13. Inverse Probability Weighting

Ian Lundberg
Cornell Info 6751: Causal Inference in Observational Settings
Fall 2022

4 Oct 2022

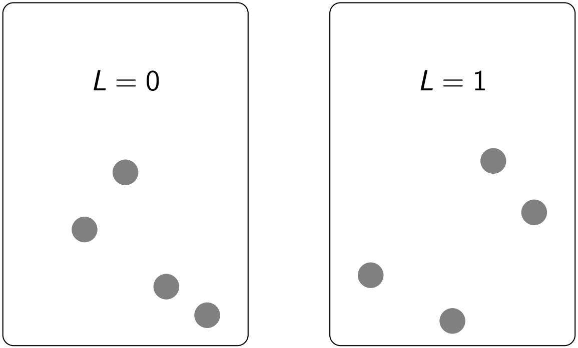
Learning goals for today

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

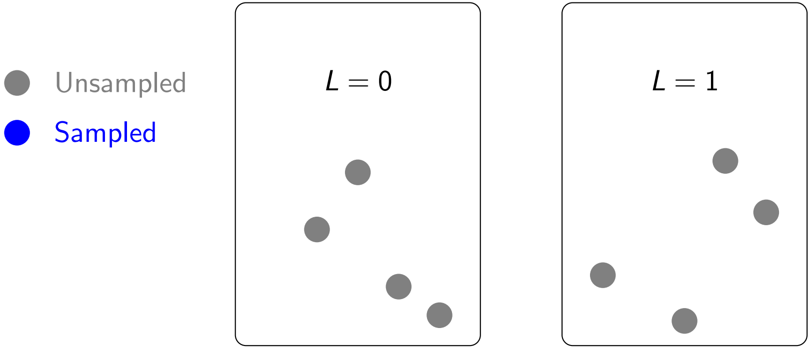
1. Trace inverse probability weighting to survey sampling
2. Apply the Horvitz-Thompson estimator for causal inference
3. Recognize the bias-variance tradeoff of trimmed weights
4. Estimate the ATE, ATT, and ATC by weighting

Inverse probability weighting: Sampling

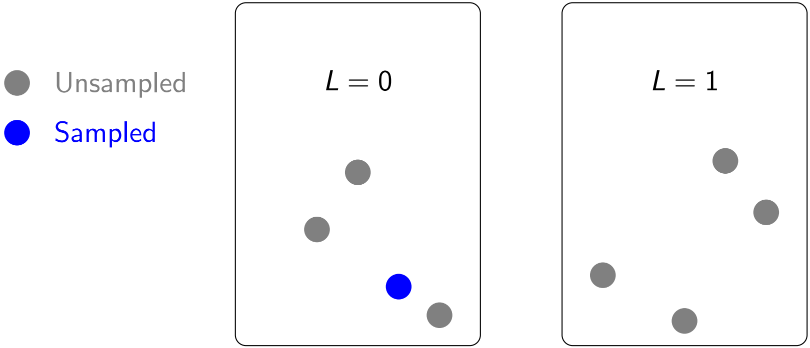
Inverse probability weighting: Sampling



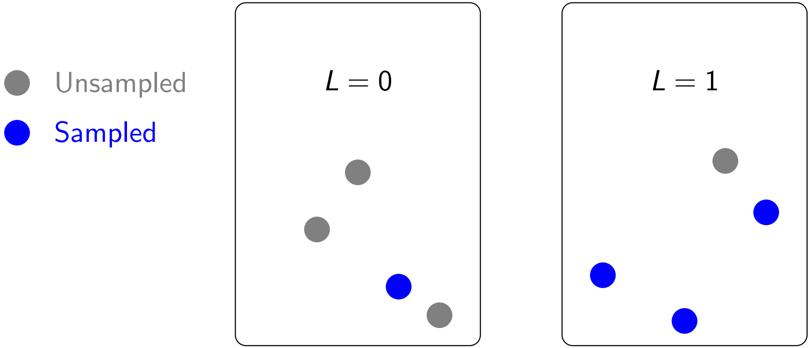
Inverse probability weighting: Sampling



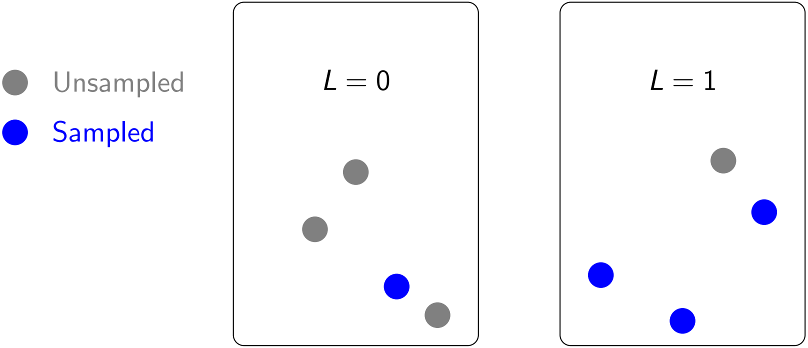
Inverse probability weighting: Sampling



Inverse probability weighting: Sampling



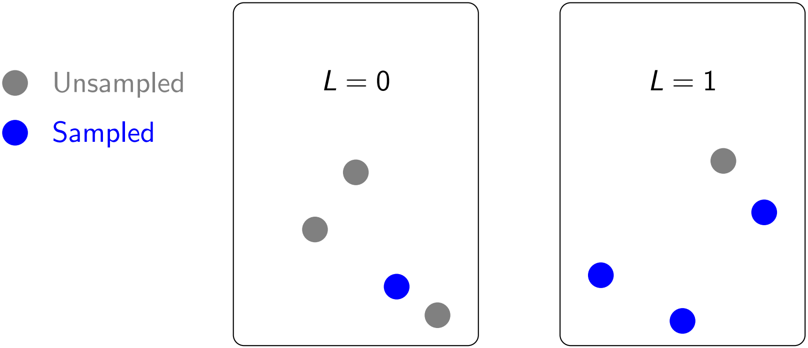
Inverse probability weighting: Sampling



$$\pi_i = \mathrm{P}(S = 1 \mid L_i = 0) = \frac{1}{4}$$

$$\pi_i = \mathrm{P}(S = 1 \mid L_i = 1) = \frac{3}{4}$$

Inverse probability weighting: Sampling



$$\pi_i = \mathrm{P}(S = 1 \mid L_i = 0) = \frac{1}{4}$$

$$\pi_i = \mathrm{P}(S = 1 \mid L_i = 1) = \frac{3}{4}$$

Each counts for: $w_i = \dfrac{1}{\pi_i} = 4$ $\qquad$ $w_i = \dfrac{1}{\pi_i} = \dfrac{4}{3}$

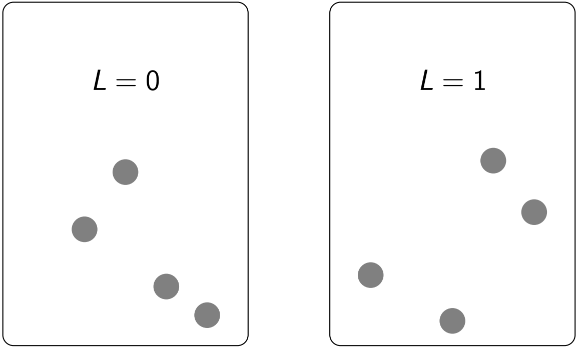
The Horvitz-Thompson estimator for a population mean:

$$\hat{\mathrm{E}}(Y) = \frac{1}{N} \sum_{i:S_i=1} \frac{Y_i}{\pi_i}$$

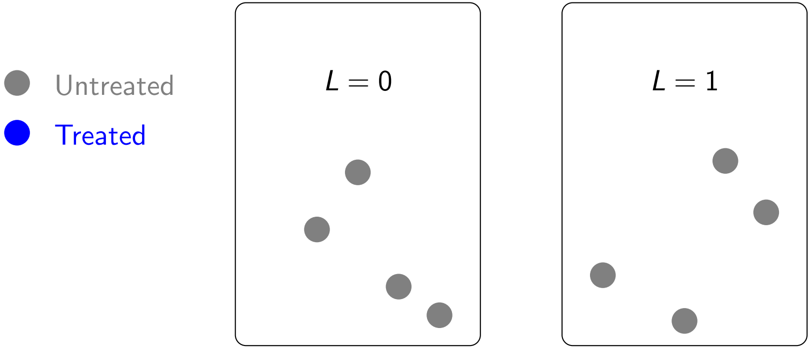
where N is the population size and
π_i is the known probability of sampling

Inverse probability weighting: Conditional randomizaton

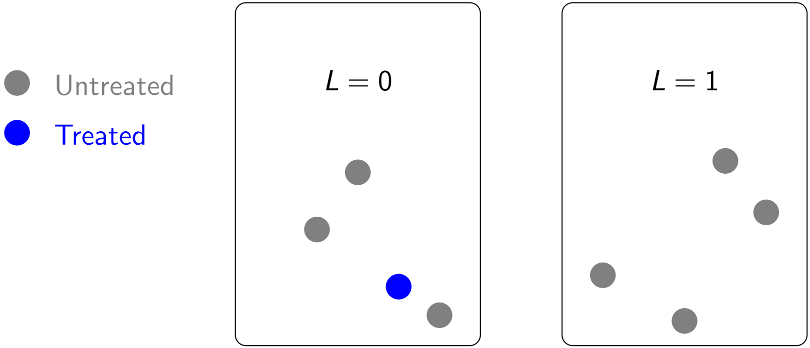
Inverse probability weighting: Conditional randomizaton



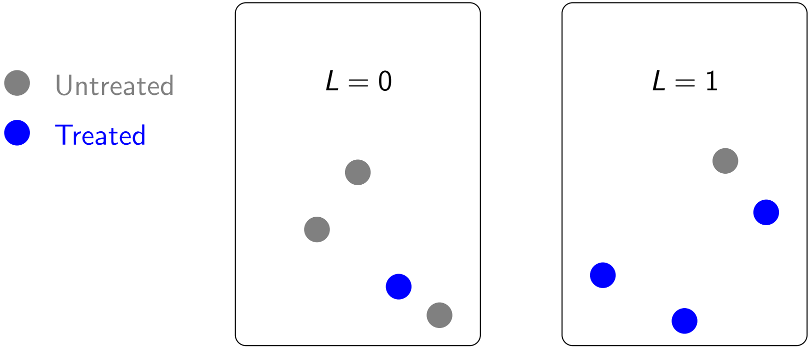
Inverse probability weighting: Conditional randomizaton



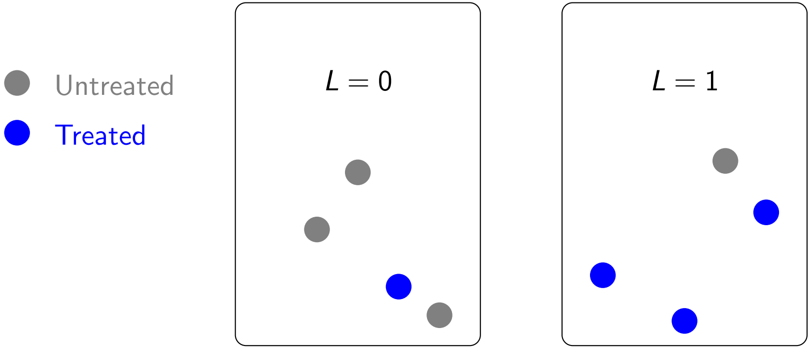
Inverse probability weighting: Conditional randomizaton



Inverse probability weighting: Conditional randomizaton

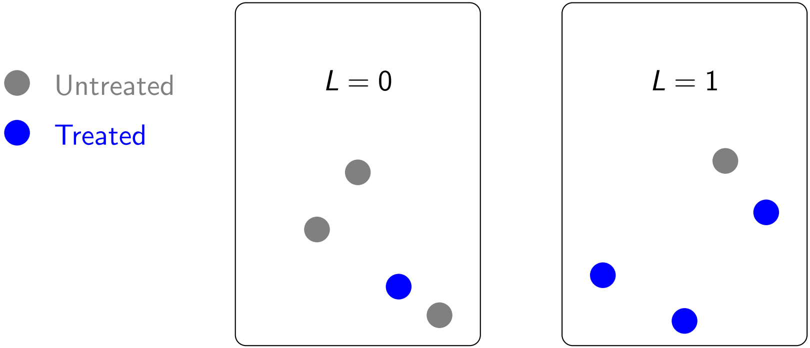


Inverse probability weighting: Conditional randomizaton



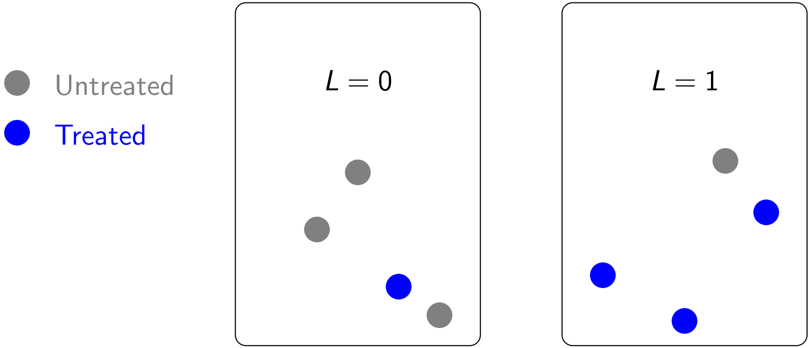
$$\pi_i = \mathrm{P}(A_i \mid L_i) = \begin{cases} \frac{1}{4} & \text{if } A_i = 1 \\ \frac{3}{4} & \text{if } A_i = 0 \end{cases}$$

Inverse probability weighting: Conditional randomizaton



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

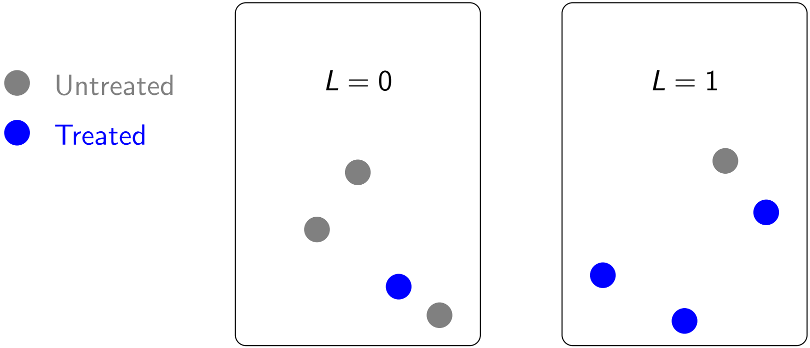
Inverse probability weighting: Conditional randomizaton



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

Each counts for:

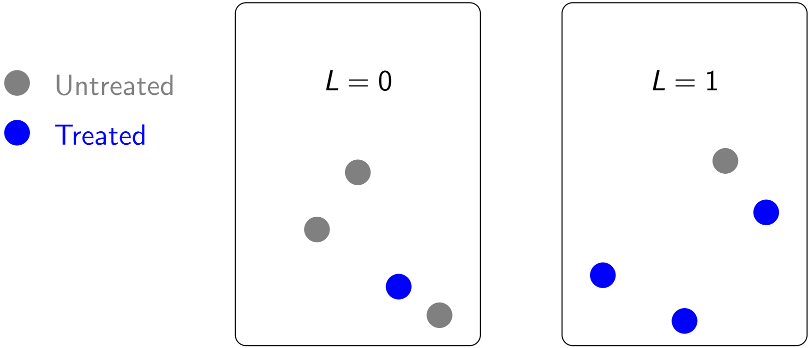
Inverse probability weighting: Conditional randomizaton



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

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



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

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 weighted (IPW) estimator
for the average treatment effect (ATE)

$$\mathrm{E}(Y^1) - \mathrm{E}(Y^0) = \frac{1}{N} \sum_{i:A_i=1} \frac{Y_i}{\pi_i} - \frac{1}{N} \sum_{i:A_i=0} \frac{Y_i}{\pi_i}$$

where $\pi_i = \mathrm{P}(A = a_i \mid \vec{L} = \vec{\ell}_i)$
is the probability of the observed treatment given confounders

Inverse probability weighting: Mathematical proof[1]

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

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

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

[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]

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

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

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

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

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

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

[1] Hernán & Robins Technical Point 2.3

Inverse probability weighting: Mathematical proof[1]

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

[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

Inverse probability weighting: Observational studies

In an experiment, we control treatment assignment.
We know the dashed edge does not exist.
We know the propensity score π.

$$
\begin{tikzcd}
 & & U \arrow[d] \arrow[dl, dashed, red] \\
\vec{L} \arrow[r] \arrow[rr, bend right] & A \arrow[r] & Y
\end{tikzcd}
$$

Inverse probability weighting: Observational studies

In an experiment, we control treatment assignment.
We know the dashed edge does not exist.
We know the propensity score π.

$$\begin{tikzcd}
 & & U \arrow[dl, dashed] \arrow[d] \\
\vec{L} \arrow[r] \arrow[rr, bend right] & A \arrow[r] & Y
\end{tikzcd}$$

In an observational study, we do not control assignment.

Inverse probability weighting: Observational studies

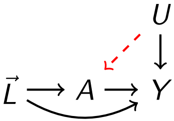
In an experiment, we control treatment assignment.
We know the dashed edge does not exist.
We know the propensity score π.

$$\begin{tikzcd}
 & & U \arrow[dl, dashed, red] \arrow[d] \\
\vec{L} \arrow[r] \arrow[rr, bend right] & A \arrow[r] & Y
\end{tikzcd}$$

In an observational study, we do not control assignment.
We assume the dashed edge does not exist.

Inverse probability weighting: Observational studies

In an experiment, we control treatment assignment.
We know the dashed edge does not exist.
We know the propensity score π.



In an observational study, we do not control assignment.
We assume the dashed edge does not exist.
We estimate the propensity score.

Inverse probability weighting: Observational studies

In an experiment, we control treatment assignment.
We know the dashed edge does not exist.
We know the propensity score π.

$$\begin{tikzcd}
 & & U \arrow[d] \arrow[dl, dashed, red] \\
\vec{L} \arrow[r] \arrow[rr, bend right] & A \arrow[r] & Y
\end{tikzcd}$$

In an observational study, we do not control assignment.
We assume the dashed edge does not exist.
We estimate the propensity score.

Otherwise all estimators (including IPW) are the same.

Inverse probability weighting: Nonparametric procedure

Inverse probability weighting: Nonparametric procedure

1. Assume a DAG where $\vec{L}$ blocks backdoor paths

Inverse probability weighting: Nonparametric procedure

1. Assume a DAG where $\vec{L}$ blocks backdoor paths
2. Stratify the by $\vec{L}$

Inverse probability weighting: Nonparametric procedure

1. Assume a DAG where $\vec{L}$ blocks backdoor paths
2. Stratify the by $\vec{L}$
3. Empirically estimate $\hat{\pi}_i$ as the proportion with $A = a_i$ within the stratum $\vec{L} = \vec{\ell}_i$ of unit i

Inverse probability weighting: Nonparametric procedure

1. Assume a DAG where $\vec{L}$ blocks backdoor paths
2. Stratify the by $\vec{L}$
3. Empirically estimate $\hat{\pi}_i$ as the proportion with $A = a_i$ within the stratum $\vec{L} = \vec{\ell}_i$ of unit i
4. Apply the IPW estimator

$$\hat{\text{E}}(Y^a) = \frac{1}{N} \sum_{i:A_i=a} \frac{Y_i}{\hat{\pi}_i}$$

Inverse probability weighting: Parametric procedure[2]

[2] On the Hajek estimator, see Hernán & Robins Technical Point 12.1

Inverse probability weighting: Parametric procedure[2]

1. Assume a DAG where $\vec{L}$ blocks backdoor paths

[2] On the Hajek estimator, see Hernán & Robins Technical Point 12.1

Inverse probability weighting: Parametric procedure[2]

1. Assume a DAG where $\vec{L}$ blocks backdoor paths
2. Estimate the propensity score $\hat{\pi}_i$ with a model

$$\hat{\text{P}}(A = 1 \mid \vec{L}) = \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right)$$

$$\hat{\pi}_i = \begin{cases} \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right) & \text{if } A_i = 1 \\ 1 - \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right) & \text{if } A_i = 0 \end{cases}$$

[2] On the Hajek estimator, see Hernán & Robins Technical Point 12.1

Inverse probability weighting: Parametric procedure[2]

1. Assume a DAG where $\vec{L}$ blocks backdoor paths
2. Estimate the propensity score $\hat{\pi}_i$ with a model

$$\hat{\text{P}}(A = 1 \mid \vec{L}) = \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right)$$

$$\hat{\pi}_i = \begin{cases} \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right) & \text{if } A_i = 1 \\ 1 - \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right) & \text{if } A_i = 0 \end{cases}$$

3. Apply an IPW estimator

[2] On the Hajek estimator, see Hernán & Robins Technical Point 12.1

Inverse probability weighting: Parametric procedure[2]

1. Assume a DAG where $\vec{L}$ blocks backdoor paths
2. Estimate the propensity score $\hat{\pi}_i$ with a model

$$\hat{\text{P}}(A = 1 \mid \vec{L}) = \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right)$$

$$\hat{\pi}_i = \begin{cases} \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right) & \text{if } A_i = 1 \\ 1 - \text{logit}^{-1}\left(\hat{\alpha} + \hat{\vec{\gamma}}\vec{L}\right) & \text{if } A_i = 0 \end{cases}$$

3. Apply an IPW estimator

$$\hat{\text{E}}(Y^a) = \frac{1}{N} \sum_{i:A_i=a} \frac{Y_i}{\hat{\pi}_i} \qquad \text{(Horvitz-Thompson)}$$

or

$$\hat{\text{E}}(Y^a) = \frac{1}{\sum_{i:A_i=a} \frac{1}{\hat{\pi}_i}} \sum_{i:A_i=a} \frac{Y_i}{\hat{\pi}_i} \qquad \text{(Hajek)}$$

[2] On the Hajek estimator, see Hernán & Robins Technical Point 12.1

Problem: Extreme weights create high variance

Problem: Extreme weights create high variance

Suppose a stratum $\vec{L} = \vec{\ell}$ contains

- 100 untreated units
- 1 treated unit

Problem: Extreme weights create high variance

Suppose a stratum $\vec{L} = \vec{\ell}$ contains

- 100 untreated units
- 1 treated unit

The treated unit gets a weight of 100.

Problem: Extreme weights create high variance

Suppose a stratum $\vec{L} = \vec{\ell}$ contains

- 100 untreated units
- 1 treated unit

The treated unit gets a weight of 100.

The estimate depends heavily on which treated unit happens to be included in the sample $\rightarrow$ high-variance estimator

Problem: Extreme weights create high variance

Suppose a stratum $\vec{L} = \vec{\ell}$ contains

- 100 untreated units
- 1 treated unit

The treated unit gets a weight of 100.

The estimate depends heavily on which treated unit happens to be included in the sample $\rightarrow$ high-variance estimator

Two solutions

1. Trim the weights
2. Truncate the weights

Problem: Extreme weights create high variance

Suppose a stratum $\vec{L} = \vec{\ell}$ contains

- 100 untreated units
- 1 treated unit

The treated unit gets a weight of 100.

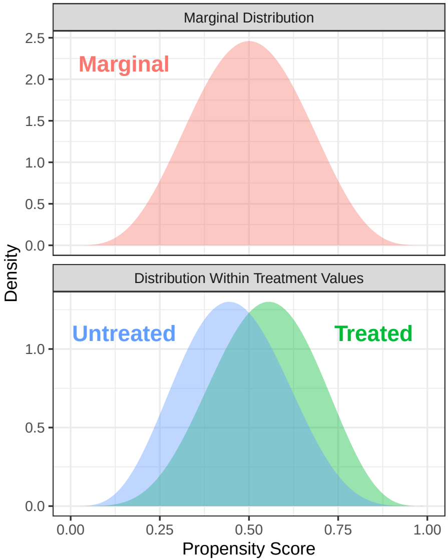
The estimate depends heavily on which treated unit happens to be included in the sample $\rightarrow$ high-variance estimator

Two solutions

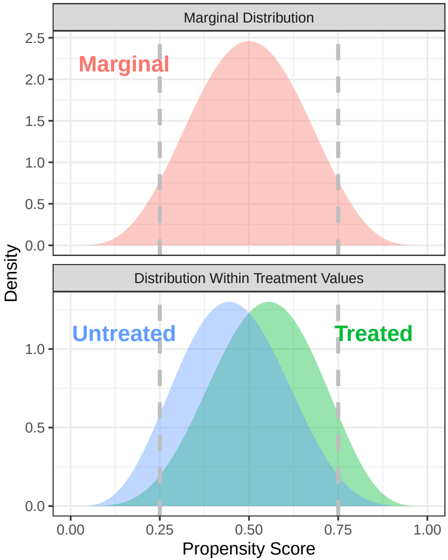
1. Trim the weights
2. Truncate the weights

Both solutions accept bias in order to reduce variance

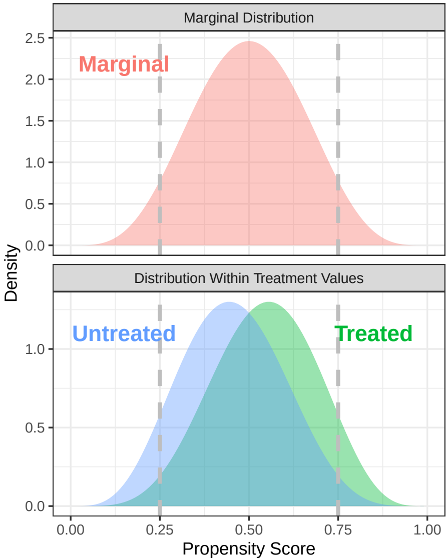
Accepting bias to reduce variance: Trimming



Accepting bias to reduce variance: Trimming

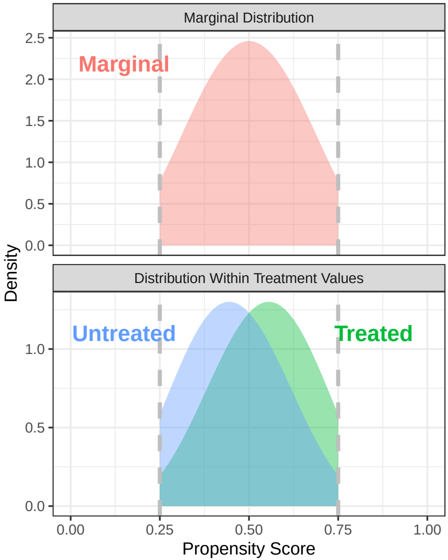


Accepting bias to reduce variance: Trimming



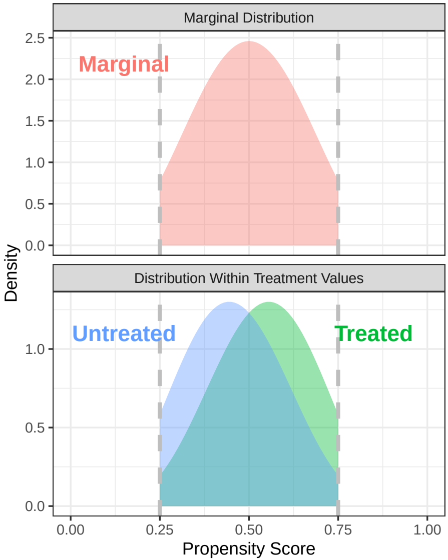
Drop units with extreme weights

Accepting bias to reduce variance: Trimming



Drop units with extreme weights

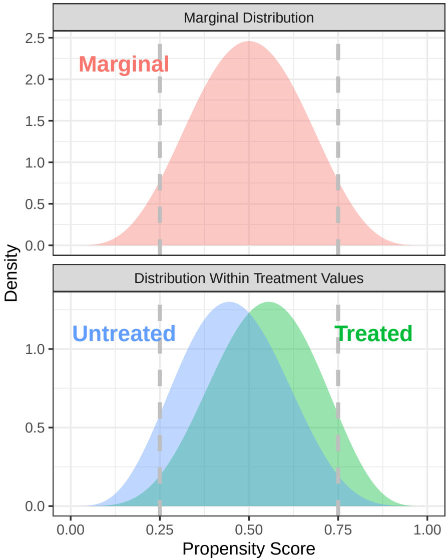
Accepting bias to reduce variance: Trimming



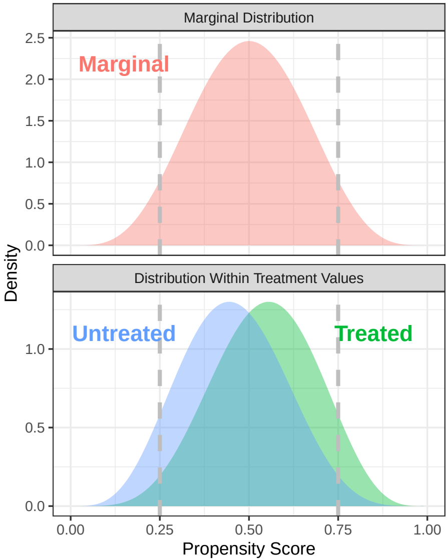
Drop units with extreme weights

Changes target population — Biased for full population

Accepting bias to reduce variance: Weight truncation

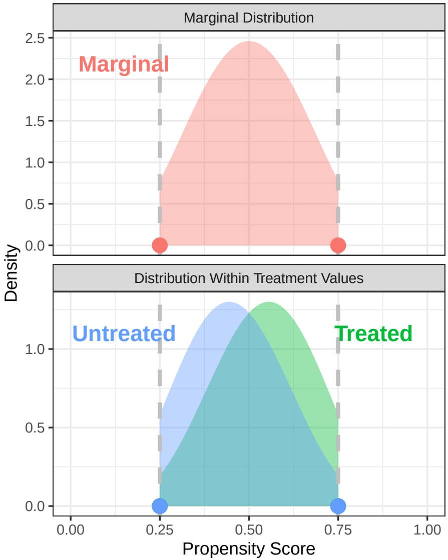


Accepting bias to reduce variance: Weight truncation



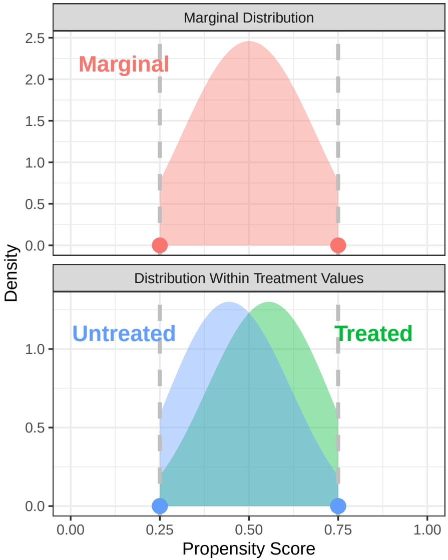
Truncate values of extreme weights

Accepting bias to reduce variance: Weight truncation



Truncate values of extreme weights

Accepting bias to reduce variance: Weight truncation

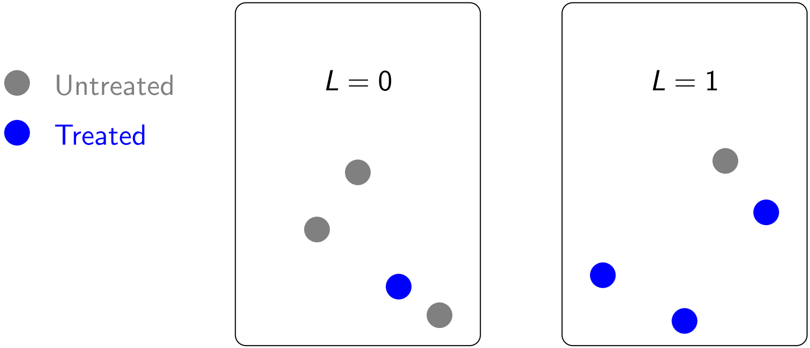


Truncate values of extreme weights

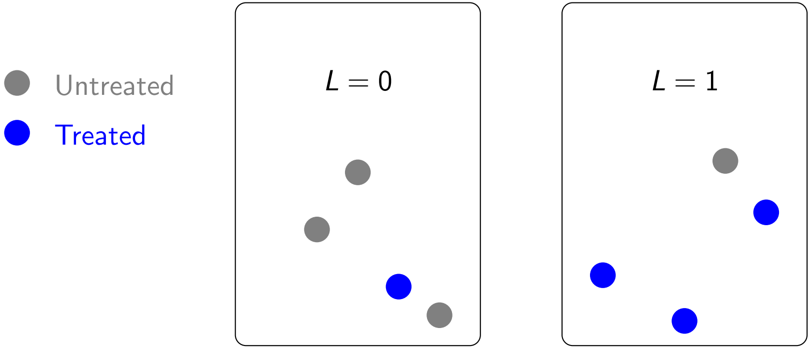
Biased: Ignores some confounding

Weighting for the ATE, ATT, ATC

Weighting for the ATE, ATT, ATC



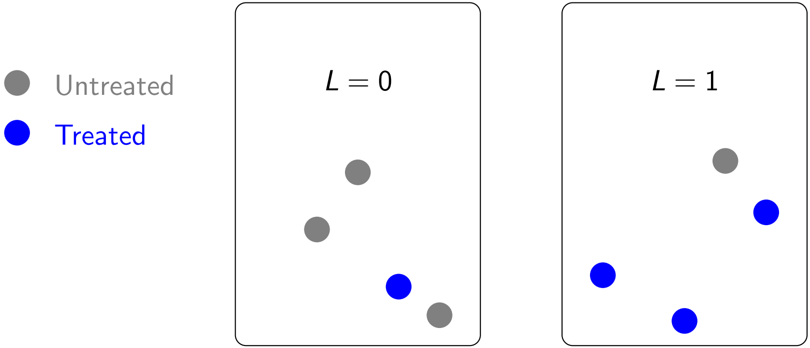
Weighting for the ATE, ATT, ATC



For ATE:

$w_i = \frac{1}{\mathrm{P}(A=a_i|L=\ell_i)}$

Weighting for the ATE, ATT, ATC

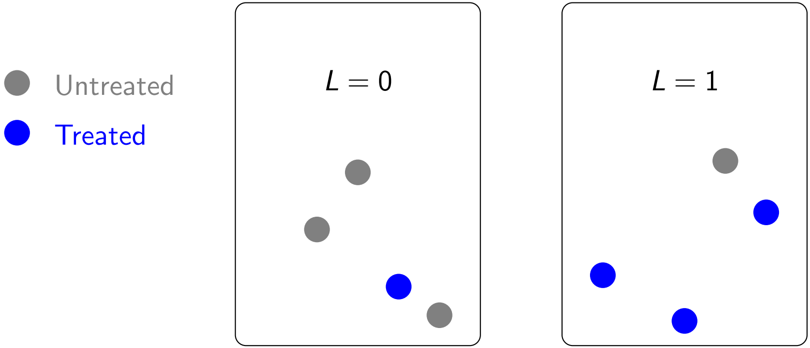


$$\frac{1}{3/4} = \frac{4}{3}$$

For ATE:

$w_i = \frac{1}{\text{P}(A=a_i|L=\ell_i)}$

Weighting for the ATE, ATT, ATC



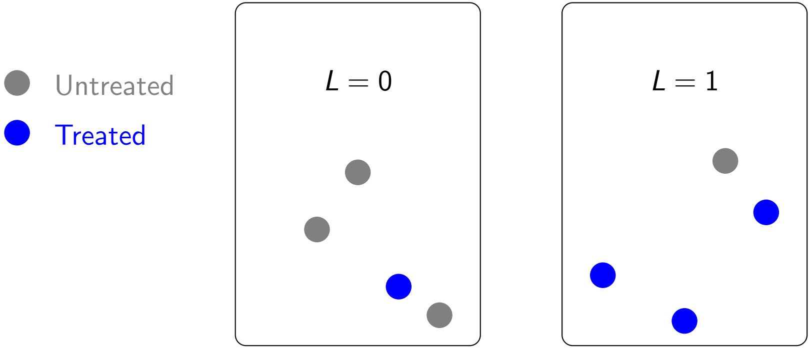
For ATE:

$w_i = \frac{1}{\mathrm{P}(A{=}a_i | L{=}\ell_i)}$

$$\frac{1}{3/4} = \frac{4}{3}$$

$$\frac{1}{1/4} = \frac{4}{1}$$

Weighting for the ATE, ATT, ATC



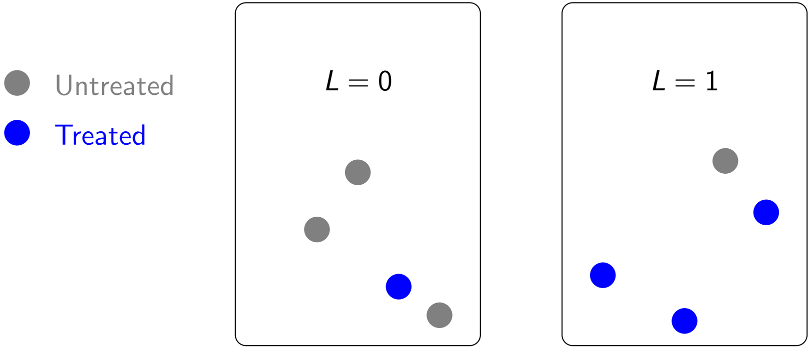
For ATE:

$w_i = \frac{1}{\mathrm{P}(A=a_i|L=\ell_i)}$

$$\frac{1}{3/4} = \frac{4}{3} \qquad \frac{1}{1/4} = \frac{4}{1}$$

$$\frac{1}{1/4} = \frac{4}{1}$$

Weighting for the ATE, ATT, ATC

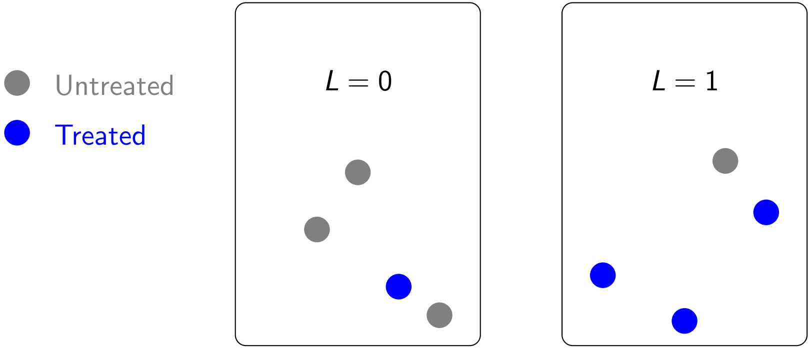


For ATE:

$w_i = \frac{1}{\text{P}(A=a_i|L=\ell_i)}$

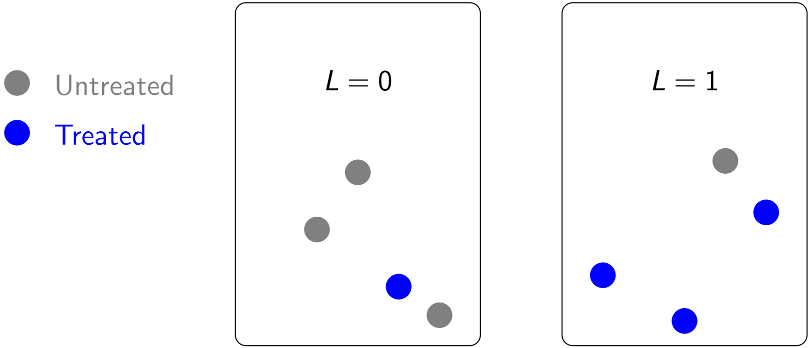
$L = 0$	$L = 1$
$\frac{1}{3/4} = \frac{4}{3}$	$\frac{1}{1/4} = \frac{4}{1}$
$\frac{1}{1/4} = \frac{4}{1}$	$\frac{1}{3/4} = \frac{4}{3}$

Weighting for the ATE, ATT, ATC



For ATT?

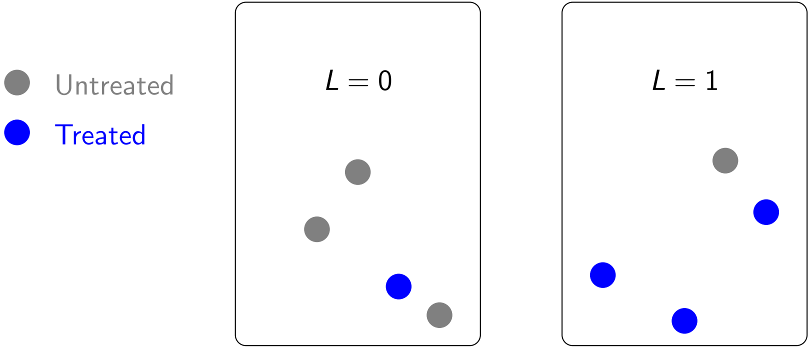
Weighting for the ATE, ATT, ATC



$$\frac{1}{3}$$

For ATT?

Weighting for the ATE, ATT, ATC

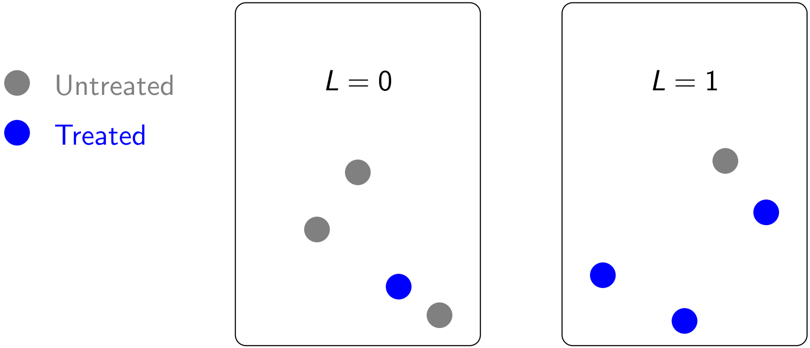


For ATT?

$\frac{1}{3}$

1

Weighting for the ATE, ATT, ATC

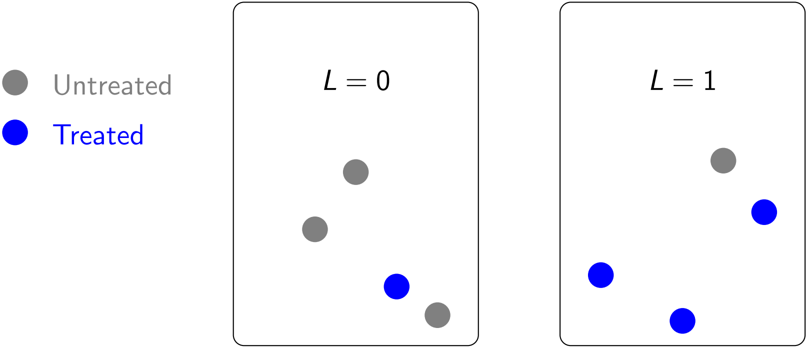


$\frac{1}{3}$ $\frac{3}{1}$

For ATT?

1

Weighting for the ATE, ATT, ATC

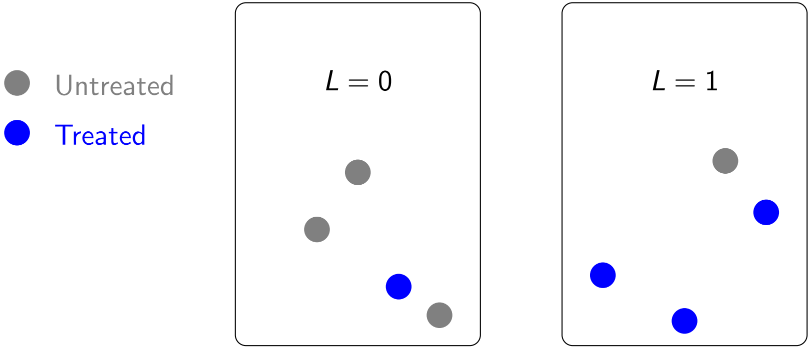


For ATT?

$\frac{1}{3}$ $\frac{3}{1}$

1 1

Weighting for the ATE, ATT, ATC

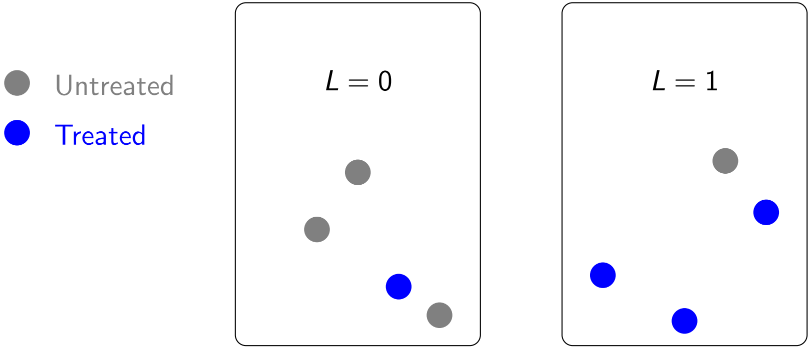


For ATT:
$w_i = \frac{P(A=1|L=\ell_i)}{P(A=a_i|L=\ell_i)}$

$\frac{1}{3}$ $\frac{3}{1}$

1 1

Weighting for the ATE, ATT, ATC

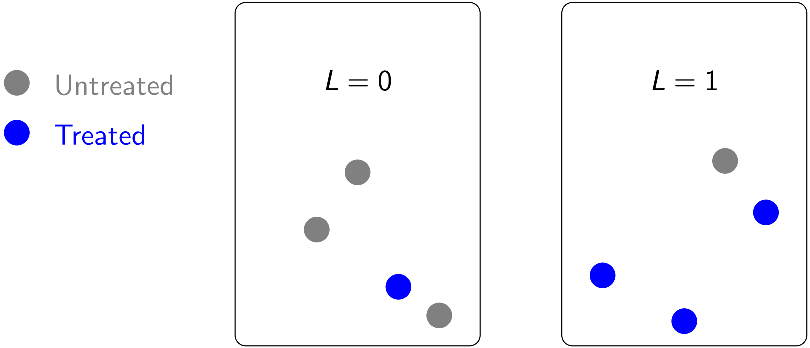


For ATT:
$w_i = \frac{P(A=1|L=\ell_i)}{P(A=a_i|L=\ell_i)}$

$L = 0$: $\frac{1/4}{3/4} = \frac{1}{3}$, $\frac{1/4}{1/4} = 1$

$L = 1$: $\frac{3/4}{1/4} = \frac{3}{1}$, $\frac{3/4}{3/4} = 1$

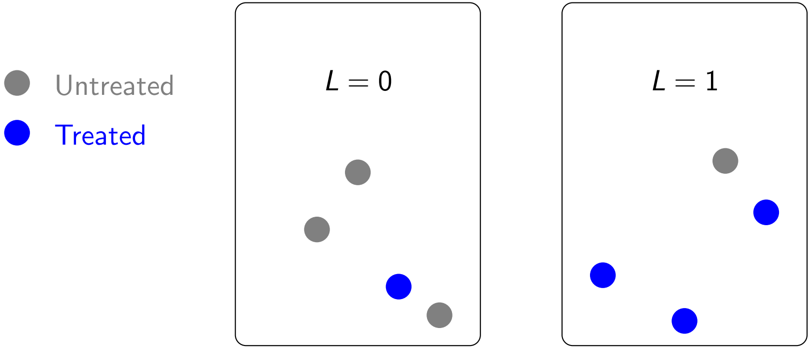
Weighting for the ATE, ATT, ATC



For ATC:

$w_i = \frac{\mathrm{P}(A{=}0 \mid L{=}\ell_i)}{\mathrm{P}(A{=}a_i \mid L{=}\ell_i)}$

Weighting for the ATE, ATT, ATC

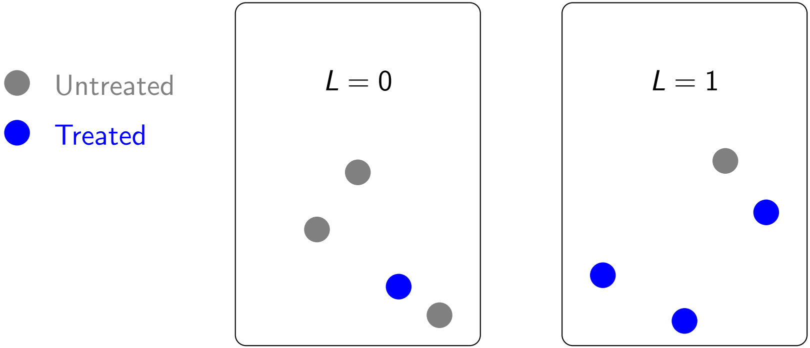


$$\frac{3/4}{3/4} = 1$$

For ATC:

$w_i = \frac{\text{P}(A=0|L=\ell_i)}{\text{P}(A=a_i|L=\ell_i)}$

Weighting for the ATE, ATT, ATC

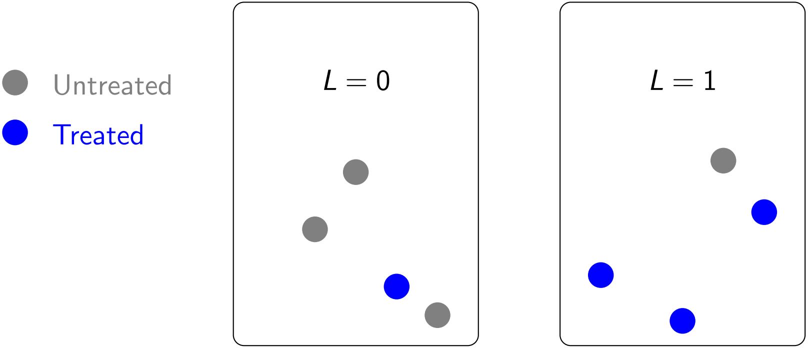


$$\frac{3/4}{3/4} = 1$$

For ATC:
$w_i = \frac{\mathrm{P}(A=0|L=\ell_i)}{\mathrm{P}(A=a_i|L=\ell_i)}$

$$\frac{3/4}{1/4} = 3$$

Weighting for the ATE, ATT, ATC



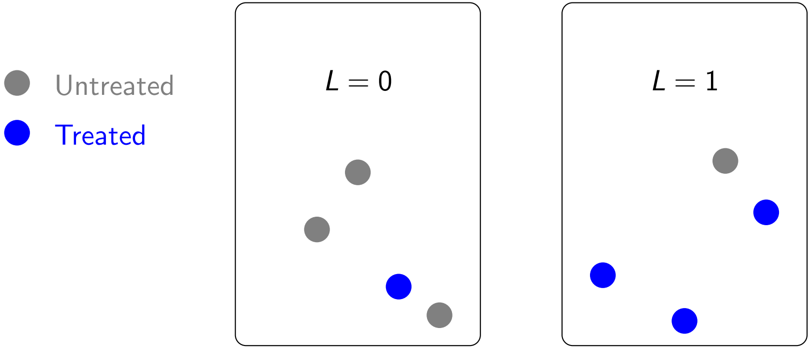
$$\frac{3/4}{3/4} = 1 \qquad \frac{1/4}{1/4} = 1$$

For ATC:

$$w_i = \frac{\mathrm{P}(A{=}0 \mid L{=}\ell_i)}{\mathrm{P}(A{=}a_i \mid L{=}\ell_i)}$$

$$\frac{3/4}{1/4} = 3$$

Weighting for the ATE, ATT, ATC



$L=0$: $\frac{3/4}{3/4} = 1$, $\frac{3/4}{1/4} = 3$

$L=1$: $\frac{1/4}{1/4} = 1$, $\frac{1/4}{3/4} = \frac{1}{3}$

For ATC:

$$w_i = \frac{\mathrm{P}(A=0 | L=\ell_i)}{\mathrm{P}(A=a_i | L=\ell_i)}$$

Weighting for the ATE, ATT, ATC

General formula:

$$w_i = \frac{\text{Size of target population in } \vec{L} = \vec{\ell}_i}{\text{Size of factual population in } \vec{L} = \vec{\ell}_i}$$

Intuition:

- Normalize the factual population across strata (denominator)
- Upweight to the counterfactual population (numerator)

Inverse probability weighting: Reading

Hernán & Robins

- 2.4
- 12.1–12.3

Learning goals for today

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

1. Trace inverse probability weighting to survey sampling
2. Apply the Horvitz-Thompson estimator for causal inference
3. Recognize the bias-variance tradeoff of trimmed weights
4. Estimate the ATE, ATT, and ATC by weighting

Let me know what you are thinking

tinyurl.com/CausalQuestions

Office hours TTh 11am-12pm and at
calendly.com/ianlundberg/office-hours
Come say hi!