Stat 256: Causal Inference (Fall 2024)

Amanda Coston
University of California, Berkeley
Email: acoston@berkeley.edu

- Units: 4
- Lectures: Tuesday and Thursday 11:00 am – 12:29
- Office hours
    - GSI: TBD
    - Instructor: Thursday 4-5 pm Evans 427
- GSI lab sessions: Mondays 10 am–12pm and 4-6 pm Evans 340 (mainly for students in Stat 156)
- GSI: Mingrui Zhang (PhD student in Biostatistics)
- Course website: https://stat156.berkeley.edu/fall-2024/

This course provides an introduction to causal inference using the potential outcomes framework and causal diagrams. We will cover topics such as randomized experiments, observational studies, instrumental variables, and mediation analysis. We will discuss applications of these methods to a variety of fields such as medicine and public policy.

The course emphasizes both theory and data analysis. Lectures present concepts and methods, and lab sections focus on applications, derivations, and worked problems often using the statistical language R.

Consult the course website for updates, homework assignments, reading assignments, and course-related announcements. We will use bCourses to release solutions for homework and the midterm. Grades will be released on Gradescope.

Students with disabilities please use the DSP to register any accommodations you may need for physical, mental or learning disabilities. The Academic Accommodations Hub provides many resources relevant to accommodations and supportive measures.

Textbooks

Course readings will be drawn primarily from the following textbooks:

- Ding, P. (2024). *A first course in causal inference*. CRC Press. Available for purchase in print here. Add promotion code JSM24 for additional discount (good thru end of September). A slightly older version is available on arXiv for free.
- Robins, J., & Hernán, M. A. (2020). *Causal inference: what if.* CRC Press. Available online here for free.

Prerequisites

- Stat 201 or Stat 210A highly recommended
- Linear models and generalized linear models
- Software R and LaTex

Evaluation

Grades will be determined according to:

- Reading assignments (20%): Each week you will read research papers and write a report describing the research and your evaluation of the work. Each report should be 2-3 pages double-spaced. See end of syllabus for a list of the papers. The GSI will provide details on how to submit reports.
- Present two published papers and record the presentations (20% + 20%). The papers must be related to causal inference and published after 2012. At least one of the papers must be a journal paper. Recommended journals include Annals of Statistics, Journal of the American Statistical Association, Journal of the Royal Statistical Society Series B, Biometrika, Annals of Applied Statistics, Journal of the Royal Statistical Society Series A, Journal of the Royal Statistical Society Series C, Biometrics, Econometrica, Journal of Econometrics, Data & Policy, Journal of Machine Learning Research, PNAS, Science, Nature. Other recommended venues are International Conference on Artificial Intelligence and Statistics (AISTATS), International Conference on Machine Learning (ICML), Neural Information Processing Systems (NeurIPS), and the ACM Conference on Fairness, Accountability, and Transparency (FAccT). Your recorded presentation should be 20 minutes per paper.

  The first presentation is due October 15, 2024 and the second is due on November 14, 2024. Email the presentation to the instructor directly. You are welcome (but not required) to share your presentation on your website and/or youtube.
- Scribe notes (5%). Students will be in charge of scribing lecture notes in LaTex. Each student will sign up to scribe one lecture, along with 1-2 of their peers, in a shared Overleaf. Tuesday scribes must release the lecture notes by the following Friday and Thursday scribes must release the lecture notes by the following Monday.
- Group project with a presentation & final report (20% for final report; 15% for presentation). In groups of 3, you will prepare a research presentation and final report. Presentations will occur in class during the last lectures of the semester (on 12/3 and 12/5) to allow students time to incorporate feedback into the final report. The final report should be 15-20 pages. If your project involves analyzing data, make sure the data is publicly available.

Late assignments will not be accepted. If you have concerns about meeting a deadline, talk to the instructor as early as possible.

Academic Integrity

Code of Student Conduct is available at https://conduct.berkeley.edu/code-of-conduct/.

Wellness

Wellness is important but can sometimes feel challenging during the demands of the semester. I encourage you to make mental wellbeing a priority. The university has resources for ensuring basic needs. If you are in a crisis, call a counselor at (510) 642-9494 (Monday - Friday, 8:00 AM - 5:00 PM) or (855) 817-5667 (other hours). UHS offers many resources for maintaining mental health. You may find wellness resources on recalibrate. I am also available to chat, listen and share my own wellbeing journey and practices.

Course Outline

The planned topics for the course are

- Association and paradoxes (week 1)
- Potential outcomes framework (week 2)
- Randomized experiments (week 2, 3)
- Unconfounded observational studies (weeks 4, 5, 6)
- Instrumental variables (weeks 7,8)
- Sensitivity analysis (week 9)
- Negative controls (week 10)
- Principal stratification and mediation (week 11)
- Modern methods (week 12)

Reading Assignments

- Bickel et al 1975 Science: Sex Bias in Graduate Admissions: Data from Berkeley [week 1, link]
- Holland 1986 JASA Statistics and Causal Inference [week 2, link]
- Miratrix 2013 JRSSB Adjusting treatment effect estimates by post-stratication in randomized experiments [week 3, link]
- Lin 2013 AOAS: Agnostic notes on regression adjustments to experimental data: Reexamining Freedmans critique [week 4, link]
- Li, Ding and Rubin 2018 PNAS Asymptotic theory of rerandomization in treatment-control experiments [week 5, link]
- Rosenbaum and Rubin 1983 Biometrika: The central role of the propensity score in observational studies for causal effects [week 6, link]
- Lunceford and Davidian 2004 SiM Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study [week 7, link]
- Angrist Imbens and Rubin 1996 JASA Identification of causal effects using instrumental variables [week 8, link]
- Imbens 2014 StatSci Instrumental Variables: An Econometricians Perspective [week 9, link]
- Ding and VanderWeele 2016 Epidemiology: Sensitivity Analysis Without Assumptions [week 10, link]
- Pearl 1995 Biometrika Causal diagrams for empirical research [week 11, link]

- Frangakis and Rubin 2002 Biometrics Principal stratification in causal inference [week 12, link]