Inverse Probability Weighting

INFO/STSCI/ILRST 3900: Causal Inference

11 Sep 2025

Learning goals for today

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

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

Logistics

- Ch 1.3 and 2.4 in Hernan and Robins 2023
- Problem Set 1 Peer Review due Sep 16
- Problem set 2 posted today, due on Sep 19
- Quiz 1 will be in class on Sep 18
    - 10 minutes
    - paper + pen/pencil
    - Please email me for SDS accomodations

Measures of association/causation[2]

- Average Causal Effect (Average treatment effect or Risk Difference)

$$E(Y^{a=1}) - E(Y^{a=0})$$

[1]For binary outcomes, $E(Y^{a=1} = 1) = Pr(Y^{a=1} = 1)$
[2]Ch 1.2 and 1.3 of Hernan and Robins

Measures of association/causation[2]

- Average Causal Effect (Average treatment effect or Risk Difference)

$$E(Y^{a=1}) - E(Y^{a=0})$$

- Causal Risk Ratio[1]:

$$\frac{E(Y^{a=1})}{E(Y^{a=0})}$$

- Causal Odds Ratio (binary outcome and treatment):

$$\frac{E(Y^{a=1}=1)/E(Y^{a=1}=0)}{E(Y^{a=0}=1)/E(Y^{a=0}=0)}$$

[1] For binary outcomes, $E(Y^{a=1}=1) = Pr(Y^{a=1}=1)$

[2] Ch 1.2 and 1.3 of Hernan and Robins

Measures of association/causation[2]

- Average Causal Effect (Average treatment effect or Risk Difference)

$$E(Y^{a=1}) - E(Y^{a=0})$$

- Causal Risk Ratio[1]:

$$\frac{E(Y^{a=1})}{E(Y^{a=0})}$$

- Causal Odds Ratio (binary outcome and treatment):

$$\frac{E(Y^{a=1} = 1)/E(Y^{a=1} = 0)}{E(Y^{a=0} = 1)/E(Y^{a=0} = 0)}$$

- No average causal effect if $CRR = COR = 1$

[1] For binary outcomes, $E(Y^{a=1} = 1) = Pr(Y^{a=1} = 1)$

[2] Ch 1.2 and 1.3 of Hernan and Robins

Measures of association/causation[3]

- All measures roughly agree if $\mathrm{E}(Y^{a=1}) = \mathrm{E}(Y^{a=0})$
- When $\mathrm{E}(Y^{a=1}) \neq \mathrm{E}(Y^{a=0})$, the different measures may be easier/harder to interpret
- From Pfizer Covid-19 Vaccine
  - Of the individuals who were given the vaccine ($A_i = 1$), 0.04% had a positive Covid test ($Y_i = 1$)
  - Of the individuals who were given the placebo ($A_i = 0$), 0.9% had a positive Covid test ($Y_i = 1$)
  - Under consistency and exchangeability, what is the ACE and CRR?
  - When trying to advocate for the vaccine, which measure would you use?
  - When trying to advocate against the vaccine, which measure would you use?

[3] Ch 1.2 and 1.3 of Hernan and Robins

Measures of association/causation[4]

- Way causal quantities are communicated can make a difference
- Statistical significance (or non-zero causal effect) does not necessarily imply clinical/practical relevance
- When using using causal inference to make decisions, must place in broader context

[4] Ch 1.2 and 1.3 of Hernan and Robins

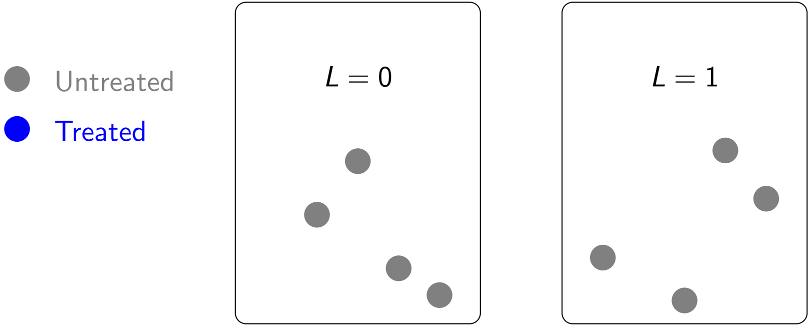
Inverse probability weighting

- 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

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

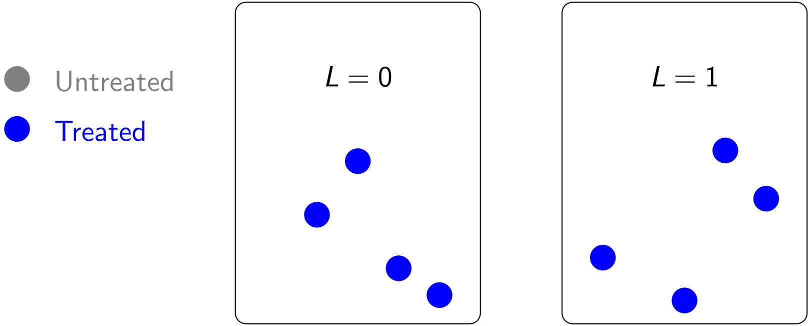
- $\pi_i = \mathrm{P}(A = a_i \mid L = \ell_i)$ is the probability of the observed treatment conditioning on confounders
- N is the total number of observations (over all treatment groups and confounder groups

Inverse probability weighting: Conditional randomizaton



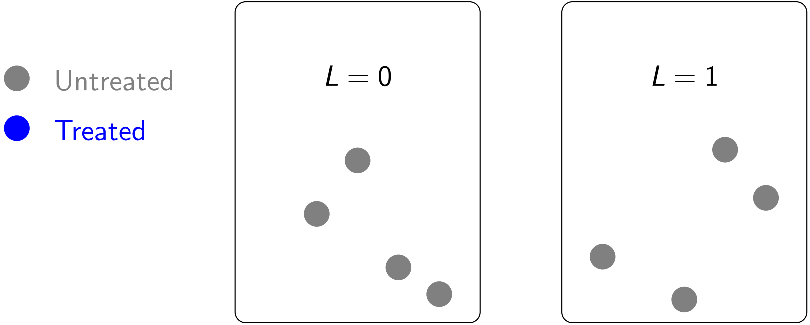
Hypothetical world where no-one is treated

Inverse probability weighting: Conditional randomizaton

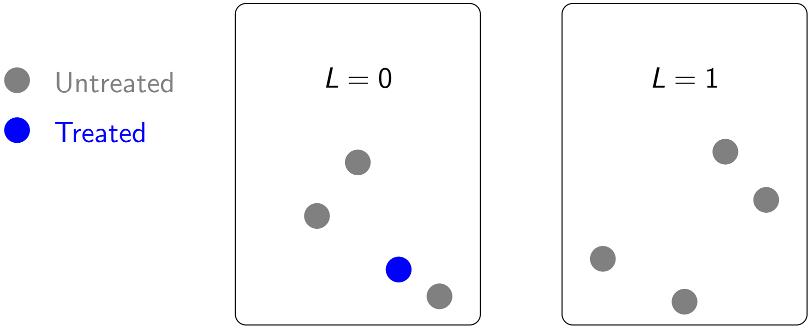


Hypothetical world where everyone is treated

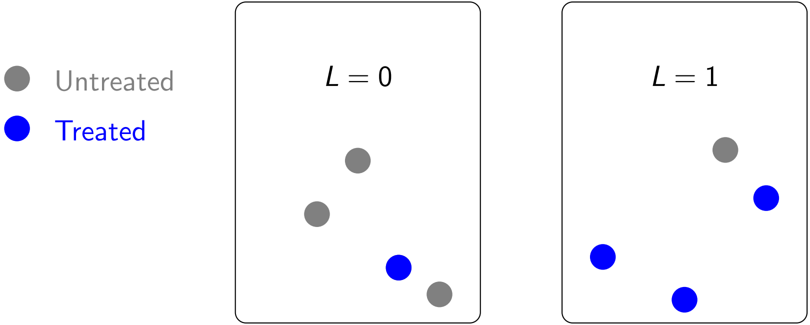
Inverse probability weighting: Conditional randomizaton



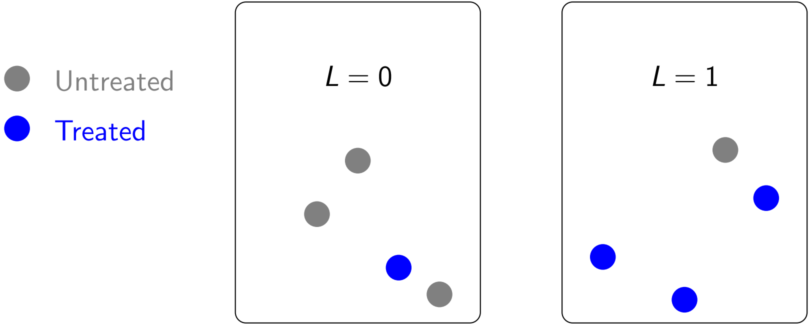
Inverse probability weighting: Conditional randomizaton



Inverse probability weighting: Conditional randomizaton

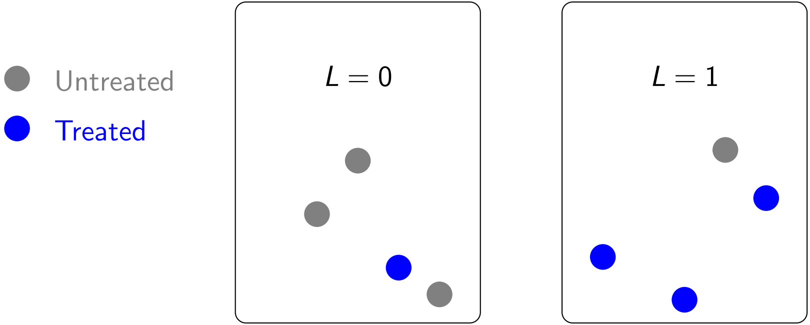


Inverse probability weighting: Conditional randomizaton



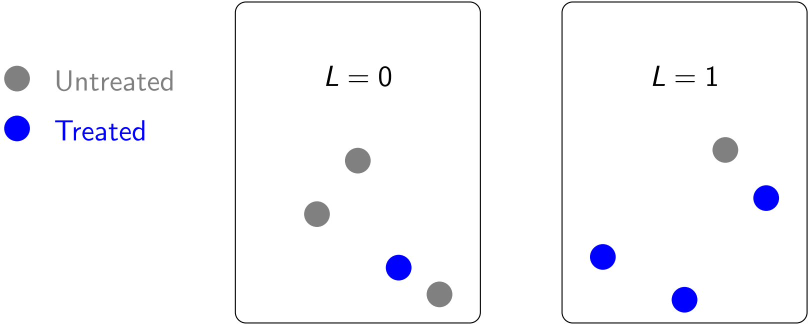
$$\pi_i = \mathrm{P}(A_i = a \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } a = 1 \\ \frac{3}{4} & \text{if } a = 0 \end{cases}$$

Inverse probability weighting: Conditional randomizaton



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

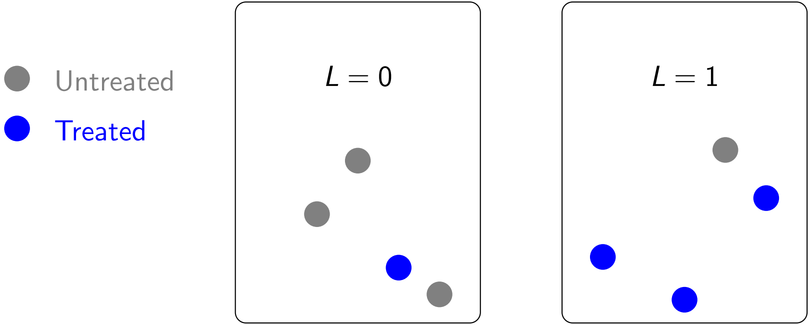
Inverse probability weighting: Conditional randomizaton



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

Each counts for:

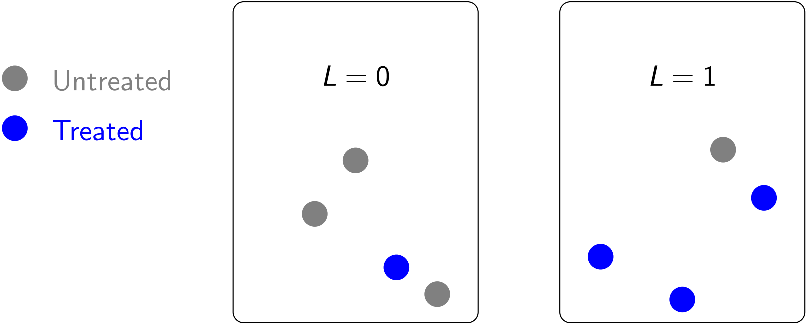
Inverse probability weighting: Conditional randomizaton



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

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

Inverse probability weighting: Conditional randomizaton



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

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

Inverse probability weighting: Mathematical proof[5]

[5] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[5]

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

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

[5]Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[5]

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

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

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

[5]Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[5]

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

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

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

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

[5] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[5]

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

[5]Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[5]

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

[5] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[5]

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

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

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

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

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

$$= \mathrm{E}(Y^a) \quad \text{iterated expectation} \tag{6}$$

[5] Hernán & Robins Technical Point 2.3

Excercise

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

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. Describe different ways to measure a causal effect
2. Estimate the average causal effect using data from a conditionally randomized experiment using inverse probability weighting
3. Explain why conditional exchangeability might be reasonable in some observational data