Observational Studies 1 (2015) 291-293 Submitted 10/15; Published 10/15

Review of the book "Causal Inference for Statistics, Social, and Biomedical Sciences" by G.W. Imbens and D.B. Rubin

Fabrizia Mealli **mealli@disia.unifi.it**
Department of Statistics, Informatics, Applications, University of Florence
50134 Florence, Italy

Research questions that motivate most studies in statistics-based sciences are causal in nature. Economists and social scientists are typically interested in estimating causal effects rather than mere associations between variables (e.g., the effects of training programs on subsequent labor market histories); the same is true for epidemiologists and medical doctors (e.g., is smoking causing lung cancer? what is the effect of pollution on health outcomes?).

Extracting information and drawing inferences about causal effects of treatments, interventions and actions is central to decision making in many disciplines and is broadly viewed as causal inference. In this groundbreaking book, Guido Imbens and Don Rubin tell us what statistics can say about causation and present statistical methods for studying causal questions. The book focuses on the most widely used statistical framework for causal inference: the potential outcome framework, also known as the Rubin Causal Model (RCM), a term coined by Holland (1986).

Some key features characterize the approach followed in this book. All causal questions are tied to specific interventions or treatments. Causal effects are defined as comparisons of potential outcomes under different treatment conditions for the same subjects. Each of these potential outcomes could have been observed had the treatment taken on the corresponding level. After the treatment has taken on a specific level, only the potential outcome corresponding to that level is realized and can be actually observed, while the other potential outcomes are missing. Causal inference is therefore fundamentally a missing data problem. As in all missing data problems, a key role is played by the mechanism that determines which data values are observed and which are missing. In causal inference, this mechanism is referred to as the assignment mechanism: the process that determines which units receive which treatments, hence which potential outcomes are realized and thus can be observed, and which are missing. In the RCM, the definition of causal effects is separated from the assignment mechanism. This is the major difference between the RCM and other frameworks for causal inference; this separation clarifies the role of potential outcomes and the assignment mechanism. The last component of the RCM is the (Bayesian) model for the potential outcomes and the covariates. The Bayesian paradigm is natural for addressing causal inference problems under the RCM and the book emphasizes the Bayesian perspective, where parameters are also seen as random variables. However, an interesting feature of the book is that it clarifies the different perspectives, frequentist and Bayesian, as well as the finite population and the super-population perspectives, also elucidating the various sources of uncertainty that arise in the estimation of causal effects.

A crucial and appealing aspect of the book is that it develops a formal statistical language in which causal effects can be unambiguously defined, and the assumptions needed for their identification are clearly stated.

The authors systematize and develop the early insights of Fisher and Neyman: They discuss why and how randomized experiments allow us to assess causal effects and then turn to observational studies. Experimental principles are translated into statistical practice; the book includes many detailed applications that arose from the authors' extensive research experience, with special focus on practical aspects for the applied researcher. These examples help to clarify and explain many important concepts and practical issues. Therefore, this book offers a wonderful introduction to fundamental principles and a practical resource for both methodologists and practitioners.

It can also be used to teach a statistics course at the graduate level. Formal statistical derivations are oftentimes relegated to appendices: This makes the reading easier and only those interested can focus on the technical parts. I have used the book for semester-length graduate and undergraduate courses at the University of Florence and at Harvard University, as well as for short courses taught mainly to students from statistics, but also from other disciplines using applied statistics, such as biostatistics, epidemiology, economics, political science.

The book is organized in twenty six chapters, divided in seven parts. The first part (Chapters 1-3) introduces the basic philosophy of the potential outcomes approach. Part II (Chapters 4-11) focuses on the analysis of classical randomized experiments, and includes Fishers p-value testing, Neymans repeated sampling estimation, model- based inference and regression methods. Part III (Chapters 12-16) discusses the design phase of an observational study that is assumed "regular." Part IV (Chapters 17-20) introduces methods for the analysis of studies with regular assignment mechanisms. Methods presented include matching, subclassification, weighting, and model-based methods. Part V (Chapters 21-22) presents some sensitivity analyses to assess the unconfoundedness assumption, which is a key requirement for a regular assignment mechanism. Part VI (Chapters 23-25) contains settings with a regular assignment mechanism but imperfect compliance with the assignment. Randomized experiments with noncompliance are used to introduce instrumental variables methods; Part VII concludes.

I see only a few minor limitations of the book, some also explicitly recognized by the authors. First, the lack of coverage of multiple and multivalued treatments, sequential treatments, and dynamic treatment regimes. Second, most of the book relies on the SUTVA assumption of no interference between units and no hidden version of the treatment. Third, the book does not provide the data and codes that are used in the various chapters. Fourth, the discussion and use of model-based Bayesian analysis for observational studies is a bit limited, and confined to Chapter 20.

Despite these few criticisms, I truly believe that this book provides unprecedented guidance for designing and analyzing both experimental and observational data, and for interpreting the results appropriately. I urge statisticians to read it, because it can change the way statistics is practiced.

In conclusion, I would like to express my congratulations to Guido and Don: with the book they are sharing with us their decades of experience and research in this area, producing this impressive text covering concepts, theory, methods and applications.

References

Holland, P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81: 945-960.