Causal Inference for the Social Sciences

Jake Bowers * Ben Hansen † Thomas (Tom) Leavitt ‡
Naomi Nubin-Sellers § Saleheh (Sally) Sharif ¶

ICPSR Session 2 (July 20, 2020)

Abstract

This course introduces methods and concepts used to infer causal effects from comparisons of intervention and control groups. 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 the interplay 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; other areas of methodological focus include assessment of covariate balance by specification tests and other methods; inference methods that are robust to small sample sizes, weak instruments, spillover and interaction effects, heterogeneous treatment effects, and/or misspecification of response surfaces; and omitted variable sensitivity analysis.

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 3–5 pm, Monday through Friday, from July 20 through August 13. There will be no class on August 14, the final day of the course. Ben is the primary instructor for the first day of the course, Tom for the next four and then Jake for the final three weeks. Tom, Naomi and Sally will be the TAs for the duration of the course.

*Political Science and Statistics Departments, U. of Illinois, Champaign-Urbana; jwbowers@illinois.edu
†Statistics Department and Survey Research Center, U. of Michigan, Ann Arbor; ben.hansen@umich.edu
‡Political Science Department, Columbia U.; tl2624@columbia.edu
§Political Science Department, U. of Houston; nnubin@umich.edu
¶Political Science Department, The Graduate Center – CUNY; sally.sharif@law.cuny.edu
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 randomized experiments, touching on considerations specific to clustered treatment assignment, “small” sample sizes and treatment effect heterogeneity. The next segment addresses conceptual and methodological challenges of applying the same models to analysis of non-experimental data. This course segment covers ignorability, selection, “common support,” covariate balance, paired comparisons, optimal matching and propensity scores. A short separate section introduces another method aiming to identify experiment-like structures in observational data, namely regression discontinuity, before a return to experiments.

With these foundations in place, the course’s second half adds conceptual depth and methodological flexibility. Central topics include instrumental variables and local average treatment effects, stratified designs with clustering, interference, omitted variable sensitivity analysis and adapting workhorse techniques such as multiple regression to the demands of causal inference. 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

Textbooks

The main texts for the course are


These three textbooks are presented in varying difficulty and we will draw from all three. Although we won’t follow these books closely, their goals and methods align with the course’s, and they will be useful as references and supplements.

Other texts that we draw on include


Several other graduate-level monographs focus or touch on causal inference. Texts that we’ve found to contain helpful discussion include:


Other readings will be assigned and distributed electronically.

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

John Fox and Sanford Weisberg. *An R Companion to Applied Regression*. Sage, 2019

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. We also recommend some work with R — for example, via working through some online R courses — before the course for students who have never used it before.

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. It can take various forms:

1. Doing in-class exercises and discussing them with your peers;
2. From time to time, making a clarification or raising a clarifying questions;
3. Contributing to in-class discussions.
4. Using a github pull request to suggest a clarification or other enhancement to a course slide or worksheet.
5. 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;
6. 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 4, 5 and 6. There’ll be an electronic sign-up for 6.

Course contents

1 Potential outcomes and random assignment

Medical- and social-science data generating processes can be difficult to capture accurately in a 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.

**Required**


**Recommended**


---

2 Random assignment as a basis for inference

2.1 Inference for causal effects: the Neyman tradition

2.1.1 Estimation of average causal effects

**Required**


**Recommended**

2.1.2 Variance estimation and hypothesis testing

Required


Recommended


2.2 Inference for causal effects: the Fisherian tradition

Required


Recommended

3 Covariance adjustment in randomized experiments

Required


Recommended


4 Noncompliance and Attrition

4.1 Noncompliance and instrumental variables

Required


**Recommended**


### 4.2 Attrition, or missing outcomes

**Recommended**


5 Observational Studies

5.1 Embedding observational studies in experiments

Required


Recommended


5.2 Matching: Introduction

Required


5.3 Propensity scores methods

Required


5.4 More on matching

Required

Chapters 8 – 9 and 13 of Paul R. Rosenbaum. *Design of Observational Studies.* Springer Verlag, 2010

Recommended


5.5 Covariate balance and outcome analysis after matching

Required


Recommended


6 Sensitivity analysis

6.1 Sensitivity analysis for sharp nulls

Required


Recommended


6.2 Sensitivity analysis for weak nulls

Required


6.3 Regression-based sensitivity analysis

Required


Recommended


7 Additional topics

7.1 Interference


7.2 Regression discontinuity designs


### 7.3 Difference-in-Differences


### 7.4 Nonbipartite matching


### 7.5 Factorial experiments


### 7.6 External validity


Magdalena Bennett, Juan Pablo Vielma, and José R. Zubizarreta. Building representative matched samples with multi-valued treatments in large observational studies. *Journal of Computational and Graphical Statistics*, Forthcoming


### 7.7 Attributable effects


### 7.8 Bayesian causal inference


