Special Topic: Causal Inference in Biomedical Studies
BIOST 578 - Spring 2024

1. Instructor

Ting Ye, PhD, Assistant Professor of Biostatistics

Guest instructors

Andrea Rotnitzky, PhD, Professor of Biostatistics
Marco Carone, PhD, Associate Professor of Biostatistics

Lectures: Mon/Wed, 1:30-2:50pm
Number of credits: 3

2. Course description

In biomedical studies (randomized or observational), many questions of interest fundamentally require assessing causal relationships rather than capturing mere associations. The overall goal of this course is to enable students to explicitly define causal questions, delineate underlying assumptions, and apply appropriate statistical methods. This course primarily focuses on causal inference on single time-point interventions using the potential outcomes framework. The main topics are estimand and causal identification, average treatment effect estimation using regression, weighting, and doubly robust methods, effect modification and interaction, matching and randomization inference, mediation analysis, instrumental variables. Some basics of causal graphs will also be introduced. Concepts and methods will be illustrated using biomedical data examples.

This course is designed for graduate students in Biostatistics, Statistics, Epidemiology, and other fields who are interested in applying causal inference methods to biomedical studies and/or conducting statistical research related to causal inference.

3. Course learning objectives

After completing this course, students will be able to:

- Understand how causal inference is different from conventional association analysis
- Delineate the causal estimand, assumptions, strengths and limitations of each method, and the conditions in which each method can be appropriately applied
- Choose and apply appropriate design and statistical methods for applied problems

4. Prerequisites

Either BIOST 536; or BIOST 515/518 (or STAT 502/504) and STAT 512/513; or BIOST 515/518 (or STAT 502/504) and BIOST 521/522, or equivalent; or permission of instructor. This course assumes an understanding of basic probability theory (e.g., independence, conditional expectation, normal distribution), linear regression and logistic regression, and R programming.

5. Grading policies

Grades will be determined by homework (50%) and a final project (50%).

For the final project, you could choose to (1) read about a statistical method related to causal inference that was not covered in class, and carry out a simulation study or data analysis to illustrate the method; (2) conduct a thorough data analysis using some of the methods covered in class.

The components of the projects are as follows:
- A one-page project proposal (10%)
- A presentation to class (20%)
- A written report of no more than 15 pages (including all materials, and in 12 point font size with one-inch margins on all sides) (20%)

6. Course materials

Lecture notes will be provided.

Recommended textbooks:
- Causal inference: What if, by Miguel Hernan and James Robins
- Causal Inference for Statistics, Social, and Biomedical Sciences, by Guido Imbens and Donald Rubin
- Causal inference in statistics, a primer, by Judea Pearl, Madelyn Glymour and Nicholas Jewell

- Design of observational studies, by Paul Rosenbaum
- Observational studies, by Paul Rosenbaum
- Mostly harmless econometrics: An empiricist's companion, by Joshua Angrist and Jorn-Steffen Pischke

7. Course session schedule

1. Motivation, potential outcomes, estimand, and causal identification
2. Causal graphs
3. Causal inference in randomized trials
4. Estimation of average treatment effect: regression, weighting, and doubly robust methods
5. Effect modification and interactions
6. Matching and randomization inference
7. Mediation analysis
8. Instrumental variables
9. Time-varying interventions (time permitting)