Stratification and Standardization

INFO/STSCI/ILRST 3900: Causal Inference

10 Sep 2023

Logistics

- Problem Set 1 due today by 5pm
- Peer reviews assigned via Canvas Thursday (9/12) @ 12pm and due Tuesday (9/17) @ 5pm
- Post questions on Ed Discussion or come to office hours!
  - Filippo: **Mon** 11am-12pm (Comstock 1187)
  - Sam: **Tue** 4-5pm (Comstock 1187)
  - Shira: **Wed** ~~3-4pm~~ **5:30-6:30pm** (Comstock 1187)
  - Mayleen: **Thu** ~~10:15-11:15am~~ **11am-12pm** (Rhodes 657)
- After class, read 2.4 of Hernán & Robins

Learning goals for today

At the end of class, you will be able to:

1. Explain marginal versus conditional exchangeability
2. Estimate the conditional average treatment effect using stratification
3. Estimate the average treatment effect using standardization

Check Your Understanding: Exchangeability

Discuss in groups, then submit your response individually to PollEverywhere. Your response won't be graded.

Registered voters are randomly assigned to two groups. The treatment group receives a phone call reminder to vote. The control group does not. After the election, we look at voter turnout in both groups. **What does exchangeability mean in this study?**

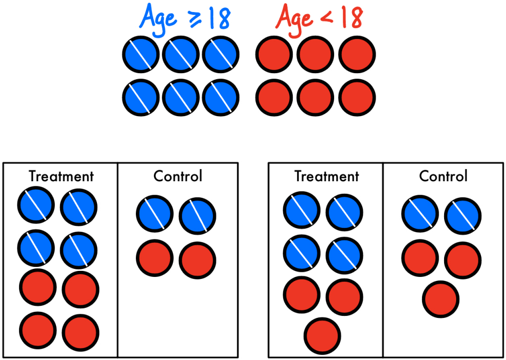
https://pollev.com/causal3900

Review: conditional randomization

- Causal effect of job training on employment rate
- Setup 1: flip same coin for everyone (left)
  - Job training with probability 2/3
- Setup 2: flip different coins depending on age (right)
  - Age ≥ 18: Job training with probability 2/3
  - Age < 18: Job training with probability 1/2

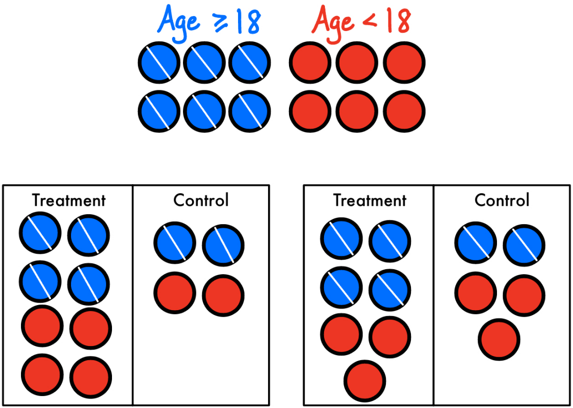
Observe: distribution of age groups is the same in treatment versus control in setup 1, but not setup 2

Key assumption:
$Y^a \not\perp\!\!\!\perp$ AGE (L)



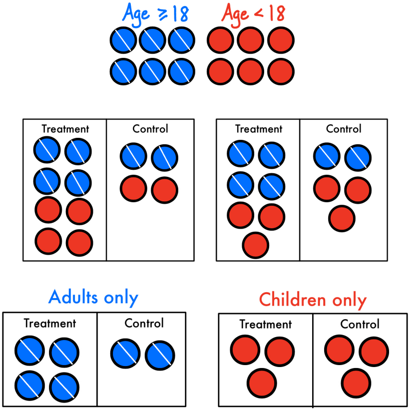
Review: conditional randomization

- **Marginal exchangeability:** $Y^a \perp\!\!\!\perp A$ for all a
- Suppose we find there is a treatment effect using the second setup. What is suspicious about that result?
- In practice, one way to reason about exchangeability is balance across covariates



Review: conditional randomization

- **Conditional exchangeability:** $Y^a \perp\!\!\!\perp A \mid L$ for all a
- When we look within sub-populations, or *strata*, marginal exchangeability holds



Check-in: Marginal vs Conditional Exchangeability

- **Marginal exchangeability:** $Y^a \perp\!\!\!\perp A$ for all a
- **Conditional exchangeability:** $Y^a \perp\!\!\!\perp A \mid L$ for all a

  The potential outcomes are independent of treatment **conditional on** L (e.g. L could be an indicator of age group)

- How do you feel about marginal versus conditional exchangeability (thumbs up or thumbs down)?
- What questions do you have?

Stratification

- Exchangeability holds *within* each sub-population
- **Stratification**: We can directly estimate causal effect within each sub-population (or stratum)
- The Conditional Average Treatment Effect (CATE)

$$CATE = E(Y^{a=1} \mid L = \ell) - E(Y^{a=0} \mid L = \ell)$$

- Conditional exchangeability (+ consistency) lets us estimate expected potential outcomes from observable averages

$$\begin{aligned} \mathrm{E}(Y \mid A = a, L = \ell) &\overset{\text{consis}}{=} \mathrm{E}(Y^a \mid A = a, L = \ell) \\ &\overset{\text{exchange}}{=} \mathrm{E}(Y^a \mid L = \ell) \end{aligned}$$

- If the treatment effect varies across sub-population, we say there is **treatment effect heterogeneity**

Stratification

- Estimate the Conditional Average Treatment Effect (CATE)
- $CATE = E(Y^{a=1} \mid L = \ell) - E(Y^{a=0} \mid L = \ell)$
- Example: medication

- Sometimes we do need the Average Treatment Effect (ATE)
- $ATE = E(Y^{a=1}) - E(Y^{a=0})$
- Example: new feature on YouTube (two thumbs up)
- How can we go from CATE to ATE?

- Law of Total Expectation:

$$E[Y^a] = \sum_{\ell} E[Y^a \mid L = \ell] \cdot Pr(L = \ell)$$

Standardization

- **Standardization** allows us to estimate the ATE by combining estimates from each sub-population
- For $L = 0, 1$

$$E(Y^a) = E(Y^a \mid L = 1) \cdot Pr(L = 1) \; + \; E(Y^a \mid L = 0) \cdot Pr(L = 0)$$

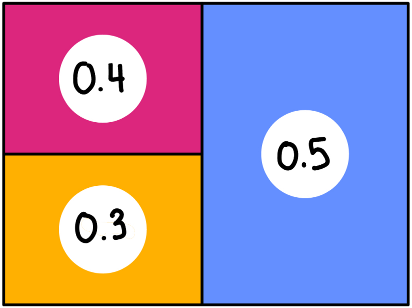
$$\begin{aligned} &= E(Y \mid L = 1, A = a) \cdot Pr(L = 1) \\ &\qquad + \; E(Y \mid L = 0, A = a) \cdot Pr(L = 0) \end{aligned}$$

- More generally, for each a

$$E(Y^a) = \sum_{\ell} E(Y \mid L = \ell, A = a) \cdot Pr(L = \ell)$$

Standardization

$$E(Y^a) = \sum_{\ell} E(Y \mid L = \ell, A = a) Pr(L = \ell)$$



$$E(Y^a) = 0.4 \cdot \frac{1}{4} + 0.3 \cdot \frac{1}{4} + 0.5 \cdot \frac{1}{2}$$

Check Your Understanding: Standardization

$$E(Y^{a=0}) = \sum_\ell E(Y \mid L = \ell, A = 0) Pr(L = \ell)$$

$$E(Y^{a=1}) = \sum_\ell E(Y \mid L = \ell, A = 1) Pr(L = \ell)$$

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

Check Your Understanding: Standardization

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$Pr(L=1) \cdot \mathrm{E}(Y \mid L=1, A=0) =$$
$$Pr(L=0) \cdot \mathrm{E}(Y \mid L=0, A=0) =$$

$$Pr(L=1) \cdot \mathrm{E}(Y \mid L=1, A=1) =$$
$$Pr(L=0) \cdot \mathrm{E}(Y \mid L=0, A=1) =$$

Check Your Understanding: Standardization

		L	A	Y
1	Rheia	0	0	0
2	Kronos	0	0	1
3	Demeter	0	0	0
4	Hades	0	0	0
5	Hestia	0	1	0
6	Poseidon	0	1	0
7	Hera	0	1	0
8	Zeus	0	1	1

		L	A	Y
9	Artemis	1	0	1
10	Apollo	1	0	1
11	Leto	1	0	0
12	Ares	1	1	1
13	Athena	1	1	1
14	Hephaestus	1	1	1
15	Aphrodite	1	1	1
16	Polyphemus	1	1	1
17	Persephone	1	1	1
18	Hermes	1	1	0
19	Hebe	1	1	0
20	Dionysus	1	1	0

$$Pr(L=1) \cdot \mathrm{E}(Y \mid L=1, A=0) = \tfrac{12}{20} \cdot \tfrac{2}{3}$$
$$Pr(L=0) \cdot \mathrm{E}(Y \mid L=0, A=0) = \tfrac{8}{20} \cdot \tfrac{1}{4}$$
$$Pr(L=1) \cdot \mathrm{E}(Y \mid L=1, A=1) = \tfrac{12}{20} \cdot \tfrac{6}{9}$$
$$Pr(L=0) \cdot \mathrm{E}(Y \mid L=0, A=1) = \tfrac{8}{20} \cdot \tfrac{1}{4}$$

Get $E(Y^{a=0})$ and $E(Y^{a=1})$ using formulas in slide 13!

Learning goals for today

At the end of class, you will be able to:

1. Explain marginal versus conditional exchangeability
2. Estimate the conditional average treatment effect using stratification
3. Estimate the average treatment effect using standardization