not to say that all clinical trials should be devoid of such analyses.

Related to the authors' (or is it the editors'?) view that clinical trials are "fundamentally hypothesis testing instruments," they also challenge the value of estimates of effect size. Significance tests (at least in well-run studies) are "always valid" (according to the text), but estimates of effect size are not (or may not be) because of selected populations, idealistic trial conditions, and so on. This is something I have long believed and I am pleased to see it written and justified, but it does leave me a bit worried. I thought 20 or so years ago, statisticians had much success in convincing clinicians that *P*-values were not all that was important—*how much* benefit (or harm!) was at least as important. I would be left a little worried if ever some of my clinical colleagues were to pick up on this challenge to the perceived wisdom of estimation. And besides, while absolute effects are almost certainly unbelievable, relative effects (or differences from comparator) might be more believable.

In addition to chapters broadly applicable across most trials, there are two quite specific chapters: one on survival analysis and one on longitudinal data. Both are very good (and useful), but others could have been considered: binary data or nonparametric methods, as examples. There is a general chapter on "Selected Issues in the Analysis" which covers (among other things) missing data, subgroup analyses, and multiplicity. Overall, I was very impressed by all the content but if I had to pick out the best of the bunch, I would cite a chapter on data monitoring and interim analyses contributed by Kyungmann Kim, Thomas Cook, and Dave DeMets. The (technical) statistical content is the main focus of the book and this is what helps it to stand apart from most others on clinical trials (even the more obviously statistically orientated ones). It takes the reader to quite a technical background that would serve him or her well if moving on to research problems in the various areas covered, yet does not lose sight of practical issues.

Production and finishing is generally to a high standard although there are a few typographical/style "errors" that should have been picked up in the copy editing. There are some differences in style of listing references, probably symptomatic of its being an edited book with many contributors. But these are all minor issues and I did not notice any gross and important errors. I was, however, amused by an entry in the references to Sir Austin Bradford-Hill, listed as "Hill, S. A. B." Ironically, the adjacent reference is to "Hill, A."

So despite a few niggling points (which have perhaps been overemphasized in this review), my overall summing up has to be very positive. There is much material in this book that is not for the casual reader and it is very much aimed at statisticians. Nonstatisticians will be comfortable with some of the content but will occasionally turn a page and remember why they gave up maths at school! For those of us with the interest (and need) to grapple with these more statistical issues I wholeheartedly recommend it.

Simon Day
Roche Products Ltd.
Welwyn Garden City
United Kingdom

CHOW, S.-C., SHAO, J., and WANG, H. **Sample Size Calculations in Clinical Research**, 2nd edition. Chapman and Hall/CRC, New York, 2008. x + 465 pp. $88.95/£44.09. ISBN 9781584889823.

This is the second edition of a widely used book on sample size selection. The first edition of this book has gotten quite favorable reviews. Shanmugam (2004) noted that it is "a good reference book for researchers in clinical trials." In a second review, Shanmugam (2005) described the first edition as "a fascinating book." He added that "applied statisticians, health and medical researchers will like it a lot. Statistical consultants will be fond of the book as a reference guide."

The second edition has been carefully updated from the first edition. For example, the discussion of power and sample size for the false discovery rate contains a careful review of the work of Jung, Chow, and Young (2005) that came out after the first edition.

This edition of the book, like the last one, is aimed at all players in medical and pharmaceutical research, including biostatisticians, doctors, and scientists. Somewhat drolly, the authors note that the aim of this book is to avoid the "(i) wrong test for the right hypotheses, (ii) right test for the wrong hypotheses, (iii) wrong test for wrong hypotheses, or (iv) right test for the right hypotheses with insufficient power."

The book begins with a comprehensive discussion of the American Food and Drug Administration regulatory requirements for sample size calculations. The brief review of confounding, interaction, crossover and parallel designs, subgroup and interim analysis, randomization, and rare events is a helpful summary of trial design aspects that can affect sample size.

The remaining chapters include formulae to calculate sample size for the comparison of means or variances, large sample tests for proportions, goodness of fit and contingency tables, time-to-event data, group sequential methods, bioequivalence testing, dose–response studies, microarray studies, Bayesian sample size calculation, and nonparametrics. The discussion of vaccine clinical trials is well done, and difficult to find in most other discussions of power and sample size.

One of the strengths of the book is the organizational structure. Each chapter contains comprehensive references, examples, and practical considerations. The book is clearly laid out and easy to read. The table of contents and the index are comprehensive, which makes the book quite useful as a reference. Despite the utility of the formulae, one deficit of the book is the lack of any software, or reference to software. Users of this book have to locate or program the formulae and algorithms on their own.

In summary, this is a useful, comprehensive compendium of almost every possible sample size formula. The strong organization and carefully defined formulae will aid any researcher designing a study.

References

Jung, S. H., Chow, S. C., and Young, S. S. (2005). Sample size calculation for multiple testing in microarray data analysis. *Biostatistics* **6,** 157–169.

Shanmugam, R. (2004). Book review. *Journal of Statistical Computation & Simulation* **74,** 387–390.

Shanmugam, R. (2005). A review of advanced biostatistics and statistical genetics books. *Journal of Statistical Computation & Simulation* **75,** 761–769.

D. H. Glueck
Colorado School of Public Health
University of Colorado Denver
Denver, Colorado 80262, U.S.A.

RUBINSTEIN, R. Y. and KROESE, D. P. **Simulation and the Monte Carlo Method, 2nd edition**. John Wiley and Sons, Hoboken, New Jersey, 2008. xvii + 345 pp. $99.95/€77.99. ISBN 9780470177945.

Except for the title and the first author, I would not recognize this as the second edition of a book I first read over 25 years ago. I had to check carefully, but I now am sure that "second edition" is in reference to the book by the first author in 1981. In the preface, the authors use the standard phrase from promotional blurbs, "long-awaited second edition." The first edition was reviewed in *Biometrics* in 1983 (vol. 39, p. 302), but this edition is substantially different.

The book is organized into nine chapters and an appendix. The first chapter, "preliminaries," has twelve sections on basic topics in probability, two sections on statistics, and one section on optimization and duality. The organization and contents of this chapter illustrate my main criticism of the book: the units, either chapters or sections, are too diverse to allow sufficient depth of coverage, and further, the units do not have a strong logical continuity. After the first two chapters, on preliminaries and random number generation, the subsequent chapters seem to be a random selection of topics. The topics themselves, for example, "counting via Monte Carlo" or "the cross-entropy method," are all interesting and include useful examples, but I feel that a more unified organization would have resulted in a much better book. I would not choose this book as the textbook for a course, but I would readily refer to it as a source of supplemental examples.

Many items appear essentially in laundry lists. For example, when the Kullback–Leibler divergence measure is introduced, the general φ-divergence measure is defined (with a typo; the "p" should be an "h") and five versions of this measure are given. There is no further comment; the only one used subsequently is the Kullback–Leibler divergence measure, which is used in the cross-entropy method of choosing a reference vector in importance sampling. Would other measures work equally well, or are there instances where one should be preferred over the others? Otherwise, why develop the general formula and list the five instances?

Each chapter has 10 to 20 exercises. They generally appear to be interesting, but, with some notable exceptions, are rather easy. Some require programming. No particular software is required, although the authors give sample `Matlab` programs in various places in the text. There is a "detailed'' (in the authors' words) solutions manual, but I did not see it.

The book abounds with acronyms. Fortunately the authors provide a list, but many terms, such as "VM" (variance minimization) or "SF" (score function), do not seem to justify the overhead required for the reader's recognition, especially because a given acronym is used only in one chapter. (The problem is when the reader returns to that chapter for a quick fact check, but does not want to proceed sequentially through the chapter.) My dislike of the multitude of acronyms is related to the diversity of topics.

If an acronym is very standard in a field, or if in a single document, a great deal of discussion is devoted to a concept, use of an acronym is justified.

The references are listed at the ends of the chapters, which given the diversity of topics makes sense. The references also illustrate my criticism of lack of unity; in one chapter one book is given as the reference for a particular topic, and in a different chapter, another book given for the same topic (and, in my opinion, at least in some cases, if I were to give the two references for the given topic, I would have reversed the selection because of the context). Most of the references are from the last century. (For example, in chapters 6 and 7, of the 46 references, 31 are from the 1900s).

Despite my criticisms about the organization, I enjoyed reading the book, and found the individual examples quite interesting.

James E. Gentle
Department of Computational
and Data Sciences
George Mason University
Fairfax, Virginia 22030, U.S.A.

SHMULEVICH, I. and DOUGHERTY, E. R. **Genomic Signal Processing**. Princeton University Press, Princeton, New Jersey, 2007. xiii + 298 pp. US$60.00/£35.00, ISBN 0-69111-762-4.

This book is written from an engineering perspective on topics related to genomic signal processing (GSP), all of which are essential building blocks in the emerging field of biological network modelling and systems biology. According to the authors, "GSP is the engineering discipline that studies the processing of genomic signals ... (the goal of GSP) is to integrate the theory and methods of signal processing with the global understanding of functional genomics." The goal of the book is to provide a rigorous mathematical foundation in the area while also connecting readers to real-world problems. The authors are eminently qualified to write about the topic based on their rich experiences and outstanding publications in the field.

The authors first present a terse, self-contained introduction to molecular biology, genomics, and proteomics. As can be imagined, 21 pages of coverage is neither elementary enough for a novice outside the field nor advanced enough for a researcher already inside the field. The main body of the book consists of two parts. The first part is dedicated to genetic networks with deterministic models in Chapter 2 and stochastic models in Chapter 3. The basic properties of graphic models and insightful discussions of popular Boolean networks and Bayesian networks are covered. Differential equation models are touched upon in a short section. In the second part of the book, the authors discuss topics in supervised machine learning, covering the fundamental concepts of classification in Chapter 4 and an important topic of regularization for feature selection in Chapter 5. The content