ECON220B Discussion Section 5

Selection on Observables

Lapo Bini

Roadmap

1. From RCT to Selection on Observables
2. Inverse Probability Weighting
3. Regression Adjustment Estimator
4. Propensity Score and Heavy-Tailed Distributions

Potential Outcome Framework

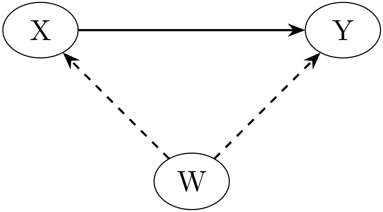
- Treatment: $x_i = \mathbb{1}\{treated\}$
- Observed outcome: $y_i = x_i y_i(1) + (1 - x_i) y_i(0)$
- RCT assumption: $x_i \perp (y_i(1), y_i(0))$

RCT: randomization equalizes everything other than the treatment in the treatment and control group. Fine for randomized experiment, what about observational studies?

Cancer & Smoking

"*Considerable propaganda is now being developed to convince the public that cigarette smoking is dangerous.*" - **Sir. Ronald Fisher (1958)**

What was his argument?



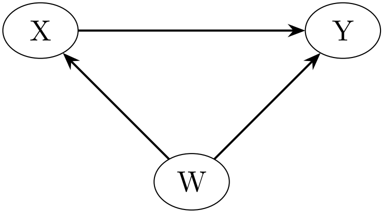
"*[...] run the risk of failing to recognize, and therefore failing to prevent, other and more genuine causes*"

Different Identification Strategies

Starting point is conditional independence, or uncounfoundedness, or selection on observables:

$$x_i \perp (y_i(1), y_i(0)) | w_i$$

i.e. all confounders have already been identified and accounted for by the set of covariates. Then two possible ways to estimate $\tau_{ATE} \equiv E[y_i(1) - y_i(0)]$



Inverse Probability Weighting (1/2)

- We model how the treatment take-up decision x_i is related with the covariates w_i, re-weighting each observation by the likelihood of receiving the treatment.
- Overlap condition: $0 < P(x_i = 1 | w_i = \mathbf{w}) < 1$, i.e. for a particular characteristic $w_i = \mathbf{w}$ if we observe some treated unit, then we should be able to observe some untreated unit as well.
- $E\left[\frac{x_i y_i}{e(w_i)}\right] =$

Inverse Probability Weighting (2/2)

- The propensity score $e(w_i) \equiv P(x_i = 1 | w_i = \mathbf{w})$ is a balancing score: after conditioning on the propensity score, the distribution of the treatment is the same for treated and untreated:

$$x_i \perp w_i | e(w_i)$$

- $e(w_i)$ is all you need to know: sufficient statistic for x_i.

More About Overlap Condition

Note: $0 < P(x_i = 1 | w_i = \mathbf{w}) < 1 \iff f_{w|x_i=1}(\mathbf{w}) > 0, f_{w|x_i=0}(\mathbf{w}) > 0$

Proof

Regression Adjustment Estimator

Do we really need overlap condition? No $\rightarrow$ linearity assumption: interaction effect of covariates and treatment

$$\tau_{ATE} = E[y_i(1) - y_i(0)] = E[y_i(1)] - E[y_i(0)]$$
$$\tau_{ATE} = E[E[y_i(1)|w_i] - E[y_i(0)|w_i]]$$
$$\tau_{ATE} = E[E[y_i(1)|w_i, x_i = 1] - E[y_i(0)|w_i, x_i = 0]]$$
$$\tau_{ATE} = E[E[y_i|w_i, x_i = 1] - E[y_i|w_i, x_i = 0]]$$
$$\tau_{ATE} = E[g_1(w_i) - g_0(w_i)] = E[w_i^T \delta_1 - w_i^T \delta_0] = \mu_w^T(\delta_1 - \delta_0)$$

Propensity Score and Heavy-Tailed Distributions

Consider the inverse probability weighting estimator:

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \frac{x_i y_i}{e(w_i)}$$

where x_i is the binary indicator of treatment status, y_i is the outcome variable, and w_i represents the covariates. For simplicity, we assume that the propensity score, $e(wi) = P[x_i = 1 | w_i]$, is known. In addition, assume y_i is bounded

Heavy-Tailed Distributions

Some useful definitions:

- A distribution is heavy-tailed if $E[tX] = \infty \; \forall t > 0$
- A distribution is light-tailed if it is not heavy-tailed.
- A distribution is light-tailed if $E[X^k] < \infty \; \forall t > 0$

Propensity Score and Heavy-Tailed Distributions

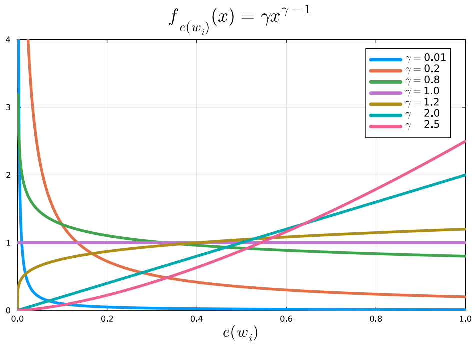
Q.1 - Assume strong overlap, show that $E\left[\left|\frac{x_i y_i}{e(w_i)}\right|^2\right] < \infty$ then conclude $\mathrm{Var}\left(\hat{\theta}\right) < \infty$.

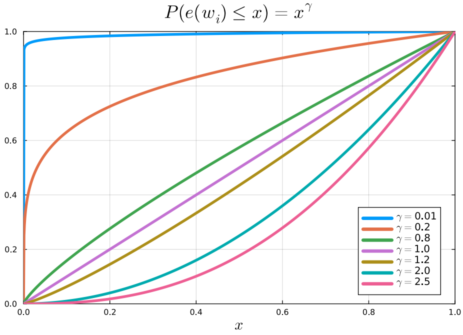
Propensity Score and Heavy-Tailed Distributions

Q.2 - Now assume the propensity score can be arbitrarily close to zero $P\left(e(w_i) \leq \delta\right) = \delta^{\gamma}$. What will happen if γ is small?

Propensity Score and Heavy-Tailed Distributions

Q.3 - Take $\gamma = 2$. Plot the density function of the propensity score. Is this sufficient to show that $E\left[\left|\frac{x_i y_i}{e(w_i)}\right|^2\right] < \infty$?





Propensity Score and Heavy-Tailed Distributions

Q.3 - Take $\gamma = 2$. Plot the density function of the propensity score. Is this sufficient to show that $E\left[\left|\frac{x_i y_i}{e(w_i)}\right|^2\right] < \infty$?

Propensity Score and Heavy-Tailed Distributions

Q.4 - Discuss what will happen if $\gamma < 1$.