Estimating Treatment Effects - Randomization

Februray 18, 2025

Policy Evaluation - Assessing Evidence

- What criteria should we use to assess existing evidence?
  - Strength of the methodology
    - Does the evidence provide a causal interpretation?
    - Is the evidence rooted in the appropriate historical context?
  - Quality of the data
  - External validity

Estimating Treatment Effects

- Goal: Calculate the **average treatment effect** of a policy or intervention on a specified outcome.

Estimating Treatment Effects

- Goal: Calculate the **average treatment effect** of a policy or intervention on a specified outcome.
- Average Treatment Effect (ATE) of an intervention among a sample of n people:

$$ATE = Avg_n[Y_i^1 - Y_i^0]$$

Where Y_i^1 is the *potential* outcome for person i with treatment and Y_i^0 is the *potential* outcome for person i without treatment.

Estimating Treatment Effects

- Suppose we're interested in the effect of a new surgical intervention ($D_i = 1$) for cancer on longevity compared to standard chemotherapy ($D_i = 0$).

Outcomes for ten patients receiving surgery (Y^1) or chemotherapy (Y^0)

Patient	Y^1	Y^0
1	7	1
2	5	6
3	5	1
4	7	8
5	4	2
6	10	1
7	1	10
8	5	6
9	3	7
10	9	8

- Calculate the ATE ($ATE = Avg_n[Y_i^1 - Y_i^0]$).

Estimating Treatment Effects

- Goal: Calculate the **average treatment effect** of a policy or intervention on a specified outcome.

- Average Treatment Effect (ATE) of an intervention among a sample of n people:

$$ATE = Avg_n[Y_i^1 - Y_i^0]$$

  Where Y_i^1 is the *potential* outcome for person i with treatment and Y_i^0 is the *potential* outcome for person i without treatment.

- Problem:

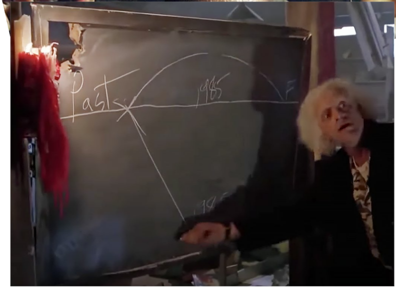
Estimating Treatment Effects

- Goal: Calculate the **average treatment effect** of a policy or intervention on a specified outcome.
- Average Treatment Effect (ATE) of an intervention among a sample of n people:
  $$ATE = Avg_n[Y_i^1 - Y_i^0]$$
  Where Y_i^1 is the *potential* outcome for person i with treatment and Y_i^0 is the *potential* outcome for person i without treatment.
- Problem: We can never obtain the ATE this way because we don't observe the ***counterfactual***.

Estimating Treatment Effects

Estimating Treatment Effects

Estimating Treatment Effects



Estimating Treatment Effects

Estimating Treatment Effects

- Solution?

Estimating Treatment Effects

- Solution? Let's *estimate* the ATE by comparing outcomes for those who are treated and those who aren't treated.

Estimating Treatment Effects

- Solution? Let's *estimate* the ATE by comparing outcomes for those who are treated and those who aren't treated.
- Define treatment status D_i as:

$$D_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{otherwise} \end{cases}$$

Estimating Treatment Effects

- Solution? Let's *estimate* the ATE by comparing outcomes for those who are treated and those who aren't treated.
- Define treatment status D_i as:

$$D_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{otherwise} \end{cases}$$

- Average outcome (Y) conditional on treatment:

Estimating Treatment Effects

- Solution? Let's *estimate* the ATE by comparing outcomes for those who are treated and those who aren't treated.
- Define treatment status D_i as:

$$D_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{otherwise} \end{cases}$$

- Average outcome (Y) conditional on treatment:

  $Avg_n[Y_i^1 | D_i = 1]$

Estimating Treatment Effects

- Solution? Let's *estimate* the ATE by comparing outcomes for those who are treated and those who aren't treated.
- Define treatment status D_i as:

$$D_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{otherwise} \end{cases}$$

- Average outcome (Y) conditional on treatment:

$$Avg_n[Y_i^1 | D_i = 1]$$
$$Avg_n[Y_i^0 | D_i = 0]$$

Estimating Treatment Effects

- Solution? Let's *estimate* the ATE by comparing outcomes for those who are treated and those who aren't treated.
- Define treatment status D_i as:

$$D_i = \begin{cases} 1 & \text{if treated} \\ 0 & \text{otherwise} \end{cases}$$

- Average outcome (Y) conditional on treatment:

$$Avg_n[Y_i^1 | D_i = 1]$$
$$Avg_n[Y_i^0 | D_i = 0]$$

- Estimated Effect of treatment on Y:

$$ATE_{est} = Avg_n[Y_i^1 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0]$$

Estimating Treatment Effects

- Suppose we're interested in the effect of a new surgical intervention ($D_i = 1$) for cancer on longevity compared to standard chemotherapy ($D_i = 0$).

Outcomes for ten patients receiving surgery (Y^1) or chemotherapy (Y^0)

Patient	Y^1	Y^0	D	Y
1	7	1	1	7
2	5	6	0	6
3	5	1	1	5
4	7	8	0	8
5	4	2	1	4
6	10	1	1	10
7	1	10	0	10
8	5	6	0	6
9	3	7	0	7
10	9	8	1	9

- $ATE = ?$
- $ATE_{est} = Avg_n[Y_i^1 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0] = ?$

Estimating Treatment Effects

- $ATE_{est} = Avg_n[Y_i^1 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0]$
- When does $ATE_{est} = ATE$ and when does it not?

Estimating Treatment Effects

- $ATE_{est} = Avg_n[Y_i^1 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0]$
- When does $ATE_{est} = ATE$ and when does it not?
  - Depends on whether treatment assignment is independent of potential outcomes.

Estimating Treatment Effects

- $ATE_{est} = Avg_n[Y_i^1 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0]$

- When does $ATE_{est} = ATE$ and when does it not?
  - Depends on whether treatment assignment is independent of potential outcomes.
    - $(Y^1, Y^0) \perp\!\!\!\perp D$

Estimating Treatment Effects

- $ATE_{est} = Avg_n[Y_i^1 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0]$

- When does $ATE_{est} = ATE$ and when does it not?
  - Depends on whether treatment assignment is independent of potential outcomes.
    - $(Y^1, Y^0) \perp\!\!\!\perp D$

- If independence is violated, then treatment effects differ by treatment status:

Estimating Treatment Effects

- $ATE_{est} = Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0]$
- When does $ATE_{est} = ATE$ and when does it not?
  - Depends on whether treatment assignment is independent of potential outcomes.
    - $(Y^1, Y^0) \perp\!\!\!\perp D$
- If independence is violated, then treatment effects differ by treatment status:
  - Average Treatment Effect for the Treated:

$$ATT = Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 1]$$

Estimating Treatment Effects

- $ATE_{est} = Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0]$
- When does $ATE_{est} = ATE$ and when does it not?
  - Depends on whether treatment assignment is independent of potential outcomes.
    - $(Y^1, Y^0) \perp\!\!\!\perp D$
- If independence is violated, then treatment effects differ by treatment status:
  - Average Treatment Effect for the Treated:

$$ATT = Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 1]$$

  - Average Treatment Effect for the Untreated:

$$ATU = Avg_n[Y_i^1|D_i = 0] - Avg_n[Y_i^0|D_i = 0]$$

Estimating Treatment Effects

- Suppose we're interested in the effect of a new surgical intervention ($D_i = 1$) for cancer on longevity compared to standard chemotherapy ($D_i = 0$).

Outcomes for ten patients receiving surgery (Y^1) or chemotherapy (Y^0)

Patient	Y^1	Y^0	D	Y
1	7	1	1	7
2	5	6	0	6
3	5	1	1	5
4	7	8	0	8
5	4	2	1	4
6	10	1	1	10
7	1	10	0	10
8	5	6	0	6
9	3	7	0	7
10	9	8	1	9

- $ATE = ?$
- $ATE_{est} = Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0] = ?$
- $ATT = Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 1] = ?$
- $ATU = Avg_n[Y_i^1|D_i = 0] - Avg_n[Y_i^0|D_i = 0] = ?$

Estimating Treatment Effects

- With no independence:

$$ATE = \pi ATT + (1 - \pi) ATU$$

where π is the share of the sample receiving treatment.

Estimating Treatment Effects

- Suppose we're interested in the effect of a new surgical intervention ($D_i = 1$) for cancer on longevity compared to standard chemotherapy ($D_i = 0$).

Outcomes for ten patients receiving surgery (Y^1) or chemotherapy (Y^0)

Patient	Y^1	Y^0	D	Y
1	7	1	1	7
2	5	6	0	6
3	5	1	1	5
4	7	8	0	8
5	4	2	1	4
6	10	1	1	10
7	1	10	0	10
8	5	6	0	6
9	3	7	0	7
10	9	8	1	9

- $ATE = ?$
- $ATE_{est} = ?$
- $ATT = ?$
- $ATU = ?$
- $ATE = ATT(\pi) + ATU(1 - \pi) = ?$

Estimating Treatment Effects

- With no independence:

$$ATE = \pi ATT + (1 - \pi) ATU$$

Estimating Treatment Effects

- With no independence:

$$ATE = \pi ATT + (1-\pi)ATU$$

- Do some algebra:

$$ATE_{est} = ATE + \{Avg_n[Y_i^0|D_i = 1] - Avg_n[Y_i^0|D_i = 0]\}$$
$$+(1-\pi)(ATT - ATU)$$

Estimating Treatment Effects

- With no independence:

$$ATE = \pi ATT + (1-\pi)ATU$$

- Do some algebra:

$$ATE_{est} = ATE + \underbrace{\{Avg_n[Y_i^0|D_i=1] - Avg_n[Y_i^0|D_i=0]\}}_{\text{Selection Bias}}$$

$$+ \underbrace{(1-\pi)(ATT - ATU)}_{\text{Heterogeneous Treatment Effect Bias}}$$

Estimating Treatment Effects - Selection Bias

- Some notes on selection bias:

Estimating Treatment Effects - Selection Bias

- Some notes on selection bias:
  - Sometimes called *omitted variable bias*.

Estimating Treatment Effects - Selection Bias

- Some notes on selection bias:
  - Sometimes called *omitted variable bias*.
  - Problem because selection into treatment depends on factors that are correlated with potential outcomes.

Estimating Treatment Effects - Selection Bias

- Some notes on selection bias:
  - Sometimes called *omitted variable bias*.
  - Problem because selection into treatment depends on factors that are correlated with potential outcomes.
  - To the extent that we can observe and control for these correlates, then our estimates are free from bias (e.g., economy).

Estimating Treatment Effects - Selection Bias

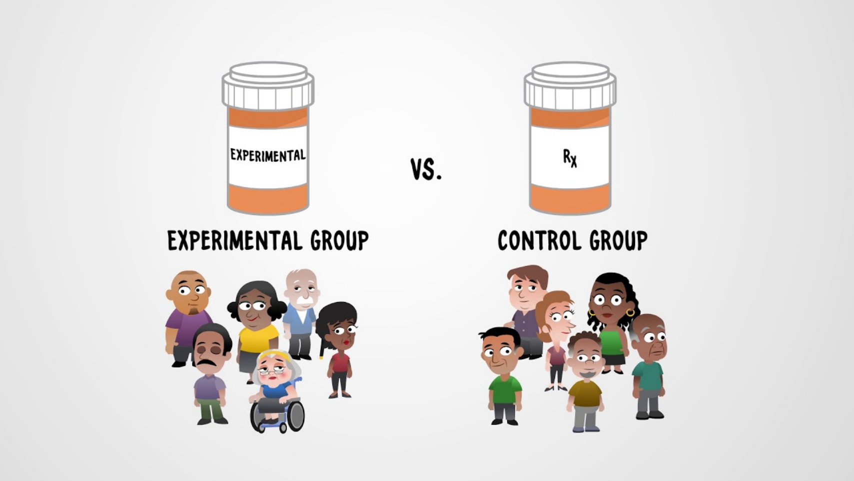
- Some notes on selection bias:
  - Sometimes called *omitted variable bias*.
  - Problem because selection into treatment depends on factors that are correlated with potential outcomes.
  - To the extent that we can observe and control for these correlates, then our estimates are free from bias (e.g., economy).
  - If these correlates are unobserved, then our estimates will be biased (e.g., risk aversion).

Estimating Treatment Effects - Randomization & Selection Bias

- So how do we mitigate selection bias?

Estimating Treatment Effects - Randomization & Selection Bias

- So how do we mitigate selection bias?



Estimating Treatment Effects - Randomization & Selection Bias

- Why does randomization mitigate omitted variable bias?

Estimating Treatment Effects - Randomization & Selection Bias

- Why does randomization mitigate omitted variable bias?
- If D is randomly assigned, then $(Y^1, Y^0) \perp\!\!\!\perp D$:
  - 1. $Avg_n[Y_i^0|D_i = 1] = Avg_n[Y_i^0|D_i = 0]$
  - 2. $ATT = ATU$

$$ATE_{est} = ATE + \underbrace{\{Avg_n[Y_i^0|D_i = 1] - Avg_n[Y_i^0|D_i = 0]\}}_{\text{Selection Bias} = 0}$$

$$+ \quad \underbrace{(1 - \pi)(ATT - ATU)}_{\text{Heterogeneous Treatment Effect Bias} = 0}$$

Estimating Treatment Effects - Randomization & Selection Bias

- Why does randomization mitigate omitted variable bias?
- If D is randomly assigned, then $(Y^1, Y^0) \perp\!\!\!\perp D$:
  - 1. $Avg_n[Y_i^0 | D_i = 1] = Avg_n[Y_i^0 | D_i = 0]$
  - 2. $ATT = ATU$

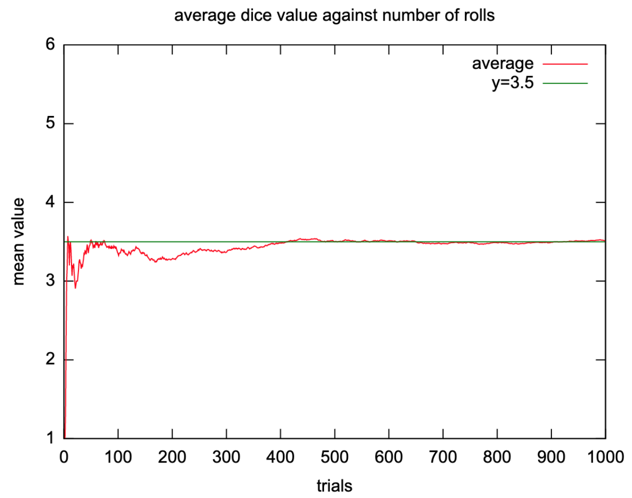
$$ATE_{est} = ATE + \underbrace{\{Avg_n[Y_i^0 | D_i = 1] - Avg_n[Y_i^0 | D_i = 0]\}}_{\text{Selection Bias} = 0}$$

$$+ \quad \underbrace{(1 - \pi)(ATT - ATU)}_{\text{Heterogeneous Treatment Effect Bias} = 0}$$

- **IMPORTANT:** This only works when the sample is "sufficiently large".

Estimating Treatment Effects - Randomization & Selection Bias

- Law of Large Numbers:



Estimating Treatment Effects - Randomization & Selection Bias

- Randomization and the LLN are even more powerful than they seem.

Estimating Treatment Effects - Randomization & Selection Bias

- Randomization and the LLN are even more powerful than they seem.

- Sample Average Treatment Effect = $Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0]$

Estimating Treatment Effects - Randomization & Selection Bias

- Randomization and the LLN are even more powerful than they seem.
- Sample Average Treatment Effect = $Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0]$
- If n is "sufficiently large", then we can replace conditional *averages* with conditional *expectations*.

- Randomization and the LLN are even more powerful than they seem.
- Sample Average Treatment Effect = $Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0]$
- If n is "sufficiently large", then we can replace conditional *averages* with conditional *expectations*.
- Population Average Causal Effect = $E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 0]$

- Randomization and the LLN are even more powerful than they seem.
- Sample Average Treatment Effect = $Avg_n[Y_i^1|D_i = 1] - Avg_n[Y_i^0|D_i = 0]$
- If n is "sufficiently large", then we can replace conditional *averages* with conditional *expectations*.
- Population Average Causal Effect = $E[Y_i^1|D_i = 1] - E[Y_i^0|D_i = 0]$
- Note that "sufficiently large" depends on the population mean and standard deviation of the outcome of interest.

Estimating Treatment Effects - Randomization & Selection Bias

- So randomization is the ideal and the "gold standard" of research methodologies.

Estimating Treatment Effects - Randomization & Selection Bias

- So randomization is the ideal and the “gold standard” of research methodologies.
- But...in policy analysis it’s often (though not always) impossible to randomize individuals in our sample into treatment?

- So randomization is the ideal and the "gold standard" of research methodologies.
- But...in policy analysis it's often (though not always) impossible to randomize individuals in our sample into treatment?
- How can we attempt to identify causal effects of policy when randomization is infeasible?