Making Causal Inferences: Statistical Inference and Adjustment for Randomized and Observational Studies

Instructor: Jake Bowers
jwbowers@illinois.edu
http://jakebowers.org/

Fall 2023

Overview

Where & When

We will meet Wednesdays, 1:00 pm – 3:00 pm CT on Zoom. I currently plan to use Campuswire for coursea communication, GitHub for assignments and other files, and Moodle for grades and some other course assignments and readings.

To fit the the schedules of the most students, the class days will be: Sep 6, Sep 13, Sep 20, Sep 27, Oct 4, Oct 11. We may reschedule the Oct 18 and Oct 25 meetings to be later in the term using a poll of the students.

a

Office Hours

Please make an appointment on Calendly if you want to come to office hours to ensure that we can meet and talk. I'm happy to schedule other times if those don't work for you.

Introduction

How do you know when you have "controlled for" or "adjusted" a relationship potentially confounding covariates enough? When might an approach to adjustment confuse more than clarify? What do we mean when we say that a hypothesis test or estimator is "justified by randomization"? What might a p-value, confidence interval, or claim of unbiasedness in an estimator mean if we do not have a randomized design? When is a hypothesis test a good one? When is an estimator useful? How should we choose estimators or tests or approaches to adjustment? When might a confidence interval or p-value mislead you? How can you convince yourselves and others that what you are doing with statistics makes sense?

In this course we will introduce a way of learning about theoretical causal claims using research design and statistical inference that is tightly tied to the research design. We use research design to justify adjustment: For example, we might say that we do not worry about the potentially confounding effect of education in a study of a get out the vote policy because we used a randomized design. For another example, we might say that we do not worry about the potentially confounding effect of education in that study because only compare people who have the same education level — we match on education. When it comes time to calculate a p-value or confidence interval about counterfactual hypotheses or estimate an otherwise unobservable effect defined in counterfactual terms, this course will emphasize ways to justify choices of test and estimator using the research design used for adjustment in the first place as well as formal sensitivity analysis for observational studies.

As currently envisioned, this course will introduce the basics of randomized experiments, instrumental variables, difference-in-differences, and matching on binary variables (bipartite matching), matching on continuous or multivalued variables (non-bipartite matching), matching with longitudinal data, and formal sensitivity analysis. I will be open to adding and subtracting topics depending on student interest.