DEPARTMENT: BIOS/EPI

COURSE NUMBER: 761/760 **SECTION NUMBER:**

CREDIT HOURS: 4 **SEMESTER:** Spring 2021

COURSE TITLE: Causal inference

CLASS HOURS AND LOCATION:
TBD

INSTRUCTOR NAME: David Benkeser

INSTRUCTOR CONTACT INFORMATION

EMAIL: benkeser@emory.edu

PHONE: (404)712-9975

SCHOOL ADDRESS OR MAILBOX LOCATION: 1518-002-3AA

OFFICE HOURS

Teaching Assistant(s): TBD

COURSE DESCRIPTION

This course is an elective for Biostatistics PhD. students, a required course in the spring semester of the second year for Epidemiology PhD. students and is open to PhD students from other departments on a limited, case-by-case basis.

This course provides a survey of modern topics in causal inference. Fundamental concepts in causal inference will be covered including: counterfactual random variables, assessing identifiability of causal effects, graphical frameworks, G-computation, inverse probability of treatment weighting, methods for efficient, doubly (multiply) robust estimation of causal effects, and causal mediation. Where possible, the course emphasizes the use of modern regression (e.g., machine learning) in causal effect estimation and provides an applied introduction to this area is provided as well.

Methods from the course will be implemented in the R programming language and using relevant R packages. Intermediate R programming skills are expected (familiarity with manipulating data, writing functions in R, reading R documentation, the glm function are particularly important). Students who are not confident in these skills are encouraged to complete related online courses (e.g., Coursera or DataCamp) to brush up on skills ahead of enrollment in the course.

Knowledge of basic probability and statistical theory is required including the following topics: random variables, distributions of random variables, independence, expectations, conditional distributions, expectations, conditional expectations. Knowledge of statistics is also required including: sampling distributions, consistency, the central limit theorem, regression (linear and logistic), basic principles of maximum likelihood estimation.

A suggested companion textbook for the course is "Causal Inference: What If" by Hernan and Robins. For more advanced topics and further reading, "Targeted Learning: Causal Inference for Observational and Experimental Data" by van der Laan and Rose is also recommended.

At the end of the course students should be able to:
- Use common frameworks to translate a research question into a causal effect parameter.
- Use graphical frameworks to assess identifiability of causal effects.
- Implement simple approaches to estimation and inference on various causal effects.
- Understand the motivation for and implementation of efficient, doubly robust approaches to estimation of various causal effect using existing software packages.
- Understand the motivation for incorporating modern regression (e.g., machine learning) into causal effect estimation and the difficulties in so doing.
- Describe issues that arise due to non-compliance. Implement estimators of quantities that are appropriate for answering causal questions in this context.
- Describe fundamental principles of causal mediation. Implement estimators of various mediation parameters.
- Understand time-varying confounding and why simple regression approaches fail to adequately answer causal questions in these contexts.
- Implement estimators of causal parameters describing the effects of longitudinal interventions.
- Interpret and write about results of analyses and describe limitations of various approaches.

EVALUATION

Homework (weekly): 30%

Students will complete weekly homework assignments. The homework assignments will be a mixture of theoretical problems that complement lecture material (e.g., using a graph to determine whether a causal effect is identifiable), computational problems (e.g., implementing a particular estimator using a software package), and interpretation (e.g., writing a mock Methods section description of a method for an applied journal).

Midterm (in class): 20%

Final (in class): 20%

Final group project: 30%

Students will form small groups (up to 4 people per group) to complete a project involving the analysis of a real data set or a simulation study comparing different estimators of causal effects. Students will produce a written report, give a short oral presentation in class, and answer questions from the class and professor.

Grade scale*:

- A = 93 -- 100%
- A- = 90 -- 93%
- B+ = 87 – 90%
- B = 83 – 87%
- B- = 80 – 83%
- C = 65 – 80%
- F = <65%

*The lower number is inclusive, so e.g., an A is an grade greater than or equal to 93, an A- is any grade that is greater than or equal to 90 and strictly less than 93, a B+ is any grade that is greater than or equal to an 87 and strictly less than 90, etc…

COURSE STRUCTURE

The course will be organized into weekly lectures consisting of a combination of electronic slides, whiteboard problem solving, and computational demonstrations. Students are expected to ask and answer questions in class.

COURSE POLICIES

Students are expected to attend lectures and participate in discussions during class. At times, students may be encouraged, but not required, to bring a laptop to class to follow along with code demonstrations.

As the instructor of this course I endeavor to provide an inclusive learning environment. However, if you experience barriers to learning in this course, do not hesitate to discuss them with me and the Office for Equity and Inclusion, 404-727-9877.

RSPH POLICIES

Accessibility and Accommodations

Accessibility Services works with students who have disabilities to provide reasonable accommodations. In order to receive consideration for reasonable accommodations, you must contact the Office of Accessibility Services (OAS). It is the responsibility of the student to register with OAS. Please note that accommodations are not retroactive and that disability accommodations are not provided until an accommodation letter has been processed.

Students who registered with OAS and have a letter outlining their academic accommodations are strongly encouraged to coordinate a meeting time with me to discuss a protocol to implement the accommodations as needed throughout the semester. This meeting should occur as early in the semester as possible.

Contact Accessibility Services for more information at (404) 727-9877 or accessibility@emory.edu. Additional information is available at the OAS website at http://equityandinclusion.emory.edu/access/students/index.html

Honor Code

You are bound by Emory University's Student Honor and Conduct Code. RSPH requires that all material submitted by a student fulfilling his or her academic course of study must be the original work of the student. Violations of academic honor include any action by a student indicating dishonesty or a lack of integrity in academic ethics. *Academic dishonesty refers to cheating, plagiarizing, assisting other students without authorization, lying, tampering, or stealing in performing any academic work, and will not be tolerated under any circumstances.*

The RSPH Honor Code states: "Plagiarism is the act of presenting as one's own work the expression, words, or ideas of another person whether published or unpublished (including the work of another student). A writer's work should be regarded as his/her own property."
(http://www.sph.emory.edu/cms/current_students/enrollment_services/honor_code.html)

COURSE CALENDAR AND OUTLINE

Topics and dates are subject to change as the semester progresses. CIWI = "Causal Inference, What If?", TL1 = "Targeted Learning: Causal Inference for Observational and Experimental Data", TL2 = "Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies"

Date	Topics	Readings
(TBD)	Causation vs. association	CIWI Chpt 1, TL Chpt 1
	Identification	CIWI Chpt 3
	Identification using graphs	CIWI Chpt 3, TL Chpt 2.2
	G-computation and IPTW	CIWI Chpt 12/13, TL Chpt 2.3
	G-computation and IPTW	CIWI Chpt 12.2-12.3, 13; TL Chpt 2.3
	Super learning	TL Chpt 3
	Super learning	TL Chpt 3
	Efficient estimators (AIPTW/TMLE)	TL Chpt 4
	Efficient estimators (AIPTW/TMLE)	TL Chpt 4
	Other approaches: matching, propensity score adjustment	CIWI Chpt 15
	Multi-level/continuous valued treatments (marginal structural models)	CIWI Chpt 12.4
	Multi-level/continuous valued treatments (stochastic interventions)	TL2 Chpt 11 Optional: Diaz and van der Laan 2012 Biometrics data analysis
	Effect modification/heterogeneity	CIWI Chpt 4, TL Chpt 9
	Optimal treatments	TL2 Chpt 22
	Optimal treatments	TL2 Chpt 22
	MIDTERM	
	Instrumental variables	CIWI Chpt 16
	Instrumental variables	CIWI Chpt 16
	Principal stratification (with examples in compliance)	Optional: Frangakis and Rubin 2002 Biometrics
	Principal stratification (with examples in compliance)	Optional: Pearl 2011 Int. Journal of Biostatistics
	Causal effects of longitudinal interventions	CIWI Chpt 19
	Causal effects of longitudinal interventions	CIWI Chpt 19
	Causal effects of longitudinal interventions	CIWI Chpt 20
	Optimal dynamic treatments	TL2 Chpt 22
	Causal mediation	Valeri and VanderWeele 2013 Psych. Methods
	Causal mediation	Vansteelandt 2017 Epidemiology
	Causal mediation	
	Causal sensitivity analysis	vanderWeele and Ding 2017 Ann. Int. Med.
	FINAL EXAM	