Thinking Causally

Experimental Research

Prof. Yamil Velez (9/13/22)

Last class

- Definitions of experiments
- Key concepts
  - Units
  - Treatments
  - Outcomes
  - Randomization
- Causation
  - Difference in potential outcomes
- Fundamental problem of causal inference

Neyman-Rubin causal model

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	1	0	1	1
2	1	1	0	0
3	0	0	0	0
4	0	1	1	-1

Neyman-Rubin causal model

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	1	0	1	1
2	1	1	0	0
3	0	0	0	0
4	0	1	1	-1

Neyman-Rubin causal model

	Happiness when assigned to breakfast $Y_i(1)$	Happiness when not assigned to breakfast $Y_i(0)$	Treatment assignment T_i	Individual treatment effect τ_i
1	90	85	1	5
2	80	90	0	-10
3	85	85	0	0
4	100	0	1	100

Neyman-Rubin causal model

	Happiness when assigned to breakfast $Y_i(1)$	Happiness when not assigned to breakfast $Y_i(0)$	Treatment assignment T_i	Individual treatment effect τ_i
1	90	?	1	?
2	?	90	0	?
3	?	85	0	?
4	100	?	1	?

What can we estimate?

- Estimating individual treatment effects is impossible
  - “Acts demolish their alternatives, that is the paradox”
  - Once a unit is assigned to a treatment condition, only potential outcomes tied to treatment assignment are realized.
  - Formally:

$$Y_i = T_i Y_i(1) + (1 - T_i) Y_i(0)$$

Neyman-Rubin Causal Model

	$Y_i(1)$	$Y_i(0)$	T_i	Y_i
1	90	85	1	1*90 + 0*85
2	80	90	0	0*80 + 1*90
3	85	85	0	0*85 + 1*85
4	100	0	1	1*100 + 0*0

Neyman-Rubin Causal Model

	$Y_i(1)$	$Y_i(0)$	T_i	Y_i
1	90	85	1	90
2	80	90	0	90
3	85	85	0	85
4	100	0	1	100

Neyman-Rubin Causal Model

	T_i	Y_i
1	1	90
2	0	90
3	0	85
4	1	100

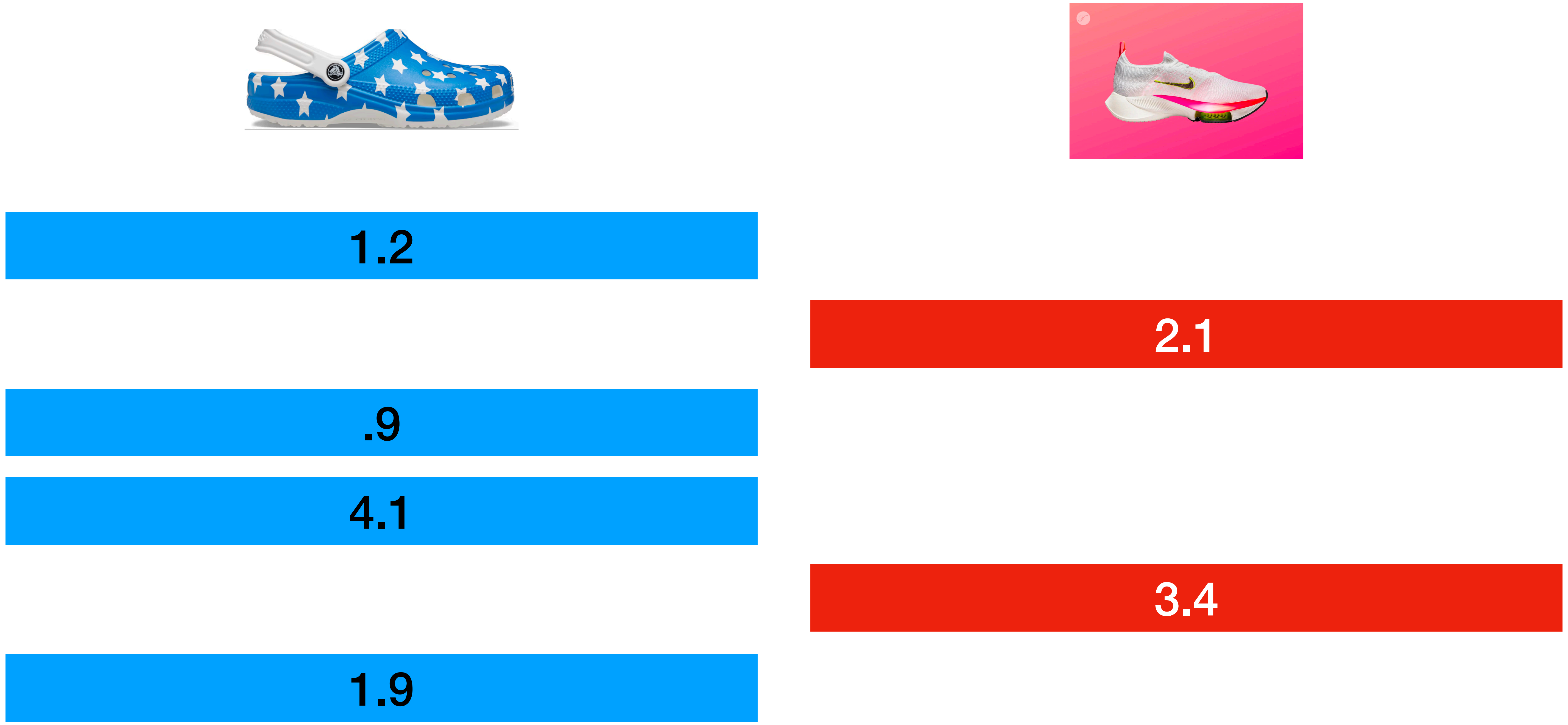
What can we estimate?

- ~~Individual treatment effects~~
- Average treatment effects
  - Though we cannot estimate how each individual will respond to an intervention, it is possible to calculate the average treatment effect
  - The average treatment effect is made up of two components
    - Mean treated potential outcome
    - Mean untreated (control) potential outcome
  - ATE = $\mathbb{E}[Y_i(1) - Y_i(0)]$ = $\mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)]$
  - The average difference in potential outcomes (i.e., average of individual treatment effects) is equal to difference in averages between treatment and control units.

What can we estimate?

1.2	1.5
2.5	2.1
.9	1.1
4.1	4.9
3.7	3.4
1.9	2.3

What can we estimate?



Experiments can be thought of as providing random samples of different possible worlds.

What can we estimate?

- ATE is a causal effect, but depends on two assumptions
  - Stable unit treatment value assumption
    - No spillovers or hidden treatments
      - The treatment status of one unit has no effect on another unit
      - “15 minutes of exercise”
  - Independence assumption
    - Treatment status is independent of potential outcomes
    - Whether a unit receives or does not receive the treatment is independent of their outcomes had they received treatment (or not)
    - Average potential outcomes are equal to average observed outcomes, given treatment assignment

Average treatment effect

The average treatment effect:

$$\mathbb{E}[Y_i(1) - Y_i(0)]$$

Average difference in potential outcomes for those in the treatment condition minus those in the control condition.

Average treatment effect

The average treatment effect:

$$\mathbb{E}[Y_i(1) - Y_i(0)]$$

Average difference in potential outcomes for those in the treatment condition minus those in the control condition.

This is a causal effect if there is zero selection bias (treatment assignment is independent of potential outcomes)

- Suppose we're interested in whether SAT prep courses improve SAT scores
  - $T_i = 0,1$
  - $Y_i = [0,1600]$
  - Can I just compare those who attended SAT prep courses to those who did not?
    - Average score (SAT Prep = 1): 1000
    - Average score (SAT Prep = 0): 1250
  - Why not?

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	900	700	1	200
2	1200	1200	0	0
3	1100	900	1	200
4	1300	1300	0	0

The problem of self-selection

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	900	**700**	1	200
2	1200	**1200**	0	0
3	1100	**900**	1	200
4	1300	**1300**	0	0

The problem of self-selection

Those who are likely to perform worse seek out SAT assistance whereas those who do not need it opt out.

Non-experimental designs

$$\mathbb{E}[Y_i(1) - Y_i(0) \mid T_i = 1] + (\mathbb{E}[Y_i(0) \mid T_i = 1] - \mathbb{E}[Y_i(0) \mid T_i = 0])$$

Potential outcome Under treatment

Potential outcome Under control

When assigned Treatment

Potential outcome Under control

When assigned Treatment

Potential outcome Under control

When assigned Control

Average treatment effect on the treated

Selection bias

Non-experimental designs

Population Avg. | given that… | Population Avg. | given that… | Population Avg. | given that…

$$\mathbb{E}[Y_i(1) - Y_i(0) \mid T_i = 1] + (\mathbb{E}[Y_i(0) \mid T_i = 1] - \mathbb{E}[Y_i(0) \mid T_i = 0])$$

Potential outcome Under treatment | Potential outcome Under control | When assigned Treatment | Potential outcome Under control | When assigned Treatment | Potential outcome Under control | When assigned Control

Average treatment effect on the treated

Selection bias

Non-experimental designs

$$Avg[Y_i(1) - Y_i(0) \mid T_i = 1] + (Avg[Y_i(0) \mid T_i = 1] - Avg[Y_i(0) \mid T_i = 0])$$

Potential outcome Under treatment	Potential outcome Under control	When assigned Treatment	Potential outcome Under control	When assigned Treatment	Potential outcome Under control	When assigned Control

Average treatment effect on the treated

Selection bias

Non-experimental designs

Average difference in potential outcomes for units assigned to treatment (T = 1)	Average control potential outcome for units assigned to treatment (T = 1)	Average control potential outcome for units assigned to control (T = 0)

$$Avg[Y_i(1) - Y_i(0) \mid T_i = 1] + (Avg[Y_i(0) \mid T_i = 1] - Avg[Y_i(0) \mid T_i = 0])$$

Potential outcome Under treatment

Potential outcome Under control

When assigned Treatment

Potential outcome Under control

When assigned Treatment

Potential outcome Under control

When assigned Control

Average treatment effect on the treated

Selection bias

Non-experimental designs

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	**900**	**700**	1	**200**
2	1200	1200	0	0
3	**1100**	**900**	1	**200**
4	1300	1300	0	0

$$\mathbb{E}[Y_i(1) - Y_i(0) \mid T_i = 1] + (\mathbb{E}[Y_i(0) \mid T_i = 1] - \mathbb{E}[Y_i(0) \mid T_i = 0])$$

(200 + 200)/2 = 200

Non-experimental designs

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	**900**	**700**	1	**200**
2	1200	1200	0	0
3	**1100**	**900**	1	**200**
4	1300	1300	0	0

$$\mathbb{E}[Y_i(1) - Y_i(0) \mid T_i = 1] + (\mathbb{E}[Y_i(0) \mid T_i = 1] - \mathbb{E}[Y_i(0) \mid T_i = 0])$$

200 (700 + 900)/2 = 800 (1200+1300)/2 = 1250

Non-experimental designs

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	**900**	**700**	1	**200**
2	1200	**1200**	0	0
3	**1100**	**900**	1	**200**
4	1300	**1300**	0	0

$$\mathbb{E}[Y_i(1) - Y_i(0) \mid T_i = 1] + (\mathbb{E}[Y_i(0) \mid T_i = 1] - \mathbb{E}[Y_i(0) \mid T_i = 0])$$

200 **800** - **1250**

Non-experimental designs

	$Y_i(1)$	$Y_i(0)$	T_i	τ_i
1	**900**	**700**	1	**200**
2	1200	1200	0	0
3	**1100**	**900**	1	**200**
4	1300	1300	0	0

$$\mathbb{E}[Y_i(1) - Y_i(0) \mid T_i = 1] + (\mathbb{E}[Y_i(0) \mid T_i = 1] - \mathbb{E}[Y_i(0) \mid T_i = 0])$$

200 **-450**

- Suppose we're interested in whether SAT prep courses improve SAT scores
  - $T_i = 0,1$
  - $Y_i = [0,1600]$
  - Can I just compare those who attended SAT prep courses to those who did not?
    - Average score (SAT Prep = 1): 1000
    - Average score (SAT Prep = 0): 1250
  - ATE = -250, but this is because there is selection bias.
  - ATT = 200, SB = -450

What can we estimate?

- In an actual experiment:
  - Treatment variable
  - Outcome variable
  - But, only observed outcomes are available, not potential outcomes.
- How is variation in the treatment variable determined?
  - Random assignment
  - What does this randomization get us?

Randomization

- What does it take to eliminate selection bias?
  - Ensure that treatment status and potential outcomes are statistically independent
    - What is statistical independence?
      - Having a piece of information tells us nothing about a probability or expected value
      - $\mathbb{E}(Y) = \mathbb{E}(Y \mid X)$
        - Y = Expected number of Congressional seats for a particular party
        - X = Number of meteor showers on Mercury
  - Randomization ensures statistical independence
    - Example: coin flip (fair coin)
    - Whether you observe the treated or control potential outcome for an individual is entirely due to chance

ATE = ATT + Selection Bias

$$Avg[Y_i(1) - Y_i(0) \mid T_i = 1] + (Avg[Y_i(0) \mid T_i = 1] - Avg[Y_i(0) \mid T_i = 0])$$

Potential outcome Under treatment	Potential outcome Under control	When assigned Treatment	Potential outcome Under control	When assigned Treatment	Potential outcome Under control	When assigned Control

Average treatment effect on the treated

Selection bias

Implications of statistical independence

$$(Avg[Y_i(1)] - Avg[Y_i(0)]) + (Avg[Y_i(0)] - Avg[Y_i(0)])$$

Experiments recover the ATE

$$(Avg[Y_i(1)] - Avg[Y_i(0)]) + 0$$