Causal Inference1

Weiheng Zhang

2022-04-15

Causality, Potential outcomes and the estimation of treatment effects in randomized studies

- Goal: policy/program evaluation s to assess the causal effect of policy interventions.

Basic concepts

- Treatment

$$D_i = \begin{cases} 1 \text{ if unit i received the treatment} \\ 0 \text{ otherwise} \end{cases}$$

- Outcome

  Y_i: observed outcome variable of interest for unit i

- Potential outcomes

  Y_{0i} and Y_{1i}: Potential outcomes for unit i

$$\begin{aligned} Y_{1i} &: \text{Potential outcome for unit i with treatment} \\ Y_{0i} &: \text{Potential outcome for unit i without treatment} \end{aligned}$$

- Treatment effect

  The treatment effect or causal effect of the treatment on the outcome for unit i is the difference between its two potential outcomes.

$$Y_{1i} - Y_{0i}$$

- Observed outcomes

$$Y_i = Y_{1i}D_i + Y_{0i}(1 - D_i) \text{ or } Y_i \begin{cases} Y_{1i} \text{ if } D_i = 1 \\ Y_{0i} \text{ if } D_i = 0 \end{cases}$$

- **Fundamental problem of causal inference**:

  Cannot observe both potential outcomes (Y_{0i}, Y_{1i}).

- Assumption: Stable Unit Treatment Value Assumption (SUTVA)

  Observed outcomes are realized as

$$Y_i = Y_{1i}D_i + Y_{0i}(1 - D_i)$$

Remark:

- Implies that potential outcomes for unit i are unaffected by the treatment of unit j. (Rule s out interference across units).
- Counterexample: effect of flu vaccine on hospitalization.

- ATE

  Average treatment effect is:

$$\alpha_{ATE} = E[Y_1 - Y_0]$$

- ATET

  Average treatment effect on the treated is:

$$\alpha_{ATE} = E[Y_1 - Y_0|D = 1]$$

- Example:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	?	3	1	?
2	1	?	1	1	?
3	?	0	0	0	?
4	?	1	1	0	?

Q: what is α_{ATE}

In Ideal world:

i	Y_{1i}	Y_{0i}	Y_i	D_i	$Y_{1i} - Y_{0i}$
1	3	0	3	1	3
2	1	1	1	1	0
3	1	0	0	0	1
4	1	1	1	0	0

$$\alpha_{ATE} = E[Y_1 - Y_0] = 4/4 = 1$$

$$\alpha_{ATET} = E[Y_1 - Y_0|D = 1] = 3/2 = 1.5$$

- **Problem of potential outcome famework**: Selection Bias

  Comparisons of earnings for the treated and the untreated do not ususally give the right answer:

$$\begin{aligned} E[Y|D = 1] - E[Y|D = 0] &= E[Y_1|D = 1] - E[Y_0|D = 0] \\ &= E[Y_1|D = 1] - E[Y_0|D = 1] + E[Y_0|D = 1] - E[Y_0|D = 0] \\ &= \underbrace{E[Y_1 - Y_0|D = 1]}_{ATET} + \underbrace{E[Y_0|D = 1] - E[Y_0|D = 0]}_{Bias} \end{aligned}$$

- Assignment mechanism

  Assignment mechanism is the procedure that determines which units are selected for treatment intake.

- Sol 1: Identification in randomized experiments

  Randomization implies:

$$(Y_1, Y_0) \perp\!\!\!\perp D$$

  We have that $E[Y_0|D=1] = E[Y_0|D=0]$ and therefore

$$\alpha_{ATET} = E[Y_1 - Y_0|D=1] = E[Y|D=1] - E[Y|D=0]$$

  and

$$\alpha_{ATE} = E[Y_1 - Y_0] = E[Y|D=1] - E[Y|D=0]$$

  As a result:

$$\underbrace{E[Y|D=1] - E[Y|D=0]}_{\text{Difference in Means}} = \alpha_{ATE} = \alpha_A TET$$

  where bias is vanished

- Extended from randomized experiments

  Let $Q_\theta(Y)$ be the θ-th quantile of the distriution of Y:

$$Pr(Y \leq Q_\theta(Y)) = \theta$$

  Given random assignment, $Y_0 \perp\!\!\!\perp D$ therefore

$$Y_0 \sim \quad Y_0|D=1 \sim \quad Y|D=0$$

  similarly we have

$$Y_1 \sim \quad Y|D=1$$

  So effect of the treatment at any quantile, $Q_\theta(Y_1) - Q_\theta(Y_0)$ is identified. Note does not identify the quantiles of the effect: $Q_\theta(Y_1 - Y_0)$.

- Estimation in Randomized Experiments
  - we construct an esitmator:

$$\hat{\alpha} = \bar{Y}_1 - \bar{Y}_0$$

    where

$$\bar{Y}_1 = \frac{\sum Y_i D_i}{\sum D_i} = \frac{1}{N_1} \sum_{D_i=1} Y_i$$
$$\bar{Y}_0 = \frac{\sum Y_i (1 - D_i)}{\sum 1 - D_i} = \frac{1}{N_1} \sum_{D_i=0} Y_i$$

- asympotic distribution

$$\frac{\hat{\alpha} - \alpha_{ATE}}{\sqrt{\frac{\hat{\sigma}_1^2}{N_1} + \frac{\hat{\sigma}_2^2}{N_0}}} \xrightarrow{d} N(0,1)$$

where

$$\hat{\sigma}_1^2 = \frac{1}{N_1 - 1} \sum_{D_i=1} (Y_i - \bar{Y}_1)^2$$

- test
  * Testing in Large Samples: Two sample t-Test

$$H_0 : \alpha_{ATE} = 0$$
$$H_1 : \alpha_{ATE} \neq 0$$

$$t = \frac{\hat{\alpha}}{\sqrt{\frac{\hat{\sigma}_1^2}{N_1} + \frac{\hat{\sigma}_2^2}{N_0}}}$$

  * Testing in Small Samples: Fisher's Exact Test
    + In large sample N, the test of differences in means is:

$$H_0 : E[Y_1] = E[Y_0]$$
$$H_1 : E[Y_1] \neq E[Y_0]$$

    + Fisher's Exact Test with small N:

$$H_0 : Y_1 = Y_0 \text{ shape null}$$
$$H_1 : Y_1 \neq Y_0$$

The null hypothesis means the treatment is similar to placebo.

$$Pr(\hat{\alpha} < z) = \frac{1}{\#\Omega} \sum_{\omega \in \Omega} Q\{\hat{(\omega)}\alpha < z\}$$

- Threats to the Validity of Randomized Experiments
  - Internal validity: can we estimate treatment effect for our particular sample?

    Fails when there are differences between treated and controls (other than the treeatment iteslf) that affect the outcome and that we cannot control for.

    Threats:

    + Failure of randomization
    + Non-complicance with experimental protocol
    + Attrition

- External validity: can we extrapolate our estimates to other populations?

  Fails when the treatment effect is different outside the evaluation envrionment.

  Threats:

  + Non-representative sample
  + Non-representative program
    - The treatment differs in actual implementations
    - Scale effects
    - Actual implementations are not randomzied (nor full scale)
  + Hawthorne effects