on the Law of large Numbers.” This book is a rich collection of important works, and delivers historical perspectives of some of the affluent work in respective research fields. This volume will be of interest to a wider readership from host scientific fields.

Below is a biased selection of topics of the respective articles in the volume:

- (translation) On the Central Limit Theorem of Probability Data
- On the Theory of Stochastic Processes (Existence and Uniqueness)
- On the Central Limit Theorem of Probability Theory. II
- Completely Monotone Functions and Sequences
- On the Integral Equation of Renewal Theory
- Some Geometric Inequalities
- Generalization of a Probability Limit Theorem of Cramer
- The General Form of the So-called Law of the Iterated Logarithm
- On the Normal Approximation to the Binomial Distribution
- The Fundamental Limit Theorems in Probability
- A Limit Theorem for Random Variables with Infinite Moments
- The Law of the Iterated Logarithm for Identically Distributed Random Variables
- On the Kolmogorov–Smirnov Limit Theorems for Empirical Distributions

S.E. AHMED
Brock University