Causal inference and experimental design

Macartan Humphreys

HU | Spring 2024

1 Class resources

- This syllabus: https://macartan.github.io/teaching/ci_2024/syllabus.pdf

- Slides: https://macartan.github.io/teaching/ci_2024/slides/ci_2024.html

- Puzzles: https://macartan.github.io/teaching/ci_2024/exercises.html

2 Times and locations

Raum: K12B

- Jan 12: 12-14 (short introductory lecture)
- Jan 19: 12-18
- Jan 26: 12-18
- Feb 2: 12-18
- Feb 9: 12-18

The longer sessions will be structured roughly as follows.

- 12:00 - 13:50: lecture
- 13:40 - 14:30: team presentations and discussions
- 14:30 - 15:15: late lunch break
- 15:15 - 17:00: lecture
- 17:00 - 17:10: break
- 17:10 - 18:00: team presentations and discussions

3 Abstract

The course addresses advanced topics in causal inference and experimental design. It is a hands on course in which theoretical results are introduced through lecture, demonstrated in practice, and worked on in groups. Topics likely include sampling and randomization schemes, including multilevel trials, restricted randomization, and patient preference trials; Bayesian approaches to causal inference, exact inference: sharp nulls for complex hypotheses, inverting hypothesis tests, Mediation analysis, Multiple comparisons, Open science workflows.

4 Readings

All readings are linked from this document.

I will draw material especially from my two books, both open access:

- Research Design in the Social Sciences

- Integrated Inferences
- Hernan/Robins: What if

Recommended non open access readings include:

- Green and Gerber: Field Experiments with supplementary material available here: https://isps.yale.edu/FEDAI

5 Prerequisites

You should already have background in statistics and fundamentals of causal inference.

In addition you should know some `R`. Really, the more you can invest on getting on top of `R` before the class the better.

- Resources for learning `R`: http://www.r-bloggers.com/how-to-learn-r-2/

Please make sure your `R` is up-to-date and that you are working in `R studio`. Then make sure you have the following packages installed.

```
pacman::p_load(
  rstan,
  dagitty,
  DeclareDesign,
  CausalQueries,
  ri2
  )
```

5.1 File Sharing via Git

I encourage you to set up and send me your git user names and we can access all materials on github.

5.2 Writing with `Rmd` or `qmd`

Please plan to do drafting in `Quarto`. This is a simple markup language that lets you integrate writing and coding. This document is written in `quarto` and the slides will be also.

The key thing is that you can insert code chunks like this.

```
# Define a random number
x <- rnorm(1)
```

which are run as the document compiles, and can be accessed as needed, like this: we just sampled the random number $x = 0.5565708$.

I recommend using `Rstudio` as an editor. More information here: https://quarto.org/docs/get-started/hello/rstudio.html

6 Modules and Readings

Day 1: Intro

- 1.1 Course outline, tools,
- 1.2 Introduction to Declare design

DD, ch 2 DD, ch 13

Day 2: Causality

- 2.1 Fundamental problems and basic solutions II ch 2, [Holland] Hernan Robbins Ch 6
- 2.2 General inquiries and causal identification See also: Hernan Robbins (Section 3.1, 7.2, 8.4, 10.1)

II ch 4 * DD, ch 7

Day 3: Estimation and Inference

- 3.1 Frequentist

@Lin

- 3.2 Bayesian
- **rstan** Getting started
- II ch 5

Day 4:

- 4.1 Experimental Design
- DD, ch 7
- 4.2 Design evaluation
- DD, ch 7

Day 5:

- 5.1 Topics and techniques
- 5.2 Open science