PLSC 40601

Week 1: Course orientation.

Molly Offer-Westort

Department of Political Science,
University of Chicago

Winter 2023

Overview

This course:

- Objectives.
- Course structure.
- Assignments.

Link to syllabus.

Final assignment

- Possibility of tutorial.
  - Illustrate what problem the tool is addressing, what an example use case is.
  - Still needs some amount of narrative contextualizing the problem.
  - Posted on class github.
  - Python, R, or Stata; **compiled report**.
  - Limited support with debugging.

Link to website. Link to tutorial.

How to read methods papers (Murphy, 1997)

Consider:

- Who the authors are and where the article is published.
  - Econometrica? AER? Annual Review of Economics?
  - PNAS? Science? Nature?
  - JRSS B? JASA, Statistical Science?
  - SSRN, arXiv, PsyArXiv?
  - Machine learning conference proceedings?
  - Political science methods, biostatistics journal?

How to read methods papers

Consider:

- What is the problem/gap/contribution?
- What is the background/methodological context for the paper?
  - Refer to references where it may be more clearly stated; textbooks are often clearer than original articles for understanding a new methodology.
  - Check out the bibliographic notes in a related section of Hastie et al. (2009), The Elements of Statistical Learning. (If you don't have this book already, download it right now.)
- (This is the role of final paper.)

How to read methods papers

Consider:

- How would you explain this paper to someone else?
- What is a use case for the tools presented in the paper?
- (This is the role of paper presentation/discussion)

How to read methods papers

Consider:

- How would you use the presented method on data?
  - Many of the papers have associated packages.

- When using real data, what are limitations of the assumptions required by the method?

- (Tutorial?)

How to read methods papers

Consider:

- Annotated bibliography?

How to read methods papers

References for notation:

- Hastie et al. (2009)
- Glossary of Aronow and Miller (2019) Foundations of Agnostic Statistics
- Other suggestions?

Potential outcomes framework.

Statistical setup.

- Population.
- Sample, units indexed $i = 1, \dots, N$.
- Observed outcome, $Y_i \in \mathbb{R}$;
- Treatments, $W_i \in \{0, \dots, K\}$; (or Z_i, or D_i, A_i,...)
  - What do we mean by treatments?
  - “No causation without manipulation.” (Holland, 1986)
- Covariates, $X_i \in \mathbb{R}^p$.

Statistical setup.

- Potential outcomes framework: $Y_i(w)$ represents the potential outcome for respondent i under treatment w.
  $\mathrm{Y}_i = (Y_i(1), \ldots, Y_i(K))$.

- $Y_i(1)$ is the outcome we would see only when individual i receives treatment 1.

- Alternatively, (for binary treatment, $W_i \in \{0, 1\}$)
  - Y_i^1, Y_i^0;
  - $Y_i^{w=1}, Y_i^{w=0}$;
  - Y_{i1}, Y_{i0};
  - $Y | \mathrm{do}(W = 1), Y | \mathrm{do}(W = 0)$

(Some) Causal estimands

What is an estimand? What makes an estimand *causal*?
Counterfactual comparison.

- Individual treatment effect:

$$\tau_i = Y_i(1) - Y_i(0)$$

- Average treatment effect (ATE):

$$\tau = \mathrm{E}\left[\tau_i\right]$$

  - Expectation over what?

  $\rightarrow$ The population we previously defined.

(Some) Causal estimands

- Average treatment effect on the treated (ATT):

$$\tau_{ATT} = \mathrm{E}\left[\tau_i | W_i = 1\right]$$

  - How is this different from the ATE?
  - Why might we care about this?

- Conditional average treatment effect (CATE):

$$\tau_{CATE} = \mathrm{E}\left[\tau_i | X_i = x\right]$$

Many others!

Identification

- $Y_i(0), \ldots, Y_i(K)$ are *potential* outcomes; typically, we only get to see an individual outcome under one version of treatment.

- **Fundamental problem of causal inference**: we can't see counterfactual potential outcomes for a given unit *at the same time*.

- How do we move from what we observe to what we would like, ideally, to measure?

Identification

- Identification: We call a parameter **identifiable** if in the case that we had infinite data, we could approximate the true parameter value to arbitrary precision.
  - e.g., if you have **infinite data** randomly sampled from a distribution, by taking the empirical mean, you get the mean of that distribution with arbitrary precision.
  - If we're dealing with a **finite population**, we can think about identifying a target quantity about that population, as if we were to repeat the procedure through we observed data about that population, averaging *across repetitions*, we could approximate the true quantity to arbitrary precision.
  - We can consider parameters to be **point identified** or **interval identified**.

Identification

- **Consistency/ Stable Unit Treatment Value Assumption (SUTVA)**: what we observe is interpretable in terms of potential outcomes.
  - no unobserved multiple versions of the treatment
  - no “interference between units”

Identification: randomization

- When we aren't in control of assigning treatment, we say the data is **observational.**

- SUTVA is not enough to get us to identification with observational data.

- **Random assignment** gives us:
  - $(Y_i(1), Y_i(0)) \perp\!\!\!\perp W_i$ (independence of potential outcomes and treatment)
  - $0 < \Pr[W_i = 1] < 0$ (positivity)

- In this class, we will not spend a lot of time on identification, but it may be worth considering how various methods fare when these assumptions are violated.

Causal inference as a missing data problem

Units	Covariates X_i	Treatment W_i	$Y_i(1)$	$Y_i(0)$	Observed Y_i
1	1	1	1	?	1
2	0	1	?	0	0
3	1	0	0	?	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
N	0	0	?	0	0

References I

Aronow, P. M. and Miller, B. T. (2019). *Foundations of agnostic statistics*. Cambridge University Press.

Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81(396):945–960.

Murphy, J. R. (1997). How to read the statistical methods literature: a guide for students. *The American Statistician*, 51(2):155–157.