Segment 1: Fundamentals of Causal Inference

Section 06: Average Causal Effects

Cut to Rubin 1974

Required reading

Journal of Educational Psychology
1974, Vol. 66, No. 5, 688–701

ESTIMATING CAUSAL EFFECTS OF TREATMENTS IN RANDOMIZED AND NONRANDOMIZED STUDIES[1]

DONALD B. RUBIN[2]

Educational Testing Service, Princeton, New Jersey

A discussion of matching, randomization, random sampling, and other methods of controlling extraneous variation is presented. The objective is to specify the benefits of randomization in estimating causal effects of treatments. The basic conclusion is that randomization should be employed whenever possible but that the use of carefully controlled nonrandomized data to estimate causal effects is a reasonable and necessary procedure in many cases.

Average Causal Effects

- The fundamental problem means we cannot estimate *individual-level causal effects* of the form:

$$\tau_{ind} =$$

- But we may wish to target the *sample average causal effect*:

$$\tau_{SATE} =$$

- This is an example of a *causal estimand*

Vocab Reminder: *Estimand*

- An *estimand* is a target quantity we wish to learn about
  - A *causal estimand* is a quantity that represents a causal effect
- An *estimator* is a procedure (e.g., a formula) for providing a numerical *estimate* of an *estimand*

Common Causal Estimands

Sample Average Treatment Effect (SATE):

$$\tau_{SATE} =$$

the average of individual-level causal effects within the observed sample

Conditional average treatment effect (CATE):

$$\tau_{CATE|X} =$$

the SATE among a *subset* of the sample

- E.g., the SATE among people aged 50 in the sample

Common Causal Estimands

Population Average Treatment Effect (PATE):

$$\tau_{PATE} =$$

the average of individual-level causal effects within the population.

If the study sample is a representative sample of the population, then any unbiased estimate of τ_{SATE} is also unbiased for τ_{PATE}.

The difference generally relates to the fact that, for τ_{PATE} we have to account for the fact that we observe *neither* potential outcome for some members of the population.

Randomized Experiment

Randomize 1000 patients to receive vaccine ($Z = 1$) or placebo ($Z = 0$)

- No infections in the vaccine arm
- 30% of the patients in the control arm become infected
- Use *average outcome* in placebo recipients as a "close substitute" for *average outcome* in vaccine recipients
  - Observed placebo outcomes represent what *would have happened* to the vaccine recipients
- Conclude that the vaccine *causes* a person to remain HIV free

This is an example of τ_{SATE}:

Causal Estimands
Measures of Causal Effects

- Functions of potential outcomes define causal effects that we may wish to measure or estimate
- In a statistical sense, these are the estimands - the targets for inference
- Other examples of **causal estimands:**
  - Causal risk ratio: $\frac{Pr(Y^t=1)}{Pr(Y^c=1)}$
  - Causal odds ratio: $\frac{Pr(Y^t=1)/Pr(Y^t=0)}{Pr(Y^c=1)/Pr(Y^c=0)}$
- Could also be defined conditional on some attributes: $E[Y^t|X=x] - E[Y^c|X=x]$
- Average causal effects are most common, but could define effects based on other functions (quantiles, variances, hazards, etc.) of potential outcomes