Lecture 6: Matching

Naijia Liu

Feb. 13, 2025

Problem Set I

- Due on 2/20 11:59pm, late submission due on 2/22 11:59pm
- `Tidyverse` or Base `R`

Problem Set I

- Due on 2/20 11:59pm, late submission due on 2/22 11:59pm
- `Tidyverse` or Base `R`
- **Cite** your source!
  - Superman helped me in coming up with the way to code this new variable.
  - Stackoverflow answer gave me the idea for this answer (with link).

Assumptions

- Probabilistic treatment: **Not deterministic that you will (or will never) be invited to the show.**

$$0 < \Pr(T_i = 1|X_i) < 1$$

See more of these reviews for lecture 1 slides.

Assumptions

- Probabilistic treatment: **Not deterministic that you will (or will never) be invited to the show.**

$$0 < \Pr(T_i = 1|X_i) < 1$$

- Ignorability: **If we were to know all covariates of candidates, invitation is almost randomly sent.**

$$Y_i(1), Y_i(0) \perp\!\!\!\perp T_i|X_i$$

See more of these reviews for lecture 1 slides.

Assumptions

- Probabilistic treatment: **Not deterministic that you will (or will never) be invited to the show.**

$$0 < \Pr(T_i = 1|X_i) < 1$$

- Ignorability: **If we were to know all covariates of candidates, invitation is almost randomly sent.**

$$Y_i(1), Y_i(0) \perp\!\!\!\perp T_i|X_i$$

- SUTVA: **Other people's donation does not affect your donation amount. One single version of treatment level.**

$$Y_i(T_i, Y_i(T_i')) = Y_i(T_i) \quad \forall i \neq i'$$
$$Y_i(T_i) = Y_i(T_i') \text{ if } T_i = T_i'$$

See more of these reviews for lecture 1 slides.

Identifying assumptions

- ⋆ **Valid matched set**: **On average, donation for those invited would have been the same to those in the matched set (not invited), if they were not invited.**

$$\underbrace{\mathbb{E}(\, Y_i(0) | \, T_i = 1)}_{\text{unobserved}} = \mathbb{E}(\, Y_i(0) | \, T_i = 0, i \in \mathcal{M})$$

Proof

Let's start with the assumption above.

$$\mathbb{E}_S(\, Y_i(0) | \, T_i = 1) = \mathbb{E}_S(\, Y_i(0) | \, T_i = 0, i \in \mathcal{M})$$

$$-\mathbb{E}_S(\, Y_i(0) | \, T_i = 1) = -\mathbb{E}_S(\, Y_i(0) | \, T_i = 0, i \in \mathcal{M}) \quad \text{multiple by -1}$$

Proof

Let's start with the assumption above.

$$\mathbb{E}_S(\, Y_i(0) | \, T_i = 1) = \mathbb{E}_S(\, Y_i(0) | \, T_i = 0, i \in \mathcal{M})$$
$$-\mathbb{E}_S(\, Y_i(0) | \, T_i = 1) = -\mathbb{E}_S(\, Y_i(0) | \, T_i = 0, i \in \mathcal{M}) \quad \text{multiple by -1}$$

Let's add $E(\, Y_i(1) | \, T_i = 1)$ to both sides of the equation.

$$\underbrace{\mathbb{E}(\, Y_i(1) - \, Y_i(0) | \, T_i = 1)}_{\text{Average treatment effect on the treated, unobserved}}$$

$$=$$

$$\underbrace{\mathbb{E}_S(\, Y_i(1) | \, T_i = 1) - \mathbb{E}_S(\, Y_i(0) | \, T_i = 0, i \in \mathcal{M})}_{\text{Observed from data}}$$

Mediators

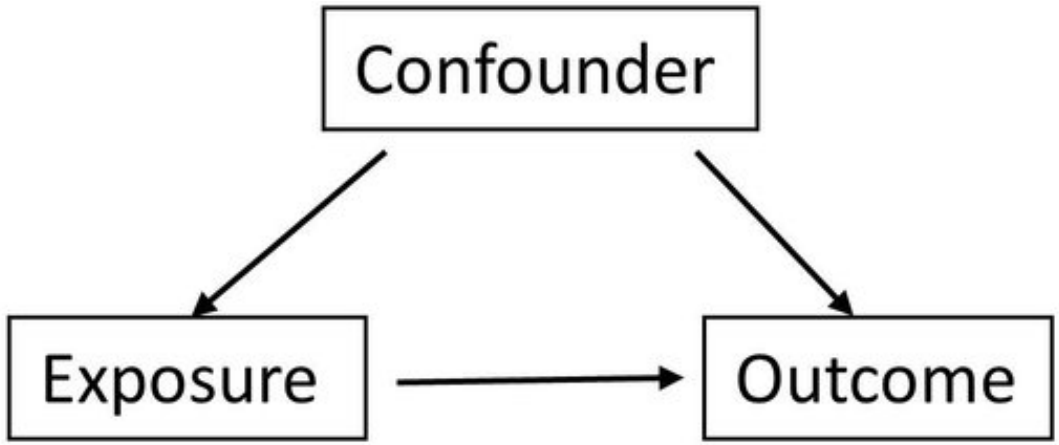
- Critique on matching method assumptions.
  - What if observables are not enough?

    Aside from incumbency, party, previous donations, what are other variables that might matter?
  - Mediators?

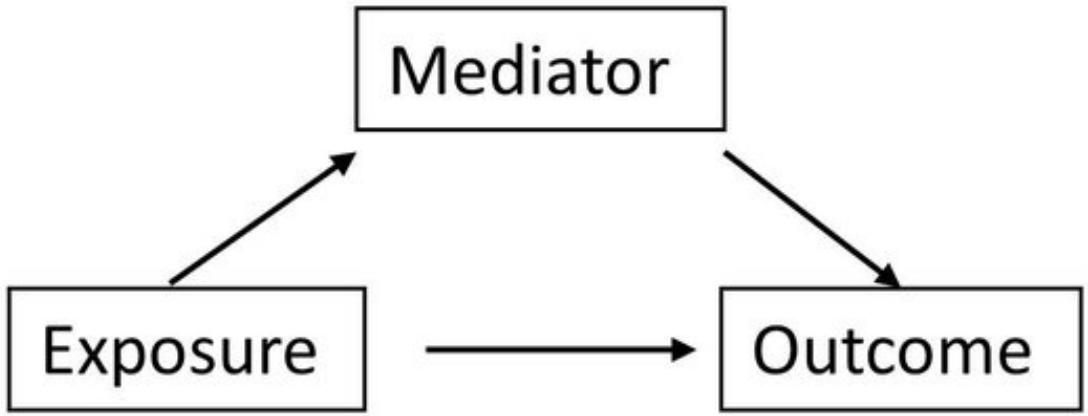
    Intermediate variable between the treatment and the outcome.

Mediators

(A) Confounding



(B) Mediation



Colber Effect

- What are the possible mediators between appearance in the show and donation?
  - Connectivity

    Appearance leads to better network and potential new donors.

Colber Effect

- What are the possible mediators between appearance in the show and donation?
  - Connectivity

    Appearance leads to better network and potential new donors.
- Matching method with above assumptions cannot match on mediators. And this is problematic when:
  - The mediators are different for treated and matched units.
  - We want to know the **direct** treatment effect of appearance on donation.

`MatchIt` Package

- Compare balance between and after matching

`MatchIt` Package

- Compare balance between and after matching
- Coarsened exact matching
  - Age groups instead of age.

`MatchIt` Package

- Compare balance between and after matching
- Coarsened exact matching
  - Age groups instead of age.
- Mahalanobis distance matching.