STAT 186
Statistical Methods for Evaluating Causal Effects
Department of Statistics
Spring 2017

Instructor: Donald B. Rubin

Time and Location: M, W 10:00 am - 11:30 am, SC Hall E

Contact: rubin@stat.harvard.edu

Office Hours: Monday 1:00 - 3:00 pm, SC 716

Teaching Fellow(s): Xinran Li(xinranli@fas.harvard.edu), Kristen Hunter(khunter@g.harvard.edu), Dongming Huang(dhuang01@g.harvard.edu)

Sections and Office Hours:

Monday 5:30-6:30 (section) SC 112, 6:30-7:30 (office hours) SC 110, Dongming
Tuesday 3-4 (section), 4-5 (office hours), SC 705, Kristen
Wednesday 6-7 (section), 7-8 (office hours), SC 705, Dongming
Thursday 4-5 (section), 5-6 (office hours), Harvard Hall 102, Xinran

Description:

Statistical methods for inferring causal effects from data from randomized experiments or observational studies. Students will develop expertise to assess the credibility of causal claims and the ability to apply the relevant statistical methods for estimating causal effects. Examples from many disciplines: economics, education, other social sciences, epidemiology, and biomedical science. Evaluations of job training programs, educational voucher schemes, changes in laws such as minimum wage laws, medical treatments, smoking, military service. The emphasis of the course will be understanding major ideas, not on memorizing equations.

Prerequisites:

Mathematics 21a or 21b (or more advanced) required; a previous statistics course such as Stat 100/104 required; additional courses having statistical content, for example, Stat 110, 111 and 139, or the equivalent courses in Economics, Psychology or Biostatistics, will make the course easier but are not required.

Section and Office Hours:

We have found that the concepts in this course are best communicated face to face, and we look forward to discussing your questions and reactions to the material during lecture, sections, and office hours.

Textbook:

The primary textbook for this class is *Causal Inference in Statistics, Social, and Biomedical Sciences: An Introduction*, by Guido W. Imbens and Donald B. Rubin.

Supplementary Books:

- *Matched Sampling for Causal Effects*, by Donald B. Rubin.
- *Experimental and quasi-experimental designs for generalized causal inference*, by William R. Shadish, Thomas D. Cook, and Donald T. Campbell.

Additional journal articles for discussion will also be made available.

Accommodations:

If you have documentation from the Accessible Education Office, please let us know early in the semester so that arrangements can be made.

Assignments and Grading:

There will be four large assignments (approximately every two-three weeks). All large assignments must be submitted to the teaching fellows (by hand) before the deadline. Class participation and attendance at section are strongly encouraged.

All quizzes will be conducted in class. Notes and laptops are allowed, but no WiFi use is allowed. There is a zero tolerance policy on cheating.

Assignment 1	10%
Assignment 2	10%
Assignment 3	10%
Assignment 4	10%
Quiz 1	5%
Quiz 2	5%
Quiz 3	5%
Quiz 4	5%
Final take-home Exam: project-oriented	30%
Participation	10%

Important Dates

This will be updated as the course progresses.

- Quiz 1: February 22

Policies:

Submit paper copies of your assignments to the teaching fellow(s) before the specified deadline, including your code with every assignment. Your code will be considered for partial credit, if necessary. If you miss a deadline but submit your assignment to a TF at or before the following lecture, we will grade your assignment but deduct 50% of the possible

points. An assignment received after this time will receive no credit. No assignments will be dropped when your grade is calculated. Extensions will be granted only in exceptional circumstances, such as serious illness documented by University Health Services or a family emergency. If you think you will miss a deadline because of interviews, your senior thesis, or other scheduling conflicts, submit the assignment early. The quizzes will be held in class only. There will be no make-ups.

Collaboration:

University policies against plagiarism will be strictly enforced. You are encouraged to discuss assignments and projects with your classmates, but each student must write solutions separately. Be sure that you have worked through each problem yourself and that the answers you submit are the result of your own efforts. You also may not share or view another student's computer code, submit output from another student's computer session, or allow another student to view your code or output. If a fellow student asks if you would like to discuss a problem, we encourage you to agree; if a fellow student asks to see your code or answer to a problem, the answer must be that it is not allowed.

You must acknowledge collaborators by writing their names at the top of your homework assignments.

Computing:

This course involves using the software R, which is available for free. No previous experience with R is necessary. However, those without previous R experience should expect to spend extra time on assignments as they familiarize themselves with the program. The TFs will provide support for coding in R, and one or two sections introducing R will be held early in the semester. We believe that learning R is worthwhile in its own right. You may choose to rely on other software for this course, but technical support and partial credit will be available for R code only.

Course Outline

Introduction (January and early February)

- Chapter 1 - The Basic Framework
- Chapter 2 - A Brief History of the Potential Outcome Approach to Causal Inference
- Chapter 3 - A Taxonomy of Assignment Mechanisms

Classical Randomized Experiments (February)

- Chapter 4 - A Taxonomy of Classical Randomized Experiments
- Chapter 5 - Fisher's Exact p-values for Completely Randomized Experiments
- Chapter 6 - Neyman's Repeated Sampling Approach to Completely Randomized Experiments

- Chapter 7 - Regression Methods for Completely Randomized Experiments
- Chapter 8 - Model-based Inference in Completely Randomized Experiments
- Chapter 9 - Stratified Randomized Experiments
- Chapter 10 - Paired Randomized Experiments

Regular Assignment Mechanisms (March)

- Chapter 12 - Unconfounded Treatment Assignment
- Chapter 13 - Estimating the Propensity Score
- Chapter 14 - Assessing Overlap in Covariate Distributions
- Chapter 15 and 16 - Design in Observational Studies: Matching, trimming, and subclassification
- Chapter 17 and 18 - Analysis in Observational Studies: Matching and subclassification

Irregular Assignment Mechanisms (March and April)

- Chapter 23 and 24 - Non-Compliance and Instrumental Variables Analysis
- Principal Stratification - Generalizing Instrumental Variables
- Additional topics depending on interests and potentially including recent developments in causal inference.