PSYCHOMETRIKA
https://doi.org/10.1007/s11336-018-9632-y

BOOK REVIEW

REVIEW OF OBSERVATION AND EXPERIMENT: AN INTRODUCTION TO CAUSAL INFERENCE BY PAUL R. ROSENBAUM

In 1922, RA Fisher set out to place the practice of statistics on a solid theoretical foundation. He wrote,

> . . . the object of statistical methods is the reduction of data . . . [This] is accomplished by constructing a hypothetical infinite population of which the actual data are regarded as constituting a random sample.

The choice of the mathematical form of the population, what Fisher called a "specification," is assumed to be known, except possibly for some parameters and is assumed to capture the essence of the real-world problem. Chance and uncertainty enter through the process of randomly sampling from the population. Once a specification has been established, the inference problem becomes that of using the data to say something about future observations, often by means of statements about the parameters in the specification, for example by providing point estimates, or interval estimates, or a posterior distribution.

Implicit in this approach is the assumption of a stable universe which allows the possibility of making accurate predictions. For example, an instructor using an initial set of measures of aptitude, performance and final grades from her previous classes of third graders might try to predict her current students' academic performance. If relevant circumstances remain stable, and by conditioning on the measured values, she can make a reliable forecast.

Causal inference is also about making predictions. However, causation is not concerned primarily with random variables under a stable set of circumstances. Rather, causation pertains to what systematic change would occur if the circumstances were altered in a specific manner. The strategy of conditioning is not adequate for causal inference. Conditioning for prediction depends on a stable set of circumstances, but analysis of causation entails consideration of a change of at least one important circumstance. Stable populations assumed by the traditional statistical paradigm can only reveal how various factors are associated, but it does not disclose how a change in one factor would produce changes in some other factor of interest. This is the problem of causal inference (see, for example, Weisberg 2010).

Before the twentieth century, causation was largely the domain of philosophers, from Aristotle to David Hume (see, for example, Loux and Crisp 2017). However, in the early 1920s, a number of foundational developments started to pave the way for learning about causality from data. Central to this development was Neyman's (1923) formalization of the concept of counterfactuals or "potential outcomes," that is, outcomes that would have been observed had conditions been different, for example, potential agricultural yields under different crop varieties. He also proposed a notation that facilitated the representation of potential outcomes. *Causal effects* are comparisons of potential outcomes under alternative treatments. Since we cannot see how a patient (or an agricultural plot or study subject) would have fared under the alternative treatment that was not actually received, causal inference is difficult because it is about something we can never see.

The goal of Paul Rosenbaum's new book, *Observation and Experiment: An Introduction to Causal Inference*, is to present the concepts of causal inference with "reasonable precision, but with a minimum of technical material." Rosenbaum is a gifted expositor, and as a result, this

Published online: 13 August 2018

book is an outstanding introduction to the topic for anyone who is interested in understanding the basic ideas and approaches to causal inference. His pedagogical approach is to introduce concepts "through examples that are first discussed in English, and then ... restated in precise terms." Readers who are interested in a more technical treatment of causal inference will find several other new books on the market; however, it is hard to imagine one that presents a more systematic, erudite and conceptual introduction to the theory and practice of causal inference. To give a sense of the clarity of his writing, consider how Rosenbaum has conceptualized the presentation of the material:

> I have drawn two red lines through causal inference, dividing the subject into three parts. To the left of the first red line are topics that can be illustrated with a scientific example ... They are not necessarily elementary topics; rather they are topics that can be discussed in a certain way.
>
> Between the two red lines are topics that require familiarity and comfort with the concepts of conditional probability for discrete random variables ... If you know conditional probability, then look at the endnotes they might provide added insight ...
>
> Beyond the second red line are topics that are a little too technical to discuss in this book. The endnotes provide pointers to places where you can read more if you are so included. (p. xi)

This book is so well written and pedagogically sound that it could be used as the text for a non-calculus, first-course in statistics taught from a causal inference perspective. The examples used throughout the book for motivation and exposition are taken from the literature and are carefully and fully developed. These are not toy examples. Rosenbaum clearly takes joy in illustrating how statistics and statistical thinking from a causal inference perspective contributes to the advancement of science.

The primary focus of Part I of the book is on randomized experiments. Here the author introduces basic statistical concepts, such as hypothesis testing and confidence intervals. From a historical and scientific perspective, the invention of the randomized experiment was truly a watershed moment. Fisher (1925) brilliantly recognized how randomization could achieve balance between treatment groups since, as Rosenbaum notes, "fair coins ignore the attributes of the patients when assigning treatment" (p. 10). Rosenbaum carefully introduces a notation for representing potential outcomes; systematically shows how randomization achieves covariate balance across treatment groups enabling inference about causal effects on populations of individuals; shows how to estimate the average causal effect; and motivates the specification of the null hypothesis distribution through the concept of uniformity trials,[1] providing a test for a population average null treatment effect.

The remaining three-fourths of the book is about making causal inferences from observational data. Rosenbaum's approach follows the work of Rubin (1974, 1978) who proposed the general use of potential outcomes not only in experiments but also in observational studies. Chapter 5 is key to making the transition between experimental and observational studies. This is where the concept of "ignorable treatment assignment" is introduced (also called "unconfounded treatment assignment"), a concept pivotal to the understanding of causal inference. In this chapter, Rosenbaum methodically takes the reader through a series of steps that introduces and illustrates the concept of ignorable treatment assignment. Rosenbaum writes,

[1] In a uniformity trial, although units are assigned randomly to treatment or control, all units are treated in the same way, i.e., all units receive the *same* treatment. "Using uniformity trials," Rosenbaum explains, "investigators learned empirically how much treated and control groups could differ when there was no treatment effect because all units were treated in the same way" (p. 33).

> . . .[I]gnorable treatment assignment means that two people, say, person i and person j, who look the same in terms of observed covariates, $x_i = x_j$, have the same probability of treatment, $0 < \pi_i = \pi_j < 1$. Stated a little more precisely, the probability, π_i, that person i receives treatment, may vary with the observed covariates, x_i, but among people with the same values of x_i, the probability π_i of treatment does not vary with their potential outcomes (p. 109).

Further, and this is indispensable for understanding the remainder of the book,

> When treatment assignment is ignorable, an analysis that does not take account of x can yield very misleading estimates of the effects caused by treatment, but a straightforward analysis that appropriately adjusts for x removes the problem (p. 98).

Of course, in an observational study, treatments/exposures are not assigned to subjects at random and therefore it is rare for an investigator to be able to claim that treatment assignment is ignorable. Thus, as anyone who has tried to draw causal conclusions from an observational study knows, "At the end of the day, scientific arguments about what causes what are almost invariably arguments about some covariate that was not measured or could not be measured" (p. 7).

How to help the scientific community reach consensus about the effectiveness of a treatment or a policy is the overarching theme of the remainder of the book. This involves issues related to weighing and evaluating evidence, often from multiple data sources, systematically eliminating alternative explanations for observed associations, including recognizing and eliminating potential sources of bias, sensitivity analyses, and the use of methods for covariate adjustment, such as matching and propensity score adjustment. All of these topics are addressed from a causal inference perspective.

The field of causal inference has advanced tremendously in the last two decades. *Observation and Experiment* is a wonderfully written and readable introduction to the field but by necessity is limited in its scope. Readers should know that there is a lot more to causal inference beyond average effects of binary treatments in single timepoint studies: dose–response curves, conditional/heterogeneous treatment effects, principal stratification, effects of dynamic interventions, optimal treatment regimes, time-varying confounding, mediation, interference, graphical methods, and so on. Readers interested in a causal inference perspective on some of these topics, especially in the longitudinal data setting, may find the forthcoming book by Hernán and Robins (2018) of interest.

In any serious introduction to such a large topic like causal inference, there will be non-trivial trade-offs between depth and breadth, and between complexity and accessibility. In *Observation and Experiment*, Paul Rosenbaum has found just the right balance, providing readers with a precise, accessible, and enjoyable entrée to the field of causal inference.

CARNEGIE MELLON UNIVERSITY *Joel B. Greenhouse and Edward H. Kennedy*

References

Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London, Series A*, *222*, 309–368.

Fisher, R. A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver and Boyd.

Hernán, M. A. & Robins, J. M. (2018). https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/. Accessed 6 August 2018.

Loux, M. J., & Crisp, T. M. (2017). *Metaphysics: A contemporary introduction* (4th ed.). New York: Routledge.

Neyman, J. (1923). On the application of probability theory to agricultural experiments, Essay on principles: Section 9, Annals of Agricultural Sciences X: 1–51 (Translated by D. M. Dabrowska and T. P. Speed (1990). *Statistical Science*, *5*(4), 465–472).

Rosenbaum, P. (2017). *Observation and experiment: An introduction to causal inference*. Cambridge: Harvard University Press.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701.
Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics*, *6*(1), 34–58.
Weisberg, H. I. (2010). *Bias and causation: Models and judgment for valid comparisons*. Hoboken: Wiley.

Manuscript Received: 5 JUL 2018