Inverse Probability Weighting

INFO/STSCI/ILRST 3900: Causal Inference

7 Sep 2023

Learning goals for today

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

1. Estimate the average causal effect using data from a conditionally randomized experiment using inverse probability weighting
2. Explain why conditional exchangeability might be reasonable in some observational data

Logistics

- Ch 2.4 and 3.2 in Hernan and Robins 2023
- Problem set 2 posted today, due on Sep 14

Conditional randomization

- **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
- **Stratification**: We can directly estimate causal effect within each sub-population (or stratum)
- We can estimate the ACE using standardization

Excercise

		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

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

$$+ Pr(L=0)E(Y \mid L=0, A=1)$$

Excercise

		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

$$\mathrm{E}(Y^{a=1}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=1)}_{6/9}$$

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

Excercise

		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

$$\mathrm{E}(Y^{a=1}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=1)}_{6/9}$$

$$+ \underbrace{Pr(L=0)}_{8/20}\underbrace{\mathrm{E}(Y \mid L=0, A=1)}_{1/4} = 1/2$$

Excercise

		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

$$E(Y^{a=0}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{E(Y \mid L=1, A=0)}_{2/3}$$

$$+ \underbrace{Pr(L=0)}_{8/20}\underbrace{E(Y \mid L=0, A=0)}_{1/4}$$

Excercise

		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

$$\mathrm{E}(Y^{a=0}) = \underbrace{Pr(L=1)}_{12/20}\underbrace{\mathrm{E}(Y \mid L=1, A=0)}_{2/3}$$

$$+ \underbrace{Pr(L=0)}_{8/20}\underbrace{\mathrm{E}(Y \mid L=0, A=0)}_{1/4} = 1/2$$

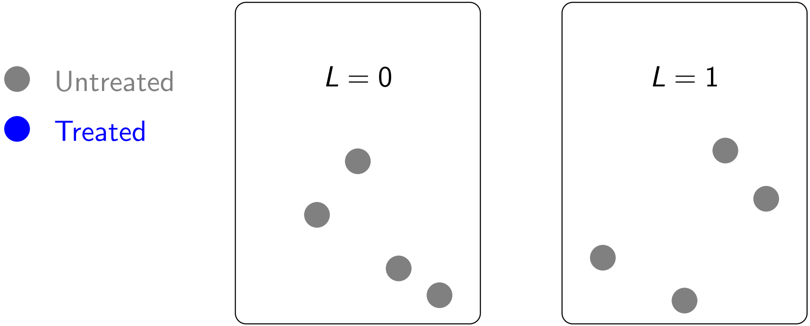
Inverse probability weighting

- Standardization: constructs an estimate of $\mathrm{E}(Y^a)$ through a weighted average
- Inverse probability weighted (IPW) estimator is equivalent to standardization
- Estimator for the ATE

$$\mathrm{E}(Y^a) = \frac{1}{N} \sum_{i:A_i=a} \frac{Y_i}{\pi_i}$$

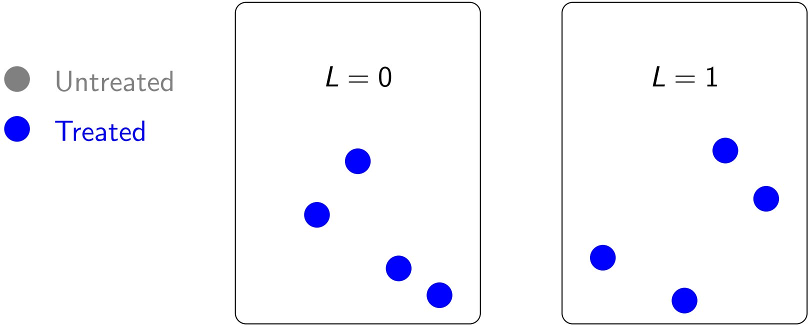
- $\pi_i = \mathrm{P}(A = a_i \mid L = \ell_i)$ is the probability of the observed treatment conditioning on confounders
- N is the total number of observations (over all treatment groups and confounder groups

Inverse probability weighting: Conditional randomizaton



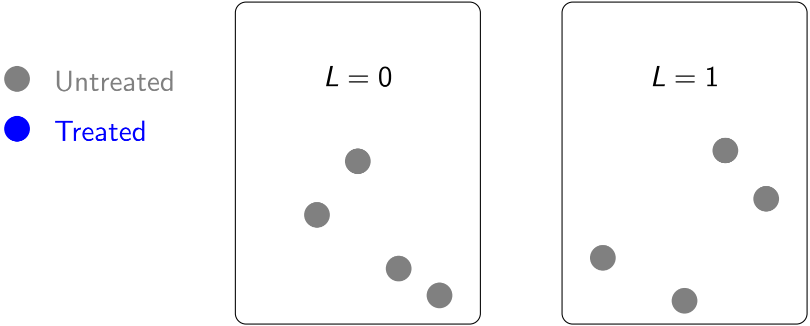
Hypothetical world where no-one is treated

Inverse probability weighting: Conditional randomizaton

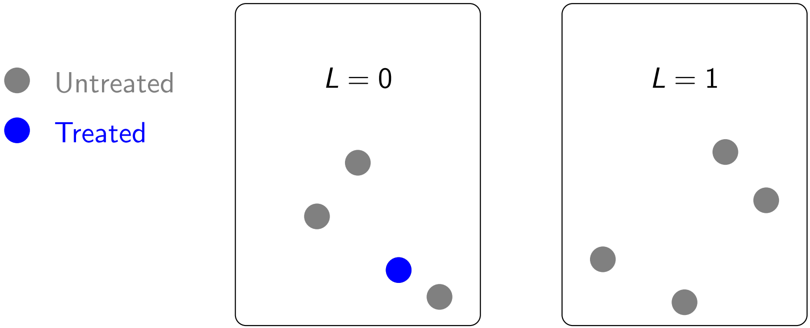


Hypothetical world where everyone is treated

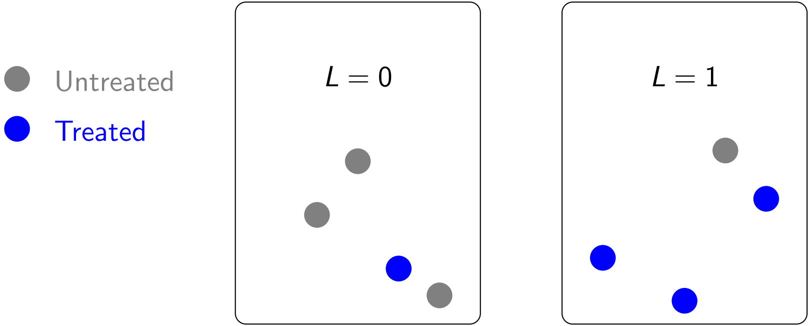
Inverse probability weighting: Conditional randomizaton



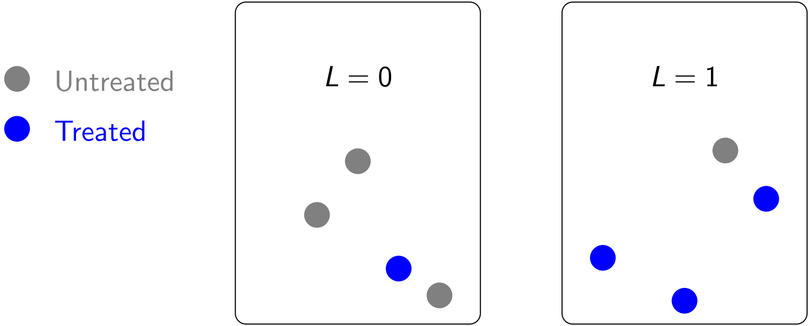
Inverse probability weighting: Conditional randomizaton



Inverse probability weighting: Conditional randomizaton

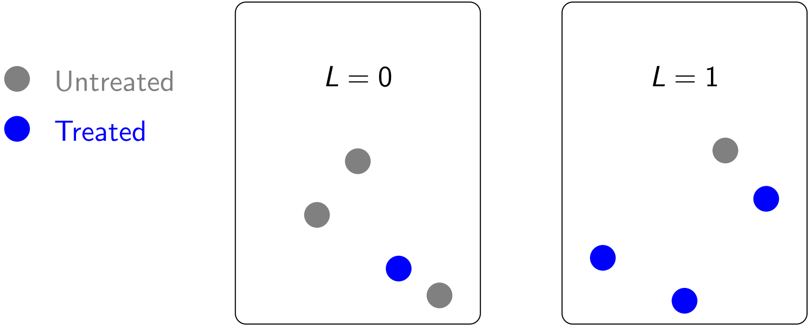


Inverse probability weighting: Conditional randomizaton



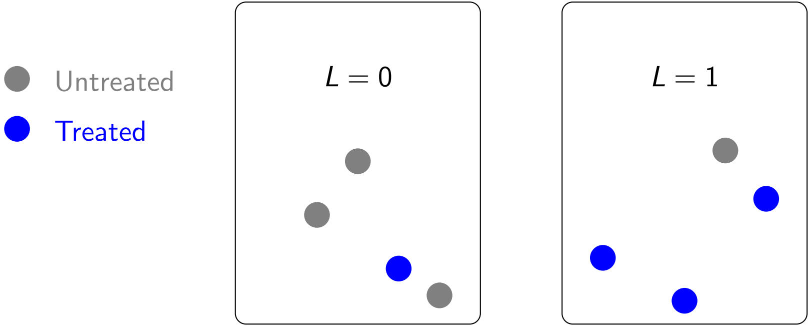
$$\pi_i = \mathrm{P}(A = a_i \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } A_i = 1 \\ \frac{3}{4} & \text{if } A_i = 0 \end{cases}$$

Inverse probability weighting: Conditional randomizaton



$$\pi_i = \mathrm{P}(A = a_i \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } A_i = 1 \\ \frac{3}{4} & \text{if } A_i = 0 \end{cases} \qquad \begin{cases} \frac{3}{4} & \text{if } A_i = 1 \\ \frac{1}{4} & \text{if } A_i = 0 \end{cases}$$

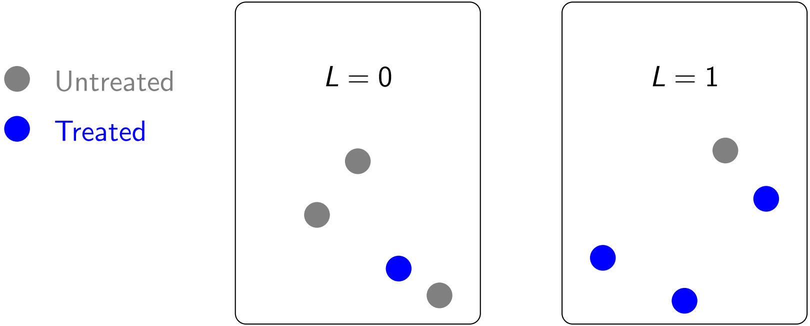
Inverse probability weighting: Conditional randomizaton



$$\pi_i = \mathrm{P}(A = a_i \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } A_i = 1 \\ \frac{3}{4} & \text{if } A_i = 0 \end{cases} \qquad \begin{cases} \frac{3}{4} & \text{if } A_i = 1 \\ \frac{1}{4} & \text{if } A_i = 0 \end{cases}$$

Each counts for:

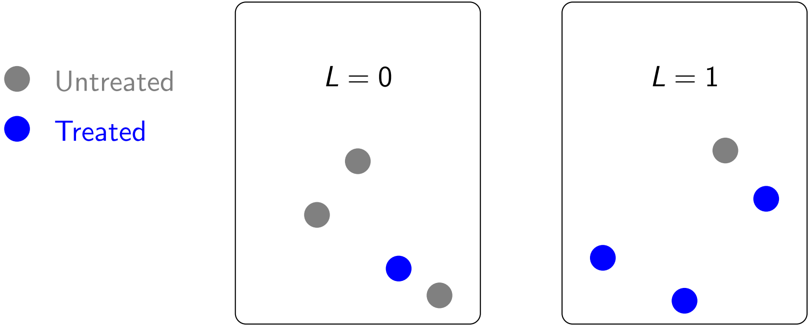
Inverse probability weighting: Conditional randomizaton



$$\pi_i = \mathrm{P}(A = a_i \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } A_i = 1 \\ \frac{3}{4} & \text{if } A_i = 0 \end{cases} \qquad \begin{cases} \frac{3}{4} & \text{if } A_i = 1 \\ \frac{1}{4} & \text{if } A_i = 0 \end{cases}$$

Each counts for: $\begin{cases} \frac{4}{1} & \text{if } A_i = 1 \\ \frac{4}{3} & \text{if } A_i = 0 \end{cases}$

Inverse probability weighting: Conditional randomizaton



$$\pi_i = \mathrm{P}(A = a_i \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } A_i = 1 \\ \frac{3}{4} & \text{if } A_i = 0 \end{cases} \qquad \begin{cases} \frac{3}{4} & \text{if } A_i = 1 \\ \frac{1}{4} & \text{if } A_i = 0 \end{cases}$$

Each counts for:

$$\begin{cases} \frac{4}{1} & \text{if } A_i = 1 \\ \frac{4}{3} & \text{if } A_i = 0 \end{cases} \qquad \begin{cases} \frac{4}{3} & \text{if } A_i = 1 \\ \frac{4}{1} & \text{if } A_i = 0 \end{cases}$$

Inverse probability weighting: Mathematical proof[1]

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

$$\text{E}\left(\frac{\mathbb{I}(A=a)}{\text{P}(A=a\mid \vec{L})}Y\right) \tag{1}$$

$$= \text{E}(Y^a) \tag{6}$$

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

$$\mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a \mid \vec{L})}Y\right) \tag{1}$$

$$= \mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a \mid \vec{L})}Y^a\right) \quad \text{consistency} \tag{2}$$

$$= \mathrm{E}(Y^a) \tag{6}$$

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

$$\text{E}\left(\frac{\mathbb{I}(A=a)}{\text{P}(A=a \mid \vec{L})}Y\right) \tag{1}$$

$$= \text{E}\left(\frac{\mathbb{I}(A=a)}{\text{P}(A=a \mid \vec{L})}Y^a\right) \quad \text{consistency} \tag{2}$$

$$= \text{E}\left(\text{E}\left[\frac{\mathbb{I}(A=a)}{\text{P}(A=a \mid \vec{L})}Y^a \mid \vec{L}\right]\right) \quad \text{iterated expectation} \tag{3}$$

$$= \text{E}(Y^a) \tag{6}$$

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

$$\mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a \mid \vec{L})} Y\right) \tag{1}$$

$$= \mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a \mid \vec{L})} Y^a\right) \qquad \text{consistency} \tag{2}$$

$$= \mathrm{E}\left(\mathrm{E}\left[\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a \mid \vec{L})} Y^a \mid \vec{L}\right]\right) \qquad \text{iterated expectation} \tag{3}$$

$$= \mathrm{E}\left(\mathrm{E}\left[\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a \mid \vec{L})} \mid \vec{L}\right] \mathrm{E}\left[Y^a \mid \vec{L}\right]\right) \qquad \text{exchangeability} \tag{4}$$

$$= \mathrm{E}(Y^a) \tag{6}$$

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

$$\mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}Y\right) \tag{1}$$

$$=\mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}Y^a\right) \quad \text{consistency} \tag{2}$$

$$=\mathrm{E}\left(\mathrm{E}\left[\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}Y^a\mid\vec{L}\right]\right) \quad \text{iterated expectation} \tag{3}$$

$$=\mathrm{E}\left(\mathrm{E}\left[\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}\mid\vec{L}\right]\mathrm{E}\left[Y^a\mid\vec{L}\right]\right) \quad \text{exchangeability} \tag{4}$$

$$=\mathrm{E}\left(\mathrm{E}\left[Y^a\mid\vec{L}\right]\right) \quad \text{since left term was 1} \tag{5}$$

$$=\mathrm{E}(Y^a) \tag{6}$$

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

$$
\begin{aligned}
&\mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}Y\right) && && (1)\\
&=\mathrm{E}\left(\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}Y^a\right) && \text{consistency} && (2)\\
&=\mathrm{E}\left(\mathrm{E}\left[\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}Y^a\mid\vec{L}\right]\right) && \text{iterated expectation} && (3)\\
&=\mathrm{E}\left(\mathrm{E}\left[\frac{\mathbb{I}(A=a)}{\mathrm{P}(A=a\mid\vec{L})}\mid\vec{L}\right]\mathrm{E}\left[Y^a\mid\vec{L}\right]\right) && \text{exchangeability} && (4)\\
&=\mathrm{E}\left(\mathrm{E}\left[Y^a\mid\vec{L}\right]\right) && \text{since left term was 1} && (5)\\
&=\mathrm{E}(Y^a) && \text{iterated expectation} && (6)
\end{aligned}
$$

[1] Hernán & Robins Technical Point 2.3

Excercise

$$E(Y^a) = \frac{1}{N} \sum_{i:A_i=a} Y_i/\pi_i \qquad \pi_i = Pr(A_i = a \mid L = \ell_i)$$

Name	L	A	Y
Rheia	0	0	0
Kronos	0	0	1
Demeter	0	0	0
Hades	0	0	0
Hestia	0	1	0
Poseidon	0	1	0
Hera	0	1	0
Zeus	0	1	1

Name	L	A	Y
Artemis	1	0	1
Apollo	1	0	1
Leto	1	0	0
Ares	1	1	1
Athena	1	1	1
Hephaestus	1	1	1
Aphrodite	1	1	1
Polyphemus	1	1	1
Persephone	1	1	1
Hermes	1	1	0
Hebe	1	1	0
Dionysus	1	1	0

Excercise

$$\mathrm{E}(Y^a) = \frac{1}{N} \sum_{i:A_i=a} Y_i/\pi_i \qquad \pi_i = Pr(A = a_i \mid L = \ell_i)$$

Name	L	A	Y
Rheia	0	0	0
Kronos	0	0	1
Demeter	0	0	0
Hades	0	0	0
Hestia	0	1	0
Poseidon	0	1	0
Hera	0	1	0
Zeus	0	1	1

Name	L	A	Y
Artemis	1	0	1
Apollo	1	0	1
Leto	1	0	0
Ares	1	1	1
Athena	1	1	1
Hephaestus	1	1	1
Aphrodite	1	1	1
Polyphemus	1	1	1
Persephone	1	1	1
Hermes	1	1	0
Hebe	1	1	0
Dionysus	1	1	0

Conditional exchangeability in observational data

- When conditional exchangeability holds, we can estimate causal effects from the observed data
- Use either standardization or inverse probability weighting
- By design, conditional exchangeability holds in conditionally randomized experiments
- Marginal exchangeability is very unlikely in observational data
- Conditional exchangeability is may be more reasonable in observational data

Excercise

Suppose we have data gathered by surveying individuals in Fall of 2021

- Whether the individual was vaccinated for Covid
  $A_i = 1$ if vaccinated, $A_i = 0$ if not vaccinated
- Whether the individual tested positive for Covid in 2021
  $Y_i = 1$ if positive test, $Y_i = 0$ if no positive test
- What additional information could you gather about each individual to make conditional exchangeability might be plausible?

$$Y^{a=1}, Y^{a=0} \perp\!\!\!\perp A \mid L$$

Conditional exchangeability in observational data

- Even if gathering data was possible for every covariate we want, when do we stop?

Conditional exchangeability in observational data

- Even if gathering data was possible for every covariate we want, when do we stop?
- Never 100% sure that conditional exchangeability holds
- Is it reasonable?

Conditional exchangeability in observational data

- Even if gathering data was possible for every covariate we want, when do we stop?
- Never 100% sure that conditional exchangeability holds
- Is it reasonable?
- In observational data, conditional exchangeability is an assumption we make (but can't typically verify)
- Requires expert knowledge

Conditional exchangeability in observational data

- Even if gathering data was possible for every covariate we want, when do we stop?
- Never 100% sure that conditional exchangeability holds
- Is it reasonable?
- In observational data, conditional exchangeability is an assumption we make (but can't typically verify)
- Requires expert knowledge
- Causal claims are data + outside knowledge

Learning goals for today

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

1. Estimate the average causal effect using data from a conditionally randomized experiment using inverse probability weighting
2. Explain why conditional exchangeability might be reasonable in some observational data