

MR3497861 62-01 62A99

Pearl, Judea (1-UCLA-SCS); **Glymour, Madelyn** (1-CMU-DQ);

Jewell, Nicholas P. (1-CA-BS)

★ **Causal inference in statistics.**

A primer.

John Wiley & Sons, Ltd., Chichester, 2016. xvii+136 pp. ISBN 978-1-119-18684-7

Despite the fact that quite a few high-quality books on the topic of causal inference have recently been published, this book clearly fills an important gap: that of providing a simple and clear primer. Its aim is explained in the preface: while causal questions motivate several data analysis problems, introductory textbooks seldom go beyond the “association does not imply causation” aphorism, possibly at most mentioning randomized experiments as the gold standard for assessment of causation.

In Chapter 1, the issues of confounding and the need for additional tools with respect to the basic statistical methods are motivated by the intriguing case of Simpson’s paradox, with some nice and clear examples. Then, basic ideas in conditional probability, expectation, regression and graphs are very briefly recalled. The set-up is non-technical and even if these topics are very broad, only what is strictly necessary for the understanding of the book’s material is reviewed.

Chapter 2 begins by explaining how the concepts of independence and conditional independence can be summarized by simple causal graphs. The ideas are then generalized to the fundamental concept of d -separation.

In Chapter 3, the concept of intervention is introduced and causal effects are formally defined and computed. The chapter begins with an intuitive introduction based upon simple examples and the comparison of controlled and observational studies. Causal effects are then computed in terms of preintervention probabilities (via the adjustment formula). The choice of the conditioning sets is thoroughly discussed. Although the chapter is focused on conditioning and on uses of simple do-calculus principles, the approach of inverse probability weighting is also mentioned before discussing mediation and linear path analysis. As in the previous chapters, intriguing examples are used to introduce and illustrate the main concepts and methods.

The final chapter is slightly more complex, and deals with counterfactuals. Use of counterfactuals is elegantly linked to the structural causal models outlined in the previous chapters. The authors discuss deterministic and nondeterministic counterfactuals, also in the context of linear models.

The book ends with some ideas on attribution and mediation. Several thought-provoking study questions, in the form of exercises, are given throughout the presentation, and they can be very helpful for a better understanding of the material and looking further into the subtleties of the concepts introduced. In summary, there is no doubt that a discussion of the basic ideas in causal inference should be included in all introductory courses of statistics. This book could serve as a very useful companion to the lectures.

Alessio Farcomeni