Potential Outcomes

Brady Neal

causalcourse.com
What are potential outcomes?

The fundamental problem of causal inference

Getting around the fundamental problem of causal inference

A complete example with estimation
What are potential outcomes?

The fundamental problem of causal inference

Getting around the fundamental problem of causal inference

A complete example with estimation
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

Take pill
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

Take pill

Don’t take pill
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

Take pill → causal effect

Don’t take pill
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

Take pill

Don’t take pill

causal effect?
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

Take pill  →  no causal effect

Don’t take pill  →  no causal effect
Potential outcomes: notation

\[
do(T = 1)
\]

\[
do(T = 0)
\]

\[T\] : observed treatment
\[Y\] : observed outcome
Potential outcomes: notation

\[ Y_i \mid \text{do}(T=1) \]

\[ \text{do}(T = 1) \]

\[ \text{do}(T = 0) \]

\( T \) : observed treatment
\( Y \) : observed outcome
\( i \) : used in subscript to denote a specific unit/individual
Potential outcomes: notation

\[ Y_i | \text{do}(T=1) \]

\[ Y_i | \text{do}(T=0) \]

- \( T \): observed treatment
- \( Y \): observed outcome
- \( i \): used in subscript to denote a specific unit/individual

\[ \text{do}(T = 1) \]

\[ \text{do}(T = 0) \]
Potential outcomes: notation

\[ Y_i \mid \text{do}(T = 1) \triangleq Y_i(1) \]

- **do**(\(T = 1\))
  - Represents observed treatment
  - \(Y_i\) : observed outcome
  - \(i\) : used in subscript to denote a specific unit/individual
  - \(Y_i(1)\) : potential outcome under treatment
Potential outcomes: notation

\[ Y_i \big| \text{do}(T=1) \triangleq Y_i(1) \]

- \( T \): observed treatment
- \( Y \): observed outcome
- \( i \): used in subscript to denote a specific unit/individual
- \( Y_i(1) \): potential outcome under treatment
Potential outcomes: notation

\[ Y_i \big| \text{do}(T=1) \triangleq Y_i(1) \]

\[ Y_i \big| \text{do}(T=0) \triangleq Y_i(0) \]

- \( T \): observed treatment
- \( Y \): observed outcome
- \( i \): used in subscript to denote a specific unit/individual
- \( Y_i(1) \): potential outcome under treatment
- \( Y_i(0) \): potential outcome under no treatment
Potential outcomes: notation

\[
do(T = 1) \quad Y_i|\do(T=1) \triangleq Y_i(1)
\]

\[
do(T = 0) \quad Y_i|\do(T=0) \triangleq Y_i(0)
\]

- **\(T\)**: observed treatment
- **\(Y\)**: observed outcome
- **\(i\)**: used in subscript to denote a specific unit/individual
- **\(Y_i(1)\)**: potential outcome under treatment
- **\(Y_i(0)\)**: potential outcome under no treatment
Potential outcomes: notation

\[ Y_i \mid \text{do}(T=1) \triangleq Y_i(1) \]

\[ Y_i \mid \text{do}(T=0) \triangleq Y_i(0) \]

- \( T \): observed treatment
- \( Y \): observed outcome
- \( i \): used in subscript to denote a specific unit/individual
- \( Y_i(1) \): potential outcome under treatment
- \( Y_i(0) \): potential outcome under no treatment

Causal effect

\[ Y_i(1) - Y_i(0) \]
Potential outcomes: notation

\[ Y_i \mid \text{do}(T = 1) \triangleq Y_i(1) \]

- \( T \): observed treatment
- \( Y \): observed outcome
- \( i \): used in subscript to denote a specific unit/individual
- \( Y_i(1) \): potential outcome under treatment
- \( Y_i(0) \): potential outcome under no treatment

Causal effect

\[ Y_i(1) - Y_i(0) \]
Potential outcomes: notation

\[ Y_i(1) = 1 \]
\[ Y_i(0) = 0 \]

**Causal effect**
\[ Y_i(1) - Y_i(0) \]
Potential outcomes: notation

\[ T \]: observed treatment
\[ Y \]: observed outcome
\[ i \]: used in subscript to denote a specific unit/individual
\[ Y_i(1) \]: potential outcome under treatment
\[ Y_i(0) \]: potential outcome under no treatment

**Causal effect**

\[ Y_i(1) - Y_i(0) = 1 \]
What are potential outcomes?

The fundamental problem of causal inference

Getting around the fundamental problem of causal inference

A complete example with estimation
Fundamental problem of causal inference

\[ \text{do}(T = 1) \quad Y_i(1) = 1 \]

\[ \text{do}(T = 0) \quad Y_i(0) = 0 \]

\[ T \quad : \text{observed treatment} \]
\[ Y \quad : \text{observed outcome} \]
\[ i \quad : \text{used in subscript to denote a specific unit/individual} \]
\[ Y_i(1) \quad : \text{potential outcome under treatment} \]
\[ Y_i(0) \quad : \text{potential outcome under no treatment} \]

Causal effect
\[ Y_i(1) - Y_i(0) = 1 \]
Fundamental problem of causal inference

Counterfactual

do($T = 1$) \quad \Rightarrow \quad Y_i(1) = 1

do($T = 0$) \quad \Rightarrow \quad Y_i(0) = 0

Factual

Causal effect

\[ Y_i(1) - Y_i(0) = 1 \]

|$T$| observed treatment
|\(Y\)| observed outcome
|i\ | used in subscript to denote a specific unit/individual
|\(Y_i(1)\)| potential outcome under treatment
|\(Y_i(0)\)| potential outcome under no treatment

Then, what does imply causation?
Fundamental problem of causal inference

Factual

\[
\text{do}(T = 1) \quad Y_i(1) = 1
\]

Counterfactual

\[
\text{do}(T = 0) \quad Y_i(0) = 0
\]

Causal effect

\[
Y_i(1) - Y_i(0) = 1
\]

\(T\) : observed treatment
\(Y\) : observed outcome
\(i\) : used in subscript to denote a specific unit/individual
\(Y_i(1)\) : potential outcome under treatment
\(Y_i(0)\) : potential outcome under no treatment

Then, what does imply causation?
## Missing data interpretation

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T$</th>
<th>$Y$</th>
<th>$Y(1)$</th>
<th>$Y(0)$</th>
<th>$Y(1) - Y(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

$T$ : observed treatment  
$Y$ : observed outcome  
$i$ : used in subscript to denote a specific unit/individual  
$Y_i(1)$ : potential outcome under treatment  
$Y_i(0)$ : potential outcome under no treatment
Question:
What is the fundamental problem of causal inference?
What are potential outcomes?

The fundamental problem of causal inference

Getting around the fundamental problem of causal inference

A complete example with estimation
Average treatment effect (ATE)

\[ Y_i(1) - Y_i(0) \]

<table>
<thead>
<tr>
<th>( i )</th>
<th>( T )</th>
<th>( Y )</th>
<th>( Y(1) )</th>
<th>( Y(0) )</th>
<th>( Y(1) - Y(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

\( T \) : observed treatment  
\( Y \) : observed outcome  
\( i \) : used in subscript to denote a specific unit/individual  
\( Y_i(1) \) : potential outcome under treatment  
\( Y_i(0) \) : potential outcome under no treatment
# Average treatment effect (ATE)

\[ E[Y_i(1) - Y_i(0)] \]

<table>
<thead>
<tr>
<th>( i )</th>
<th>( T )</th>
<th>( Y )</th>
<th>( Y(1) )</th>
<th>( Y(0) )</th>
<th>( Y(1) - Y(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- \( T \) : observed treatment
- \( Y \) : observed outcome
- \( i \) : used in subscript to denote a specific unit/individual
- \( Y_i(1) \) : potential outcome under treatment
- \( Y_i(0) \) : potential outcome under no treatment
# Average treatment effect (ATE)

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

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- \(T\) : observed treatment
- \(Y\) : observed outcome
- \(i\) : used in subscript to denote a specific unit/individual
- \(Y_i(1)\) : potential outcome under treatment
- \(Y_i(0)\) : potential outcome under no treatment
### Average treatment effect (ATE)

\[
\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]
\]

<table>
<thead>
<tr>
<th></th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- \(T\) : observed treatment
- \(Y\) : observed outcome
- \(i\) : used in subscript to denote a specific unit/individual
- \(Y_i(1)\) : potential outcome under treatment
- \(Y_i(0)\) : potential outcome under no treatment
Average treatment effect (ATE)

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]$$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T$</th>
<th>$Y$</th>
<th>$Y(1)$</th>
<th>$Y(0)$</th>
<th>$Y(1) - Y(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- $T$: observed treatment
- $Y$: observed outcome
- $i$: used in subscript to denote a specific unit/individual
- $Y_i(1)$: potential outcome under treatment
- $Y_i(0)$: potential outcome under no treatment
Average treatment effect (ATE)

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0] \]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

- \(T\) : observed treatment
- \(Y\) : observed outcome
- \(i\) : used in subscript to denote a specific unit/individual
- \(Y_i(1)\) : potential outcome under treatment
- \(Y_i(0)\) : potential outcome under no treatment
Average treatment effect (ATE)

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0] \]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\(T\) : observed treatment
\(Y\) : observed outcome
\(i\) : used in subscript to denote a specific unit/individual
\(Y_i(1)\) : potential outcome under treatment
\(Y_i(0)\) : potential outcome under no treatment

\(\frac{2}{3}\)
Average treatment effect (ATE)

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0] \]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\(T\) : observed treatment  
\(Y\) : observed outcome  
\(i\) : used in subscript to denote a specific unit/individual  
\(Y_i(1)\) : potential outcome under treatment  
\(Y_i(0)\) : potential outcome under no treatment  

\[ \frac{2}{3} \] \[ \frac{1}{3} \]
Average treatment effect (ATE)

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0]$$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T$</th>
<th>$Y$</th>
<th>$Y(1)$</th>
<th>$Y(0)$</th>
<th>$Y(1) - Y(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td></td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>?</td>
</tr>
</tbody>
</table>

$$\frac{2}{3} - \frac{1}{3} = \frac{1}{3}$$

$T$: observed treatment
$Y$: observed outcome
$i$: used in subscript to denote a specific unit/individual
$Y_i(1)$: potential outcome under treatment
$Y_i(0)$: potential outcome under no treatment
Average treatment effect (ATE)

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] \]

<table>
<thead>
<tr>
<th>( i )</th>
<th>( T )</th>
<th>( Y )</th>
<th>( Y(1) )</th>
<th>( Y(0) )</th>
<th>( Y(1) - Y(0) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\[ \frac{2}{3} - \frac{1}{3} = \frac{1}{3} \]

- \( T \): observed treatment
- \( Y \): observed outcome
- \( i \): used in subscript to denote a specific unit/individual
- \( Y_i(1) \): potential outcome under treatment
- \( Y_i(0) \): potential outcome under no treatment
# Average treatment effect (ATE)

\[
\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\[
\frac{2}{3} - \frac{1}{3} = \frac{1}{3}
\]

- \(T\): observed treatment
- \(Y\): observed outcome
- \(i\): used in subscript to denote a specific unit/individual
- \(Y_i(1)\): potential outcome under treatment
- \(Y_i(0)\): potential outcome under no treatment
# Average treatment effect (ATE)

\[
\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] 
\neq \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>i</th>
<th>T</th>
<th>Y</th>
<th>Y(1)</th>
<th>Y(0)</th>
<th>Y(1) - Y(0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>?</td>
</tr>
</tbody>
</table>

\[
\frac{2}{3} - \frac{1}{3} = \frac{1}{3}
\]
Association is not causation

Sleeping with shoes on is strongly correlated with waking up with a headache
Association is not causation

Sleeping with shoes on is strongly correlated with waking up with a headache.

Common cause: drinking the night before.
Association is not causation

Sleeping with shoes on is strongly correlated with waking up with a headache

Common cause: drinking the night before

1. Confounding
Association is not causation

Sleeping with shoes on is strongly correlated with waking up with a headache

Common cause: drinking the night before

1. Confounding
Association is not causation

Sleeping with shoes on is strongly correlated with waking up with a headache

Common cause: drinking the night before
1. Confounding
2. Shoe-sleepers differ from non-shoe-sleepers in a key way
Why? Because the groups are not comparable

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \neq \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0] \]

Went to sleep with shoes on
\((T = 1)\)

Went to sleep without shoes on
\((T = 0)\)
Why? Because the groups are not comparable

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \neq \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

Went to sleep **with shoes** on \((T = 1)\)

Went to sleep **without shoes** on \((T = 0)\)
Why? Because the groups are not comparable

$$E[Y(1)] - E[Y(0)] \neq E[Y \mid T = 1] - E[Y \mid T = 0]$$

Went to sleep with shoes on

\[ (T = 1) \]

Went to sleep without shoes on

\[ (T = 0) \]
Why? Because the groups are not comparable

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \neq \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] \]

Went to sleep **with shoes** on

\((T = 1)\)

Went to sleep **without shoes** on

\((T = 0)\)

**r/unpopularopinion**

Posted by u/jimmythang34 1 year ago 🔴

**2.1k**

**sleeping with shoes on is comfortable**

sober  drunk

sober  sober  sober  drunk

sober  sober  sober  sober

sober  sober  sober  sober

sober  sober  sober  drunk

sober  sober  sober  sober

sober  sober  sober  sober

sober  sober  sober  sober
What would comparable groups look like?

\[ E[Y(1)] - E[Y(0)] \neq E[Y \mid T = 1] - E[Y \mid T = 0] \]

Went to sleep **with shoes** on

\[ (T = 1) \]

Went to sleep **without shoes** on

\[ (T = 0) \]
What would comparable groups look like?

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \neq \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] \]

Went to sleep **with shoes** on

\(T = 1\)

Went to sleep **without shoes** on

\(T = 0\)
What would comparable groups look like?

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0] \]

Went to sleep **with shoes** on

(T = 1)

- drunk
- sober
- drunk
- sober
- drunk
- sober

Went to sleep **without shoes** on

(T = 0)

- drunk
- sober
- drunk
- sober
- drunk
- sober
- drunk
- sober
Question:
Why is association not causation?
What assumptions would make the ATE equal to the associational difference?
Ignorability: \((Y(1), Y(0)) \perp T\)
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)}
\]
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
E[Y(1)] - E[Y(0)] = E[Y(1) | T = 1] - E[Y(0) | T = 0] \quad \text{(ignorability)}
\]

\[
= E[Y | T = 1] - E[Y | T = 0]
\]
**Ignorability:** \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)}
\]

\[
= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>0</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>?</td>
<td>1</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
\begin{align*}
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] &= \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)} \\
&= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\end{align*}
\]

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>
Ignorability: \((Y(1), Y(0)) \perp T\)

\[\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0]\]  
(ignorability)

\[= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

\[\frac{2}{3}\]
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) | T = 1] - \mathbb{E}[Y(0) | T = 0] \quad \text{(ignorability)}
\]

\[
= \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0]
\]
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)}
\]

\[
= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\[
\frac{2}{3} - \frac{1}{3} = \frac{1}{3}
\]
Ignorability: $(Y(1), Y(0)) \perp T$

$$E[Y(1)] - E[Y(0)] = E[Y(1) \mid T = 1] - E[Y(0) \mid T = 0] \quad \text{(ignorability)}$$

$$= E[Y \mid T = 1] - E[Y \mid T = 0]$$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$T$</th>
<th>$Y$</th>
<th>$Y(1)$</th>
<th>$Y(0)$</th>
<th>$Y(1) - Y(0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

$$\frac{2}{3} - \frac{1}{3} = \frac{1}{3}$$
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)}
\]

\[
= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\[\frac{2}{3} - \frac{1}{3} = \frac{1}{3}\]
Ignorability: \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)}
\]

\[
= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

\[
\frac{2}{3} - \frac{1}{3} = \frac{1}{3}
\]
**Ignorability:** \((Y(1), Y(0)) \perp T\)

\[
\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)}
\]

\[
= \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]
\]

<table>
<thead>
<tr>
<th>(i)</th>
<th>(T)</th>
<th>(Y)</th>
<th>(Y(1))</th>
<th>(Y(0))</th>
<th>(Y(1) - Y(0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>?</td>
<td></td>
</tr>
</tbody>
</table>

\[
\frac{2}{3} - \frac{1}{3} = \frac{1}{3}
\]
Another perspective: exchangeability

\[ T = 1 \]

\[ \mathbb{E}[Y \mid T = 1] = y_1 \]

\[ T = 0 \]

\[ \mathbb{E}[Y \mid T = 0] = y_0 \]
Another perspective: exchangeability

\[ T = 1 \]

Group A

\[ \mathbb{E}[Y \mid T = 1] = y_1 \]

\[ T = 0 \]

Group B

\[ \mathbb{E}[Y \mid T = 0] = y_0 \]
Another perspective: exchangeability

\[ T = 1 \]

Group B

\[ \mathbb{E}[Y \mid T = 1] = y_1 \]

\[ T = 0 \]

Group A

\[ \mathbb{E}[Y \mid T = 0] = y_0 \]
Another perspective: exchangeability

\[ T = 1 \]

\[ \mathbb{E}[Y \mid T = 1] = y_1 \]

Group B

\[ T = 0 \]

\[ \mathbb{E}[Y \mid T = 0] = y_0 \]

Group A
Another perspective: exchangeability

\[ T = 1 \]

\( \mathbb{E}[Y \mid T = 1] = y_1 \)

Group B

Before switch

\[ \mathbb{E}[Y(1) \mid T = 1] = \mathbb{E}[Y(1) \mid T = 0] \]

\[ T = 0 \]

\( \mathbb{E}[Y \mid T = 0] = y_0 \)

Group A

After switch
Another perspective: exchangeability

\[ T = 1 \]
\[ \mathbb{E}[Y \mid T = 1] = y_1 \]

Group B

Before switch
\[ \mathbb{E}[Y(1) \mid T = 1] = \mathbb{E}[Y(1) \mid T = 0] = \mathbb{E}[Y(1)] \]

\[ T = 0 \]
\[ \mathbb{E}[Y \mid T = 0] = y_0 \]

Group A

After switch
Another perspective: exchangeability

\[ T = 1 \]

Group B

\[ \mathbb{E}[Y \mid T = 1] = y_1 \]

\[ T = 0 \]

Group A

\[ \mathbb{E}[Y \mid T = 0] = y_0 \]

Before switch

\[ \mathbb{E}[Y(1) \mid T = 1] = \mathbb{E}[Y(1) \mid T = 0] = \mathbb{E}[Y(1)] \]

\[ \mathbb{E}[Y(0) \mid T = 0] = \mathbb{E}[Y(0) \mid T = 1] = \mathbb{E}[Y(0)] \]

After switch
Aside: identifiability

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)} \]

\[ = \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] \]
Aside: identifiability

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T = 1] - \mathbb{E}[Y(0) \mid T = 0] \quad \text{(ignorability)} \]

Causal quantities

\[ = \mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0] \]
Aside: identifiability

\[ E[Y(1)] - E[Y(0)] = E[Y(1) | T = 1] - E[Y(0) | T = 0] \]  
(ignorability)

\[ = E[Y | T = 1] - E[Y | T = 0] \]

Causal quantities

Statistical quantities

(accessible, since we have \( P(x, t, y) \))
Aside: identifiability

\[ \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) | T = 1] - \mathbb{E}[Y(0) | T = 0] \quad \text{(ignorability)} \]

\[ = \mathbb{E}[Y | T = 1] - \mathbb{E}[Y | T = 0] \]

A causal quantity (e.g. \( \mathbb{E}[Y(t)] \)) is **identifiable** if we can compute it from a purely statistical quantity (e.g. \( \mathbb{E}[Y | t] \))
Randomized control trial (RCT)
Randomized control trial (RCT)

Went to sleep with shoes on \((T = 1)\)

Went to sleep without shoes on \((T = 0)\)
Randomized control trial (RCT)

Slept with shoes on $(T = 1)$

Went to sleep without shoes on $(T = 0)$
Randomized control trial (RCT)

\[ T = 1 \]  vs  \[ T = 0 \]

Slept with shoes on

Slept without shoes on

Brady Neal

Page 20 / 41
Randomized control trial (RCT)

Slept with shoes on $(T = 1)$

Slept without shoes on $(T = 0)$
Graphical interpretation of RCT
Graphical interpretation of RCT
Graphical interpretation of RCT
Graphical interpretation of RCT
Question: What important property does an RCT give us?
Conditional exchangeability

Exchangeability:

\[(Y(1), Y(0)) \perp T\]
Conditional exchangeability

Exchangeability:
\[(Y(1), Y(0)) \perp\!
\perp T\]
Conditional exchangeability

Exchangeability:
\((Y(1), Y(0)) \perp T\)
Conditional exchangeability

Exchangeability:
\[(Y(1), Y(0)) \perp T\]

Conditional exchangeability:
\[(Y(1), Y(0)) \perp T \mid X\]
Conditional exchangeability

Exchangeability:
\[(Y(1), Y(0)) \perp T\]

Conditional exchangeability:
\[(Y(1), Y(0)) \perp T \mid X\]
Identification of conditional average treatment effect

Conditional exchangeability: \((Y(1), Y(0)) \perp T \mid X\)
Identification of conditional average treatment effect

Conditional exchangeability: \((Y(1), Y(0)) \perp T \mid X\)

\[\mathbb{E}[Y(1) - Y(0) \mid X] = \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]\]
Identification of conditional average treatment effect

Conditional exchangeability: \((Y(1), Y(0)) \perp T \mid X\)

\[\mathbb{E}[Y(1) - Y(0) \mid X] = \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]\]
\[= \mathbb{E}[Y(1) \mid T = 1, X] - \mathbb{E}[Y(0) \mid T = 0, X]\]
Identification of conditional average treatment effect

Conditional exchangeability: \((Y(1), Y(0)) \perp T \mid X\)

\[
\mathbb{E}[Y(1) - Y(0) \mid X] = \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]
\]
\[
= \mathbb{E}[Y(1) \mid T = 1, X] - \mathbb{E}[Y(0) \mid T = 0, X]
\]
\[
= \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]
\]
Identification of conditional average treatment effect

Conditional exchangeability: \((Y(1), Y(0)) \perp T | X\)

\[
\mathbb{E}[Y(1) - Y(0) | X] = \mathbb{E}[Y(1) | X] - \mathbb{E}[Y(0) | X]
\]
\[
= \mathbb{E}[Y(1) | T = 1, X] - \mathbb{E}[Y(0) | T = 0, X]
\]
\[
= \mathbb{E}[Y | T = 1, X] - \mathbb{E}[Y | T = 0, X]
\]

What about the ATE?  \(\mathbb{E}[Y(1) - Y(0)]\)
The Adjustment Formula (identification of ATE)

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) \mid X] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]] \]
The Adjustment Formula (identification of ATE)

\[
E[Y(1) - Y(0)] = E_X E[Y(1) - Y(0) | X] \\
= E_X [E[Y | T = 1, X] - E[Y | T = 0, X]]
\]
The Adjustment Formula (identification of ATE)

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) \mid X] \]
\[ = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]] \]
The Adjustment Formula (identification of ATE)

\[
\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) \mid X]
= \mathbb{E}_X \left[ \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X] \right]
\]
Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability
Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

\[(Y(1), Y(0)) \perp T \mid X\]
Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

\[(Y(1), Y(0)) \perp T \mid X\]
Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:
\[(Y(1), Y(0)) \perp T \mid X\]
Positivity

For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),

$$0 < P(T = 1 \mid X = x) < 1$$
Positivity

For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),

$$0 < P(T = 1 \mid X = x) < 1$$

Why? Recall the adjustment formula:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$$
Positivity

For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),

$$0 < P(T = 1 \mid X = x) < 1$$

Why? Recall the adjustment formula:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$$

$$\sum_x \left( \sum_y y P(Y = y \mid T = 1, X = x) - \sum_y y P(Y = y \mid T = 0, X = x) \right)$$
Positivity

For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),

$$0 < P(T = 1 \mid X = x) < 1$$

Why? Recall the adjustment formula:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[ \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X] \right]$$

$$\sum_x \left( \sum_y y P(Y = y \mid T = 1, X = x) - \sum_y y P(Y = y \mid T = 0, X = x) \right)$$

$$\sum_x \left( \sum_y \frac{P(Y = y, T = 1, X = x)}{P(T = 1 \mid X = x)P(X = x)} - \sum_y \frac{P(Y = y, T = 0, X = x)}{P(T = 0 \mid X = x)P(X = x)} \right)$$
Positivity

For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),

$$0 < P(T = 1 \mid X = x) < 1$$

Why? Recall the adjustment formula:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$$

$$\sum_x \left( \sum_y y P(Y = y \mid T = 1, X = x) - \sum_y y P(Y = y \mid T = 0, X = x) \right)$$

$$\sum_x \left( \sum_y \frac{P(Y = y, T = 1, X = x)}{P(T = 1 \mid X = x)P(X = x)} - \sum_y \frac{P(Y = y, T = 0, X = x)}{P(T = 0 \mid X = x)P(X = x)} \right)$$
Positivity

For all values of covariates $x$ present in the population of interest (i.e. $x$ such that $P(X = x) > 0$),

$$0 < P(T = 1 \mid X = x) < 1$$

Why? Recall the adjustment formula:

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[ \mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X] \right]$$

$$\sum_x \left( \sum_y y P(Y = y \mid T = 1, X = x) - \sum_y y P(Y = y \mid T = 0, X = x) \right)$$

$$\sum_x \left( \sum_y y \frac{P(Y = y, T = 1, X = x)}{P(T = 1 \mid X = x)P(X = x)} - \sum_y y \frac{P(Y = y, T = 0, X = x)}{P(T = 0 \mid X = x)P(X = x)} \right)$$
Positivity: intuition

Total population
Positivity: intuition

Total population

\[ X = x \]
Positivity: intuition

Total population

\[ X = x \]

No one treated
Positivity: intuition

Total population

Everyone treated

\[ X = x \]

\[ T = 1 \]
\[ T = 1 \]
\[ T = 1 \]
Another perspective: overlap

Overlap between $P(X \mid T = 0)$ and $P(X \mid T = 1)$
Another perspective: overlap

\[ P(x \mid t) \]

\[ P(X \mid T = 0) \]

\[ P(X \mid T = 1) \]
Another perspective: overlap

No overlap means severe positivity violation

\[ P(x \mid t) \]

\[ P(X \mid T = 0) \quad P(X \mid T = 1) \]
Another perspective: overlap

\[ P(x \mid t) \]
Another perspective: overlap

\[ P(x \mid t) \]
Another perspective: overlap

\[ P(x \mid t) \]
Another perspective: overlap
Another perspective: overlap

Complete overlap means no positivity violation

\[ P(x \mid t) \]
Question:
What goes wrong if we don’t have positivity?
The Positivity-Unconfoundedness Tradeoff

\[ P(X \mid T = 0) \quad \text{and} \quad P(X \mid T = 1) \]
The Positivity-Unconfoundedness Tradeoff

\[ x \]
The Positivity-Unconfoundedness Tradeoff

50% overlap
The Positivity-Unconfoundedness Tradeoff

1-dimensional

50% overlap
The Positivity-Unconfoundedness Tradeoff

2-dimensional

1-dimensional

50% overlap
The Positivity-Unconfoundedness Tradeoff

2-dimensional $x_2$

1-dimensional

25% overlap

50% overlap
The Positivity-Unconfoundedness Tradeoff

- 25% overlap (2-dimensional)
- 50% overlap (1-dimensional)
- 3-dimensional – 12.5% overlap
The Positivity-Unconfoundedness Tradeoff

2-dimensional

1-dimensional

25% overlap

50% overlap

3-dimensional – 12.5% overlap

... and so on
(curse of dimensionality)
Extrapolation

$T = 0$ \quad $x$ \quad $T = 1$
Extrapolation

Adjustment formula: \[ \sum_x (\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]) \]
Extrapolation

Adjustment formula: \[ \sum_{x} (\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]) \]

Model with \( f_1(x) \)
Extrapolation

Adjustment formula:

\[ \sum_x \left( \mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x] \right) \]

Model with \( T = 1 \) and \( f_1(x) \)

Model with \( T = 0 \) and \( f_0(x) \)
Extrapolation

Adjustment formula: \[ \sum_x \left( \mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x] \right) \]

Model with \( f_1(x) \) and \( f_0(x) \)

Model with \( f_0(x) \)
Extrapolation

Adjustment formula: \[ \sum_{x} (\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]) \]

Model with \( f_1(x) \)
Model with \( f_0(x) \)

\( f_1(x) \) and \( f_0(x) \)
Extrapolation

Adjustment formula: \[ \sum_{x} (\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]) \] with

Model with \[ f_1(x) \]

and

Model with \[ f_0(x) \]
Extrapolation

Adjustment formula: \[ \sum_x (E[Y \mid T = 1, x] - E[Y \mid T = 0, x]) \]

Model with \( f_1(x) \)  
Model with \( f_0(x) \)

and \( f_1(x) \) and \( f_0(x) \)
Extrapolation

Adjustment formula: \[
\sum_x \left( \mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x] \right) \]

Model with \( f_1(x) \)

Model with \( f_0(x) \)
Extrapolation

Adjustment formula: \[ \sum_x (\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]) \]

Model with \( f_1(x) \)
Model with \( f_0(x) \)

\[ f_0(x) \]
\[ f_1(x) \]

\( T = 0 \)
\( x \)
\( T = 1 \)
Extrapolation

Adjustment formula: \[ \sum_x (\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]) \]

with \[ f_1(x) \]

and \[ f_0(x) \]
No interference

\[ Y_i(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = Y_i(t_i) \]
No interference

\[ Y_i(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = Y_i(t_i) \]
No interference

\[ Y_i(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = Y_i(t_i) \]
No interference

\[ Y_i(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = Y_i(t_i) \]

Whether friends get dogs

\[ T_1 \quad \ldots \quad T_{i-1} \quad T_i \quad T_{i+1} \quad \ldots \quad T_n \]

My happiness

Whether friends get dogs
No interference

\[ Y_i(t_1, \ldots, t_{i-1}, t_i, t_{i+1}, \ldots, t_n) = Y_i(t_i) \]
Consistency: \( T = t \implies Y = Y(t) \)
Consistency: \( T = t \implies Y = Y(t) \)

\[ T = 1 \]

“\text{I get a dog}”
Consistency: \[ T = t \implies Y = Y(t) \]

\[
\begin{array}{c|c}
T = 1 & T = 0 \\
\text{“I get a dog”} & \text{“I don’t get a dog”} \\
\end{array}
\]
Consistency: \[ T = t \implies Y = Y(t) \]

\[ \begin{align*}
T &= 1 \\
\text{“I get a dog”} & \quad T = 0 \\
\text{“I don’t get a dog”}
\end{align*} \]

\( (T = 1) \implies Y = 1 \text{ (I’m happy)} \)
Consistency:  \( T = t \implies Y = Y(t) \)

- \( T = 1 \)
  - “I get a dog”
- \( T = 0 \)
  - “I don’t get a dog”

\( (T = 1) \implies Y = 1 \) (I’m happy)

\( (T = 1) \implies Y = 0 \) (I’m not happy)
Consistency: \( T = t \implies Y = Y(t) \)

\( T = 1 \)

“I get a dog”

\( T = 0 \)

“I don’t get a dog”

\(( T = 1 \) \implies Y = 1 \) (I’m happy)

\(( T = 1 \) \implies Y = 0 \) (I’m not happy)

Consistency assumption violated
Recall:
1. What were the four main assumptions?
2. Why do positivity violations require extrapolation?
3. Can you test if unconfoundedness is satisfied?
4. What is identifiability?
Tying it all together

\[ \mathbb{E}[Y(1) - Y(0)] \]
Tying it all together

No interference

\[ E[Y(1) - Y(0)] \]
Tying it all together

No interference

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \] (linearity of expectation)
Tying it all together

No interference

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \quad \text{(linearity of expectation)} \]
\[ = \mathbb{E}_X \left[ \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X] \right] \quad \text{(law of iterated expectations)} \]
Tying it all together

No interference

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \]  
\( \text{(linearity of expectation)} \)

\[ = \mathbb{E}_X \left[ \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X] \right] \]  
\( \text{(law of iterated expectations)} \)

\[ = \mathbb{E}_X \left[ \mathbb{E}[Y(1) \mid T = 1, X] - \mathbb{E}[Y(0) \mid T = 0, X] \right] \]  
\( \text{(unconfoundedness and positivity)} \)
Tying it all together

No interference

\[ \mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \]  
(linearity of expectation)

\[ = \mathbb{E}_X [\mathbb{E}[Y(1) | X] - \mathbb{E}[Y(0) | X]] \]  
(law of iterated expectations)

\[ = \mathbb{E}_X [\mathbb{E}[Y(1) | T = 1, X] - \mathbb{E}[Y(0) | T = 0, X]] \]  
(unconfoundedness and positivity)

\[ = \mathbb{E}_X [\mathbb{E}[Y | T = 1, X] - \mathbb{E}[Y | T = 0, X]] \]  
(consistency)
What are potential outcomes?

The fundamental problem of causal inference

Getting around the fundamental problem of causal inference

A complete example with estimation
Estimands, estimates, and the Identification-Estimation Flowchart
Estimands, estimates, and the Identification-Estimation Flowchart

• Estimand - any quantity we want to estimate
Estimands, estimates, and the Identification-Estimation Flowchart

• Estimand - any quantity we want to estimate
  • Causal estimand (e.g. $\mathbb{E}[Y(1) – Y(0)]$)
Estimands, estimates, and the Identification-Estimation Flowchart

• Estimand - any quantity we want to estimate
  • Causal estimand (e.g. $\mathbb{E}[Y(1) - Y(0)]$)
  • Statistical estimand (e.g. $\mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$)
Estimands, estimates, and the Identification-Estimation Flowchart

• Estimand - any quantity we want to estimate
  • Causal estimand (e.g. $\mathbb{E}[Y(1) - Y(0)]$)
  • Statistical estimand (e.g. $\mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$)
• Estimate: approximation of some estimand, using data
Estimands, estimates, and the Identification-Estimation Flowchart

• Estimand - any quantity we want to estimate
  • Causal estimand (e.g. $\mathbb{E}[Y(1) - Y(0)]$)
  • Statistical estimand (e.g. $\mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$)
• Estimate: approximation of some estimand, using data
• Estimation: process for getting from data + estimand to estimate
Estimands, estimates, and the Identification-Estimation Flowchart

• Estimand - any quantity we want to estimate
  • Causal estimand (e.g. $\mathbb{E}[Y(1) - Y(0)]$)
  • Statistical estimand (e.g. $\mathbb{E}_x [\mathbb{E}[Y | T = 1, X] - \mathbb{E}[Y | T = 0, X]]$)

• Estimate: approximation of some estimand, using data

• Estimation: process for getting from data + estimand to estimate

The Identification-Estimation Flowchart

Causal Estimand → Identification → Statistical Estimand → Estimation → Estimate
Problem: effect of sodium intake on blood pressure
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
• Epidemiological example taken from Luque-Fernandez et al. (2018)
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
• Epidemiological example taken from Luque-Fernandez et al. (2018)
• Outcome Y: (systolic) blood pressure (continuous)
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
• Epidemiological example taken from Luque-Fernandez et al. (2018)
• Outcome Y: (systolic) blood pressure (continuous)
• Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:
- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)
- Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
- Covariates X: age and amount of protein excreted in urine
Problem: effect of sodium intake on blood pressure

Motivation: 46% of Americans have high blood pressure and high blood pressure is associated with increased risk of mortality

Data:

- Epidemiological example taken from Luque-Fernandez et al. (2018)
- Outcome Y: (systolic) blood pressure (continuous)
- Treatment T: sodium intake (1 if above 3.5 mg and 0 if below)
- Covariates X: age and amount of protein excreted in urine
- Simulation: so we know the “true” ATE is 1.05
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y | T = 1, X] - \mathbb{E}[Y | T = 0, X]]$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation:
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x [\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]]$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x [\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]]$

Model (linear regression)
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y | T = 1, X] - \mathbb{E}[Y | T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x \left[ \mathbb{E}[Y | T = 1, x] - \mathbb{E}[Y | T = 0, x] \right]$

Model (linear regression)

Estimate: 0.85
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x [\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]]$

Model (linear regression)

Estimate: 0.85

Naive: $\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x [\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]]$

Model (linear regression)

Estimate: 0.85

Naive: $\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]$

Naive estimate: 5.33
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x [\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]]$

Model (linear regression)

Estimate: 0.85

Naive: $\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]$

Naive estimate: 5.33

$\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$
Estimation of ATE

True ATE: $\mathbb{E}[Y(1) - Y(0)] = 1.05$

Identification: $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X [\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]]$

Estimation: $\frac{1}{n} \sum_x [\mathbb{E}[Y \mid T = 1, x] - \mathbb{E}[Y \mid T = 0, x]]$

Model (linear regression)

Estimate: 0.85

$\frac{|0.85 - 1.05|}{1.05} \times 100\% = 19\%$

Naive: $\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]$

Naive estimate: 5.33

$\frac{|5.33 - 1.05|}{1.05} \times 100\% = 407\%$
Using coefficient of linear regression
Using coefficient of linear regression

Assume linear parametric form: $Y = \alpha T + \beta X$
Using coefficient of linear regression

Assume linear parametric form: $Y = \alpha T + \beta X$

Run linear regression: $Y = \hat{\alpha} T + \hat{\beta} X$
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \) \( \hat{\alpha} = 0.85 \)
Using coefficient of linear regression

Assume linear parametric form: $Y = \alpha T + \beta X$

Run linear regression: $Y = \hat{\alpha} T + \hat{\beta} X \quad \hat{\alpha} = 0.85$

Continuous treatment: $\mathbb{E}[Y(1) - Y(0)]$
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \qquad \hat{\alpha} = 0.85 \)

Continuous treatment: \( \mathbb{E}[Y(t)] \)
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \) \( \hat{\alpha} = 0.85 \)

Continuous treatment: \( \mathbb{E}[Y(t)] \) \( \hat{\alpha} = 0.85 \)
Using coefficient of linear regression

Assume linear parametric form: $Y = \alpha T + \beta X$

Run linear regression: $Y = \hat{\alpha} T + \hat{\beta} X$ \hspace{1cm} \hat{\alpha} = 0.85

Continuous treatment: $E[Y(t)]$ \hspace{1cm} \hat{\alpha} = 0.85

Severe limitations:
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \)  \( \hat{\alpha} = 0.85 \)

Continuous treatment: \( \mathbb{E}[Y(t)] \)  \( \hat{\alpha} = 0.85 \)

Severe limitations: the causal effect is the same for all individuals
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \) \( \hat{\alpha} = 0.85 \)

Continuous treatment: \( \mathbb{E}[Y(t)] \) \( \hat{\alpha} = 0.85 \)

Severe limitations: the causal effect is the same for all individuals

\[ Y_i(t) = \alpha t + \beta x_i \]
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \) \( \hat{\alpha} = 0.85 \)

Continuous treatment: \( E[Y(t)] \) \( \hat{\alpha} = 0.85 \)

Severe limitations: the causal effect is the same for all individuals

\[
Y_i(t) = \alpha t + \beta x_i \quad Y_i(1) - Y_i(0) = \alpha \cdot 1 + \beta x_i \\
-\alpha \cdot 0 - \beta x_i
\]
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \) \( \hat{\alpha} = 0.85 \)

Continuous treatment: \( \mathbb{E}[Y(t)] \) \( \hat{\alpha} = 0.85 \)

Severe limitations: the causal effect is the same for all individuals

\[
Y_i(t) = \alpha t + \beta x_i \quad Y_i(1) - Y_i(0) = \alpha \cdot 1 + \beta x_i \\
-\alpha \cdot 0 - \beta x_i = \alpha
\]
Using coefficient of linear regression

Assume linear parametric form: \( Y = \alpha T + \beta X \)

Run linear regression: \( Y = \hat{\alpha} T + \hat{\beta} X \) \( \hat{\alpha} = 0.85 \)

Continuous treatment: \( \mathbb{E}[Y(t)] \) \( \hat{\alpha} = 0.85 \)

Severe limitations: the causal effect is the same for all individuals

\[
Y_i(t) = \alpha t + \beta x_i \quad Y_i(1) - Y_i(0) = \alpha \cdot 1 + \beta x_i \\
-\alpha \cdot 0 - \beta x_i = \alpha
\]

See Sections 6.2 and 6.3 of Morgan & Winship (2014) for more complete critique