

Statistics 256: Causality

In-class Activity 6: Selection on Observables — Estimation

Name(s): _____

Question 1: One target, many roads

Under conditional ignorability, every estimator we covered (subclassification, matching, PS-matching, IPW, balancing weights, regression imputation) targets the same population object:

$$\mathbb{E}_{X \sim F} [\mu_1(X) - \mu_0(X)], \quad \text{where } \mu_d(X) = \mathbb{E}[Y \mid X, D=d].$$

- (a) Explain in 1–2 sentences how this object is *akin to conditioning on X* : what is the distribution of X over which treated and controls get compared here, and how does that differ from the plain difference in means $\mathbb{E}[Y \mid D=1] - \mathbb{E}[Y \mid D=0]$?

- (b) If we set $F = p(X \mid D = 1)$, what estimand does this object target?

Question 2: Does any of this fix CI/SOO?

Suppose there is an unobserved confounder $U \notin X$. Which of the following methods deliver an unbiased $\hat{\tau}$? *Tick all that apply.*

- ☐ IPW with a correctly specified propensity score on X
- ☐ Entropy balancing on all observed X
- ☐ Lin regression with full $D \times X$ interactions
- ☐ Doubly-robust AIPW
- ☐ None of the above

In one sentence, why?

Question 3: ATT vs ATE

Your control pool is huge and very different from your treated units (think LaLonde: 2,490 PSID controls, 185 NSW treated). In 1–2 sentences, why might ATT be easier to estimate well than ATE here?

Question 4: Why doesn't matching always work?

With more than one continuous variable in the mix, even after carefully picking the nearest control to each treated unit on X , matching can give a biased $\hat{\tau}$. Why? (There is one core reason, and one thing that amplifies it quickly in higher dimensions.)

Question 5: Method-to-data fit

For each setting, name a method you'd reach for first, and one you'd avoid. One sentence each.

(a) X is two binary covariates; $n = 500$ in each arm.

(b) X has 30 continuous covariates; $n_1 = 150$, $n_0 = 5,000$; modest overlap.

(c) X has good overlap, but Y 's dependence on X is strongly nonlinear and you don't know the form.