Biometrics 66, 994–998
September 2010

BOOK REVIEWS

EDITOR:
GUILHERME J. M. ROSA

Design of Observational Studies
(P. R. Rosenbaum) *Peter B. Gilbert*

Longitudinal Data Analysis
(G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs, Editors) *Joseph W. Hogan*

Multiple Testing Problems in Pharmaceutical Statistics
(A. Dmitrienko, A. C. Tamhane, and F. Bretz, Editors) *Jason Stover*

Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families
(A. W. Marshall and I. Olkin) *Yosef Rinott*

ROSENBAUM, P. R. **Design of Observational Studies.** Springer, New York, 2010. xviii + 385 pp. $89.95/€85.55. ISBN 9781441912121.

Design of Observational Studies by P.R. Rosenbaum, published in 2010 by Springer Series in Statistics, follows Prof. Rosenbaum's previous book *Observational Studies* (2002). The new book is less technical, proving less and discussing more on philosophy, history, and heuristics of sound and creative study design. As such the book majors on the most important thing—how to think well—and is of high value for any statistician. Prof. Rosenbaum seems to emphasize the role of the statistician as a "philosopher of science" (definition from E.S. Gilbert), making the book more intellectually exciting than other statistics books that focus on technical derivations. It is obvious that Prof. Rosenbaum enjoyed writing this elegant book, with prose including the lines, "To believe that something is true because it would be convenient if it were true is a fair definition of naïveté" [on assuming strongly ignorable treatment assignments in observational studies]; "In our ignorance, we scientists spend a lot of time fishing in lakes with no fish, and you need to get the basic concepts down" [on defining power]; and "Given that you are alive reading this sentence, there is a small chance you will not live to read the last word" [on defining a hazard rate].

Consistent with Blaise Pascal's (1670) observation, "The last thing one knows in constructing a work is what to put first," Prof. Rosenbaum asks in the preface, "To end well, how should we begin?" His choice pleases this reviewer who lives in the world of randomized studies; the book opens with Cochran's basic advice, "The planner of an observational study should always ask himself the question, 'How would the study be conducted if it were possible to do it by controlled experimentation?'" (Cochran, 1965). Launching from this point, in Chapter 2 Prof. Rosenbaum discusses central concepts for making causal inferences in randomized experiments, including valuable history lessons about the role R.A. Fisher and others played in conceiving causal inferences in terms of potential outcomes. This starting point allows Prof. Rosenbaum to go on to transparently describe the conditions under which an observational study yields causal inferences of interest.

After noting that it is impossible to have assurance about the required conditions for virtually any observational study ("Keep in mind that [with these conditions] we are speaking here about mere imaginings, suppositions, hopes, aspirations and attempts, as distinct from simple, undeniable facts created by randomization in Chapter 2"), the remainder of Part I provides highly instructive discussion about principles of sound and creative study design, illustrated with numerous real examples, and Parts II and III discuss at length two central activities for rigorous evaluation of causal inferences in observational studies. In particular, Part II describes several approaches to matching groups to make them alike as possible with respect to observable variables, addressing practical issues such as missing data and pointing to `R` code for implementation, whereas Part III describes design sensitivity as a criterion for comparing different study designs paired with analytic approaches by the sensitivity of inferences to unmeasured biases. Together Parts I—III express a compelling general approach to the design of observational studies wherein one first endeavors to control or avoid all observable biases to the best extent possible, and, recognizing that unobservable biases inevitably lie in wait, to secondly conduct a sensitivity analysis, and to formally account for design sensitivity in the choice of study design.

By virtue of working almost exclusively on randomized studies, I admit to an occasional overly simplified and overly judgmental thought, "there are so many lurking potential biases, and human fancy is all too prone to over-interpret trends, that I trust only randomized studies." While Prof. Rosenbaum's book confirms this judgment for poorly designed observational studies, it corrects me for painting all observational studies with the same brush, by introducing creative

devices that sometimes empower observational studies to yield credible answers, including disambiguation, multiple control groups, and coherence among several outcomes, and, crucially, by promulgating sensitivity analysis as a central activity for demonstrating when plausible unmeasured biases cannot overturn a causal inference. Moreover, knowledge of these creative devices, which is unlikely learned from the sphere of randomized studies, proves valuable for the design of both randomized and observational studies, and the book inspires greater use of sensitivity analysis in randomized studies, which is sorely needed. *Design of Observational Studies* has sparked lines of enquiry for improving my own specialty area of statistical science, randomized vaccine efficacy trials.

A limitation—or simply a fact—of the book is that it almost exclusively considers the technique of matching for minimizing observable biases, so much so that the book may have been appropriately titled, "*Design of Matched Observational Studies*." While the principles and concepts translate well to unmatched techniques such as regression, the absence of discussion about how the translation is made may make it hard for some readers to apply the book's ideas to the design of unmatched studies. I found myself wishing for an extra section that tutors the reader on how to draw these lines. But this is slight criticism compared to my high praise for the book as a treasure of statistical history, philosophy, and principles for sound and creative study design that would benefit almost any statistician.

References

Cochran, W. G. (1965). The planning of observational studies of human populations (with Discussion). *Journal of the Royal Statistical Society, Series A* **128,** 234–265.

Pascal, B. [1966 (1670)]. *Pensées*. Translation by A. J. Krailsheimer. London: Penguin Books.

Rosenbaum, P. R. (2002). *Observational Studies*, 2nd edition. New York: Springer.

Peter B. Gilbert
Fred Hutchinson Cancer Research Center
and Department of Biostatistics
University of Washington
Seattle, Washington, USA

FITZMAURICE, G., DAVIDIAN, M., VERBEKE, G., and MOLENBERGHS, G. (eds). **Longitudinal Data Analysis.** Chapman & Hall/CRC, New York, 2008. viii + 632 pp. $94.95/£59.99. ISBN 9781584886587.

Longitudinal Data Analysis is Chapman & Hall's first entry in its new series *Handbooks of Modern Statistical Methods*. The choice of topic is natural: longitudinal data analysis has a rich history in a number of fields, and new avenues of investigation continue to be opened up. The volume's editors have assembled a world-class panel of contributors; many have made seminal contributions to the field (this includes the editors themselves).

Immediately apparent is the uniformity of notation and writing style not typically found in volumes of this kind. The editors clearly have taken great care to ensure a whole document rather than a disjointed patchwork typical of similar collections. Chapters reflect contributor diversity while suppressing distracting idiosyncrasies. Chapman & Hall would do well to emulate this feature in future entries to the series.

The book is divided into five parts: historical overview, parametric modeling, non- and semi-parametric methods, joint models, and incomplete data. Each chapter is self-contained, which works well for readers who choose to read individual chapters, but it also leads to noticeable overlap in content: for example, generalized linear models and the random effects model are exhaustively described in early chapters, but then reintroduced—in detail—in each of several subsequent chapters.

The emphasis of the volume is overwhelmingly on regression models. This is especially true of the early chapters, which cover material that is by now 25–30 years removed from first appearance in the statistical literature. But a comprehensive account of longitudinal data analysis has to start somewhere, and Part II provides a full accounting of generalized linear models, random effects models, nonlinear regression, and generalized estimating equations. Augmenting the standard descriptions are broad historical perspective, some new insights, and several interesting examples.

Part III on non- and semi-parametric modeling is a modern and forward-looking summary of smoothing and functional data analysis, a truly unique contribution that would make an outstanding monograph on its own. Part IV comprises several chapters on joint modeling, a vague term typically used to label models for joint distribution of repeated measures and event-history data. The editors have a broader meaning in mind: Part IV contains a concise but exemplary chapter on repeated measures and event times (Chapter 15, by Diggle, Henderson, and Philipson), but also includes chapters on multivariate repeated measures processes and latent variable models for high-dimensional data (where more extensions of multilevel regression are encountered). The diversity of interesting examples, rather than a coherent description of "joint modeling," is the strength of this chapter.

A brief digression: In reading Parts III and IV, one is struck by the remarkably wide-ranging utility of the hierarchical generalized linear model (a.k.a. generalized linear mixed model, random effects model) as a tool for inference from longitudinal data. Multilevel regression formulations figure prominently in growth curve modeling, regression splines, functional data analysis, dimension reduction, and multivariate longitudinal outcomes. Parts III and IV are a testament to the durability and scope of the generalized linear model in handling both low- and high-dimensional data. A number of new methods having exotic nomenclature still reduce to a (hierarchical) generalized linear model of one kind or another.

With chapters on missing observations, dropout, and causal inference, Part V takes a sweeping view of the term "missing data." While it is true that a causal inference via the potential outcomes model is a missing data problem, drawing causal inferences has become so central to the collection and analysis