Causal Inference for the Social Sciences
2014 ICPSR Summer Program, Session 2

Jake Bowers* Ben Hansen† Tom Leavitt‡(t.a.)

July 21, 2014

Abstract

This course introduces methods and concepts used to infer causal effects from comparisons of intervention and control groups. Random assignment plays a central role in the relevant statistical theories, but the techniques and principles that emerge apply to nonrandomized comparisons as well. We'll use the potential outcomes framework of causality to analyze both randomized and observational studies, distinguishing different forms of random assignment and separating observational studies that involve instruments, discontinuities and other devices, highlighting relative strengths of the designs and the implications of study design for statistical analysis. Propensity score matching is treated in depth, with explicit instruction in the use of "optmatch" and related packages in R.

The course presupposes knowledge of multiple regression at the level of the ICPSR course Regression: II, as well as multiple regression with binary dependent variables (as taught in the ICPSR courses Regression: III or Maximum Likelihood). The part of the course presenting matching requires the use of R for computation, but other methods presented in the course are readily implemented either in R or in Stata.

The course meets 3–5pm, Monday through Friday, July 21 through August 15, in 296 Dennison[1], University of Michigan Central Campus. Periodically the course will move to the computer lab in B760 East Hall[2]. Ben is primary instructor for the first two weeks, Jake for the second.

Tom holds office hours noon-3pm, in the Newberry Residence. In weeks one and two, Ben holds office hours Monday, Wednesday and Friday between 9:30 and 10:30 a.m.[3], at the Helen Newberry residence. (Either in the first-floor common area near the rear of the building or in office number 225.) In weeks three and four, Jake holds office hours Monday, Wednesday and Friday between 9:30 and 10:30 a.m. at the Helen Newberry residence. Additional office hours may be announced.

*Political Science and Statistics Departments, U. of Illinois, Champaign-Urbana; jwbowers@illinois.edu

†Statistics Department and Survey Research Center, U. of Michigan; ben.hansen@umich.edu

‡Political Science, Columbia U.; thomasleavitt87@gmail.com

[1]South end of Dennison Hall's southern extension, second floor.

[2]At the southeast corner of East Hall, the corner closest to the Church St/South University intersection. Enter from Church St.

[3]With prior notice, BH *may* be able to start 10-15 minutes earlier or stay 15-30 minutes later.

Overview

We may all warn our freshmen that association is not causation, but inferring causation has always been a central aim both for statisticians and for their collaborators. Until recently, however, inference of causation from statistical evidence depended on murky, scarcely attainable requirements; in practice, the weight of casual arguments was largely determined by the scientific authority of the people making them.

Requirements for causal inference become more clear when they are framed in terms of *potential outcomes*. This was first done by Neyman, who in the 1920s used potential outcomes to model agricultural experiments. Fisher independently proposed a related but distinct, ultimately more influential, analysis of experiments in 1935, and a rich strain of causal analysis developed among his intellectual progeny. It clarified the differing requirements for causal inference with experiments and with observational data, isolating the distinct contributions required of the statistician and of his disciplinary collaborators; generated more satisfying methods with which to address potential confounding due to measured variables; qualitatively and quantitatively advanced our grasp of unmeasured confounding and its potential ramifications; furnished statistical methods with which to eke more out of the strongest study designs, under fewer assumptions; and articulated principles with which to understand study designs as a spectrum, rather than a dichotomy between “good" experiments and “bad" observational studies. Understanding the methods and outlook of the school founded by Fisher’s student W.G. Cochran will be the central task of this course.

The course begins by applying the Fisher and Neyman-Rubin models of causality to experiments. Perhaps the best known technique to emerge from the Cochran school is Rosenbaum and Rubin’s propensity score matching, a technique used primarily in observational studies. The course continues here, covering ignorability, selection, “common support,” optimal matching, and covariate balance, probably also touching on sensitivity analysis. A short separate section introduces another method aiming to identify experiment-like structures in observational data, namely regression discontinuity, before a return to experiments. Central topics for this second segment on random assignment include instrumental variables, local average treatment effects and clustered assignment to treatment. W then treat in depth the topics of how to adapt statistical modeling strategies to comport with probability structures corresponding to experiments and observational studies attempting to mimic experiments; and studying the sensitivity of inference from observational studies to departures from random assignment. Over the course of the four weeks the course becomes progressively less lecture-oriented and more hands-on, with increasing emphasis on computing strategies in R.

Administrative

Textbook(s)

The main text for the course is

> Paul R. Rosenbaum. *Design of Observational Studies*. Springer, New York, 2010.

(Hereafter "*DOS*".) Although we won't follow it closely, its goals and methods align with the course's, and it will be useful as a reference and supplement. Other texts that we draw on include

Alan S Gerber and Donald P Green. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012.

Thad Dunning. *Natural experiments in the social sciences: a design-based approach*. Cambridge University Press, 2012.

Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: an empiricist's companion*. Princeton University Press, 2009.

A number of other graduate-level texts on or touching on causal inference have emerged in recent years. Many of them are quite good, including:

Richard A Berk. *Regression analysis: A constructive critique*, volume 11. Sage, 2004.

Andrew Gelman and Jennifer Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, 2006.

Stephen L. Morgan and Christopher Winship. *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge University Press, 2007.

Richard J Murnane and John B Willett. *Methods matter: Improving causal inference in educational and social science research*. Oxford University Press, 2010.

Other readings will be assigned and distributed electronically through a password-protected "ctools" Web site (`http://ctools.umich.edu`).

If you're new to R, we suggest getting hold of:

John Fox and Sanford Weisberg. *An R Companion to Applied Regression*. Sage, 2011.

R software will be required for several specific segments of the course. With some independent effort, students not familiar with R in advance should be able to learn enough R during the course to complete these assignments.

Assignments

Assignments are due each Tuesday, at the beginning of class. Parts of the assignment will be given at the beginning of the week, but other parts will be given during class, over the course of the week. Many of these daily assignments will be given with the expectation that they'll be completed by the next course meeting, although they'll only be collected at the end of the week. Late homework will not be accepted without cause (or prior arrangement with the teaching assistant).

You're welcome to submit a paper at the end of the course, whether or not you're taking the course for credit. In that case we'll return it with comments within a month or so of the course's completion. (If you're taking the course for a grade, the paper won't contribute to the grade unless you're on the borderline between two grades.)

Participation is expected. Some recommended forms:

1. In class, chime in occasionally with a topical question, remark or answer to someone else's question;
2. Drop by one of the professor's office hours to share a point that you and at least one classmate would like to have clarified or amplified, or to point out a connection to your field;
3. Give a 5-10 minute in-class presentation of a paper in your field that uses methods or designs we're discussing in the course.

If you are taking the course for a grade, make a point of doing at least one of 2 and 3. There'll be an electronic sign-up for 3.

Course management system(s)

Readings will be distributed via the ctools.umich.edu system.

You'll receive invitations to join both sites at the email address that appears for you on the course roster, the address you gave when you signed up for the Summer Program. Sign in to the sites using that email address. (You may have received a `umich` email address upon arrival in Ann Arbor, perhaps for using the University's wireless networks. Do *not* use this address for accessing Ctools: if you do, you may be allowed to authenticate to the system but you won't be able to see or access website for this course!)

Course contents

Experiments and the foundations of causal inference

Medical- and social-science data generating processes can be difficult to capture accurately in single regression equation, for various reasons. The statistical foundations of randomized experiments are much more satisfying, particularly when they are taken on their own terms. Fisher and Neyman did this earlier in the 1920s and 30s, in work that Rubin, Holland and others reinvigorated beginning in the 1970s. The course begins by surveying the circle of ideas to emerge from this.

In partial compensation for this early focus on experiments to the expense of observational studies, for the course's first meeting please read *DOS,* pages 3–4. It's short and it contains a rationale for beginning by understanding experiments, even if you'll be working exclusively with observational data.

Some readings for this course segment:

Chapter 1, "Introduction," of R. A. Fisher. *Design of Experiments*. Oliver and Boyd, Edinburgh, 1935.

Section 1.2, "Experimentation defined," of Donald R Kinder and Thomas R Palfrey. *Experimental foundations of political science*. University of Michigan Press, 1993. (Particularly pp. 5–10.)

Chapter 2 of Alan S Gerber and Donald P Green. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012.

Chapter 2 of Paul R. Rosenbaum. *Design of Observational Studies*. Springer Verlag, 2010.

Section 2.3 of Paul R. Rosenbaum. Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, 17(3):286–327, 2002.

Additional references:

Chapter 1 of Alan S Gerber and Donald P Green. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012.

P. W. Holland. Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, 81:945–970, 1986. (The article that brought the "Rubin Causal Model" to statisticians' attention.)

Endnote spanning pages A-32 and 33, David A. Freedman, Robert Pisani, and Roger Purves. *Statistics*. W.W. Norton and Co., 1998. (This can be read as a précis of: Jerzy Neyman. On the application of probability theory to agricultural experiments. essay on principles. section 9. *Statistical Science*, 5:463–480, 1990. transl. by D.M. Dabrowska and T.P. Speed from 1923 Polish original.)

Found experiments

Chapter 1 of DOS.

Chapter 6 of Thad Dunning. *Natural experiments in the social sciences: a design-based approach*. Cambridge University Press, 2012.

Ben B. Hansen and Jake Bowers. Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science*, 23(2):219–236, 2008.

Propensity score methods

Motivation and background:

D. B. Rubin. Using multivariate matched sampling and regression adjustment to control bias in observational studies. *Journal of the American Statistical Association*, 74:318–328, 1979.

Donald B. Rubin and Richard P. Waterman. Estimating the causal effects of marketing interventions using propensity score methodology. *Statistical Science*, 21(2):206–22, 2006.

Daniel Ho, Kosuke Imai, Gary King, and Elizabeth Stuart. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15:199–236, 2007.

Paul R Rosenbaum. Observational studies: Overview. In Neil J. Smelser and Paul B. Baltes, editors, *International Encyclopedia of the Social & Behavioral Sciences*, pages 10808–10815. Elsevier/North-Holland [Elsevier Science Publishing Co., New York; North-Holland Publishing Co., Amsterdam], 2001.

Guido Imbens and Donald Rubin. Rubin causal model. In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics*. Macmillan Publishers Ltd, 2008.

A success and a failure:

Robert Bifulco. Can nonexperimental estimates replicate estimates based on random assignment in evaluations of school choice? a within-study comparison. *Journal of Policy Analysis and Management*, 31(3):729–751, 2012.

Kevin Arceneaux, Alan S Gerber, and Donald P Green. A cautionary note on the use of matching to estimate causal effects: an empirical example comparing matching estimates to an experimental benchmark. *Sociological methods & research*, 39(2):256–282, 2010.

Matching a focal group to controls

DOS, Chap 8–9, 13.

P. R. Rosenbaum and D. B. Rubin. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *American Statistician*, 39:33–38, 1985.

Daniel Ho, Kosuke Imai, Gary King, and Elizabeth Stuart. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 15:199–236, 2007.

Ben B. Hansen. Propensity score matching to extract latent experiments from nonexperimental data: A case study. In Neil Dorans and Sandip Sinharay, editors, *Looking Back: Proceedings of a Conference in Honor of Paul W. Holland*, chapter 9, pages 149–181. Springer, 2011.

Non-bipartite matching

(Readings tba.)

Causal inference from assignment mechanism models

Instrumental Variables

Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91 (434):444–455, 1996.

Roderick JA Little and L. Yau. Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using rubin's causal model. *Psychological Methods*, 3(2):147–159, 1998.

Mike Baiocchi, Dylan S Small, Scott Lorch, and Paul R Rosenbaum. Building a stronger instrument in an observational study of perinatal care for premature infants. *Journal of the American Statistical Association*, 105(492):1285–1296, 2010.

Regression Discontinuity Designs

David S. Lee. Randomized experiments from non-random selection in u.s. house elections. *Journal of Econometrics*, 142:675–697, 2008.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

Devin Caughey and Jasjeet S. Sekhon. Elections and the regression discontinuity design: Lessons from close u.s. house races, 1942–2008. *Political Analysis*, 19(4):385–408, 2011.

Modeling the assignment mechanism

Why we might need a special story or additional steps in order to use familiar models:

David A. Freedman. Randomization does not justify logistic regression. *Statistical Science*, 23(2):237–249, 2008.

but see also

Winston Lin. Agnostic notes on regression adjustments to experimental data: reexamining freedmanâĂŹs critique. *Annals of Applied Statistics*, 7(1):295–318, 2013.

A special story and a few additional steps that enable you to use familiar models:

Paul R. Rosenbaum. Covariance adjustment in randomized experiments and observational studies. *Statistical Science*, 17(3):286–327, 2002.

In addition:

Example involving hierarchical data and a rather elaborate logistic regression model. See in particular methodological appendix.

M. Cerda, J.D. Morenoff, B.B. Hansen, K.J. Tessari Hicks, L.F. Duque, A. Restrepo, and A.V. Diez Roux. Reducing violence by transforming neighborhoods: A natural experiment in medellín. *American Journal of Epidemiology*, 175(10):1045–53, 2012.

Large samples enable some useful trickery. Example also demonstrates use of clustering.

Ben B. Hansen and Jake Bowers. Attributing effects to a cluster randomized get-out-the-vote campaign. *Journal of the American Statistical Association*, 104(487):873–85, 2009. DOI: 10.1198/jasa.2009.ap06589.

Instrumental variables.

Paul R Rosenbaum. Identification of causal effects using instrumental variables: Comment. *Journal of the American Statistical Association*, 91(434):465–468, June 1996.

Guido W. Imbens and Paul R. Rosenbaum. Robust, accurate confidence intervals with a weak instrument: Quarter of birth and education. *Journal of the Royal Statistical Society, Series A: Statistics in Society*, 168(1):109–126, 2005.

Last but not least

Sensitivity analysis

DOS, Chap 3

Carrie A. Hosman, Ben B. Hansen, and Paul W. Holland. The sensitivity of linear regression coefficients' confidence limits to the omission of a confounder. *Annals of Applied Statistics*, 4(2):849–870, 2010.

Interference

Michael E. Sobel. What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. *Journal of the American Statistical Association*, 101(476):1398–1407, 2006.

Michael G. Hudgens and M. Elizabeth Halloran. Toward causal inference with interference. *Journal of the American Statistical Association*, 103(482), 2008.

Paul R. Rosenbaum. Interference between units in randomized experiments. *Journal of the American Statistical Association*, 102(477):191–200, March 2007.

Jake Bowers, Mark M. Fredrickson, and Costas Panagopoulos. Reasoning about interference between units: A general framework. *Political Analysis*, 21:97–124, 2013.