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

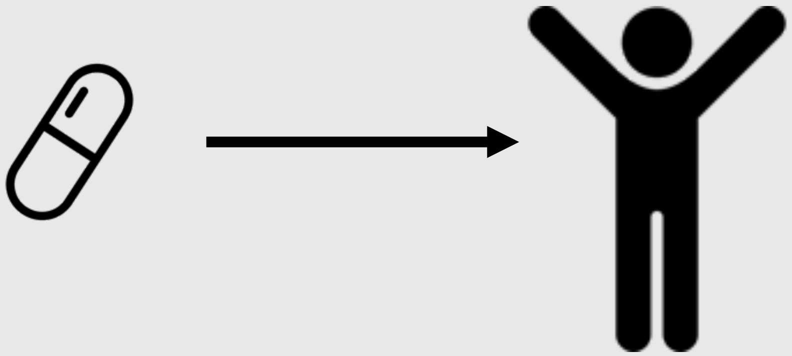
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

Potential outcomes: intuition

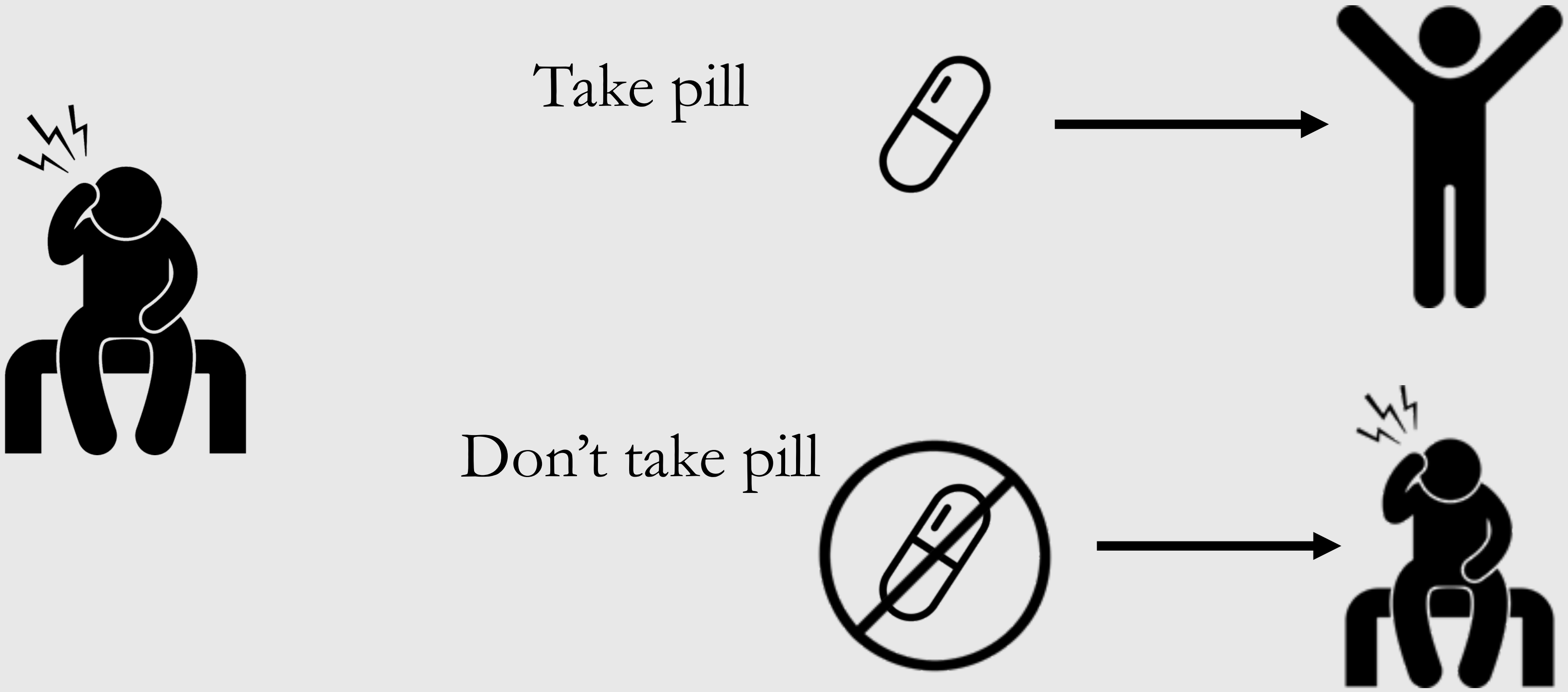
Inferring the effect of treatment/policy on some outcome

Take pill



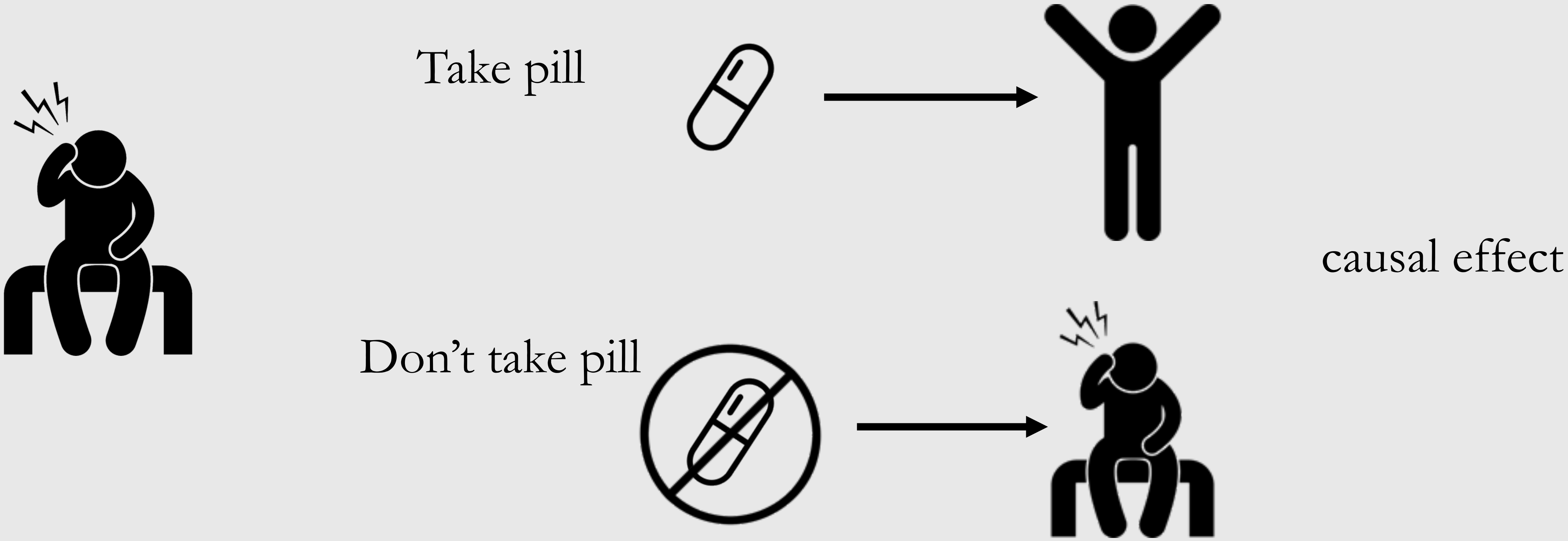
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome



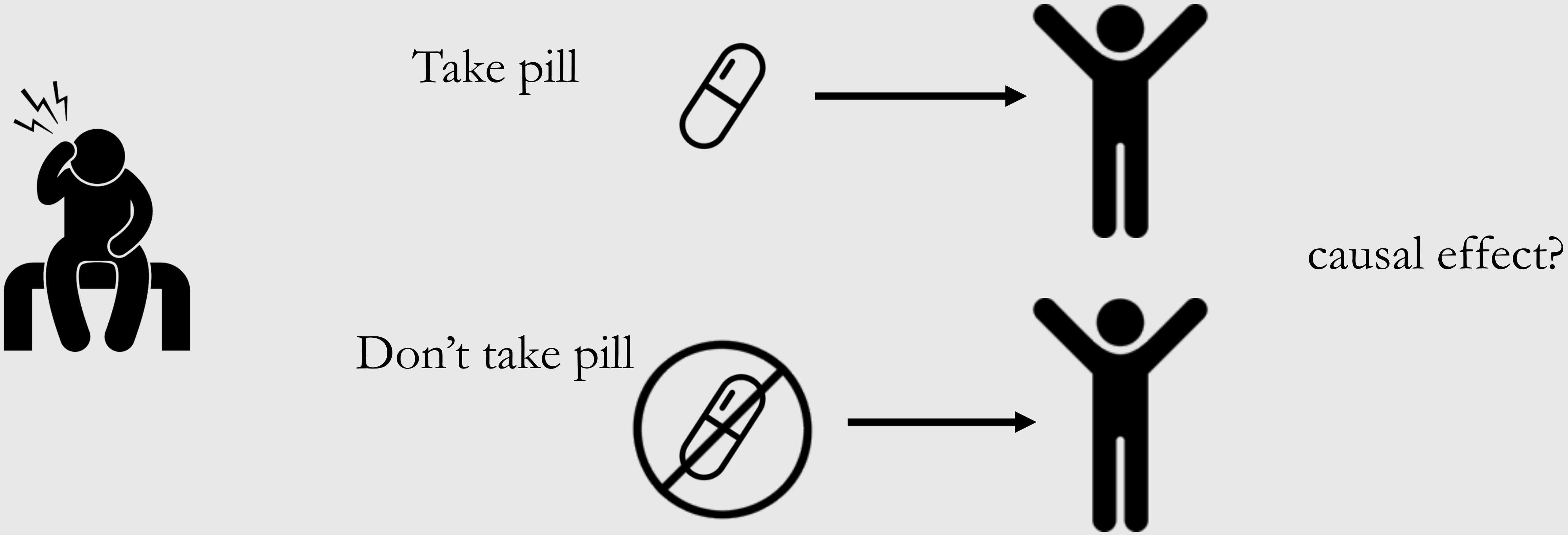
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome



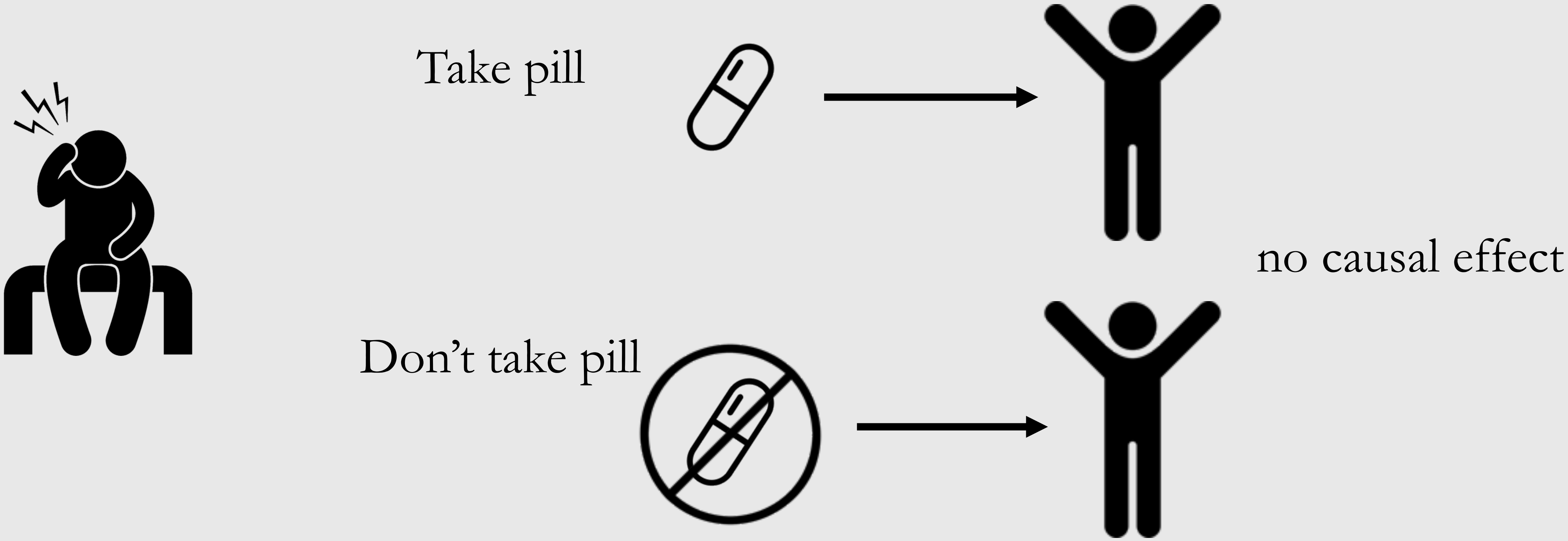
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome



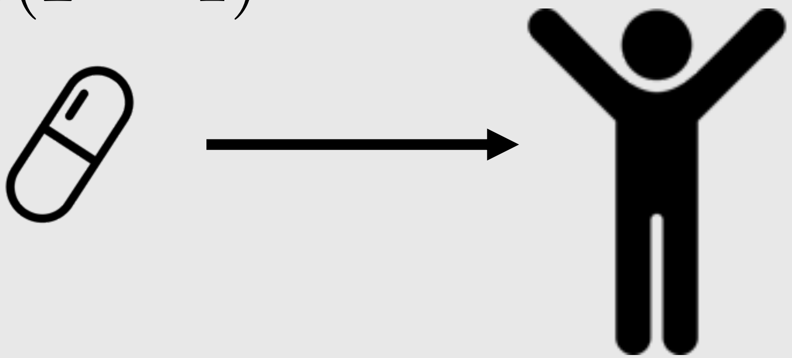
Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

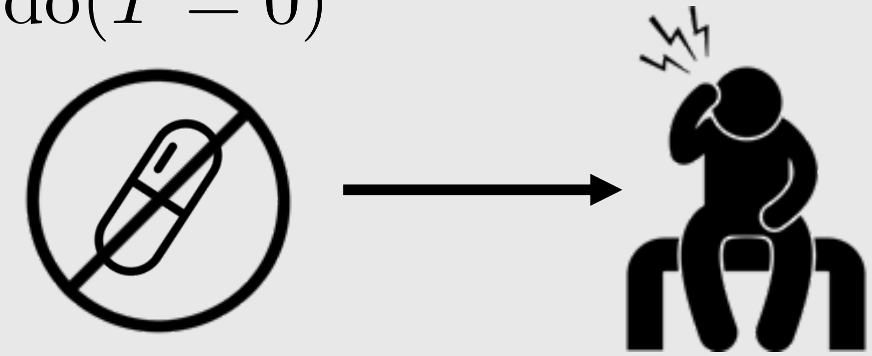


Potential outcomes: notation

$\mathrm{do}(T = 1)$



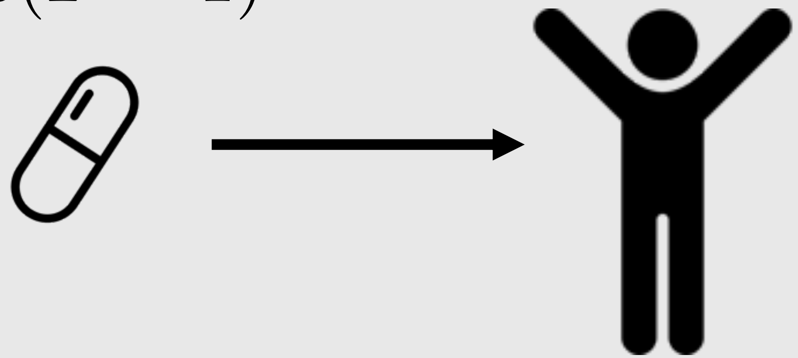
$\mathrm{do}(T = 0)$



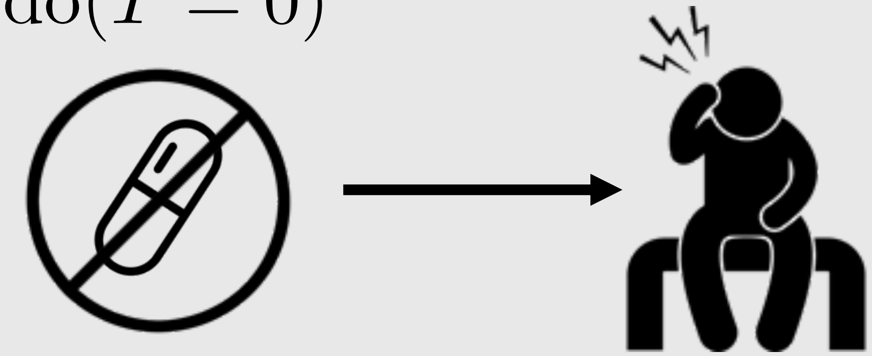
T : observed treatment
Y : observed outcome

Potential outcomes: notation

$\mathrm{do}(T = 1)$ $\qquad$ $Y_i|_{\mathrm{do}(T=1)}$

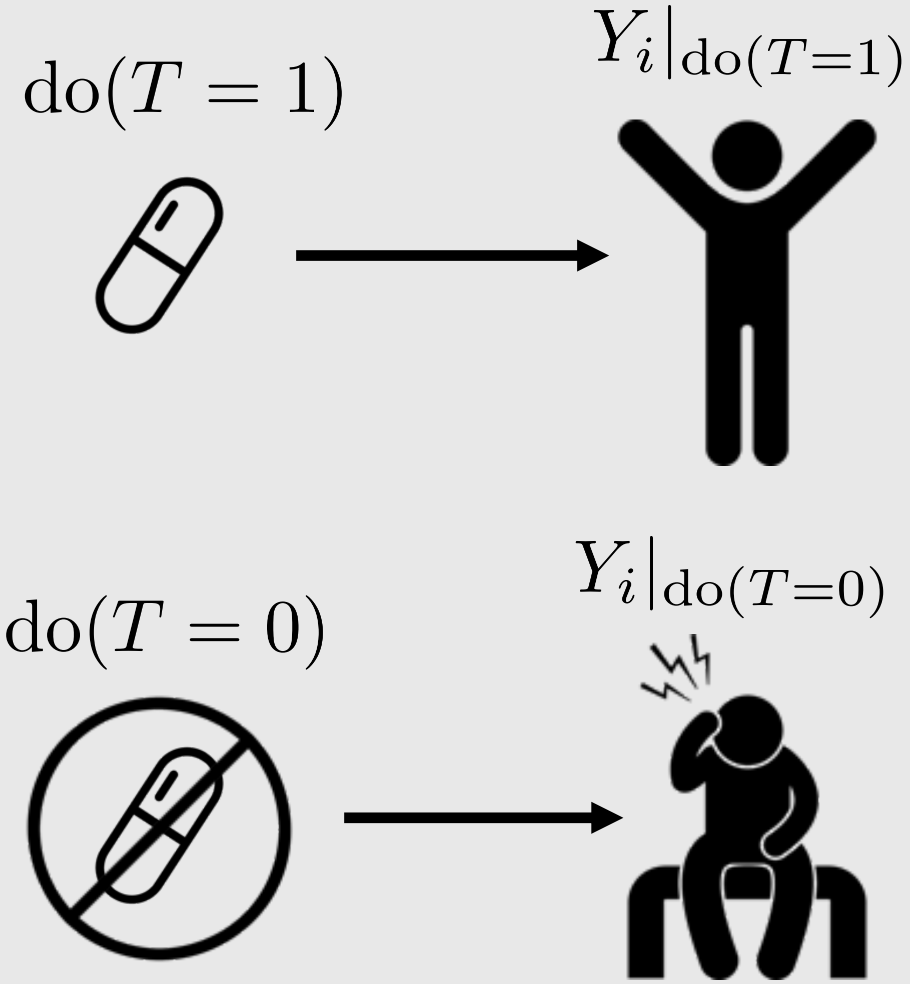


$\mathrm{do}(T = 0)$



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

Potential outcomes: notation



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

Potential outcomes: notation

$\mathrm{do}(T = 1)$

$$Y_i|_{\mathrm{do}(T=1)} \triangleq Y_i(1)$$

$\mathrm{do}(T = 0)$

$$Y_i|_{\mathrm{do}(T=0)}$$

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

$$Y_i|_{do(T=1)} \triangleq Y_i(\underline{1})$$

$\underline{do(T = 1)}$

$$Y_i|_{do(T=0)}$$

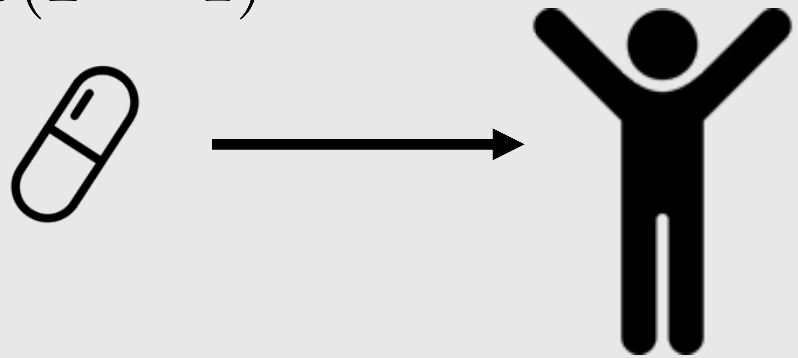
$do(T = 0)$

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

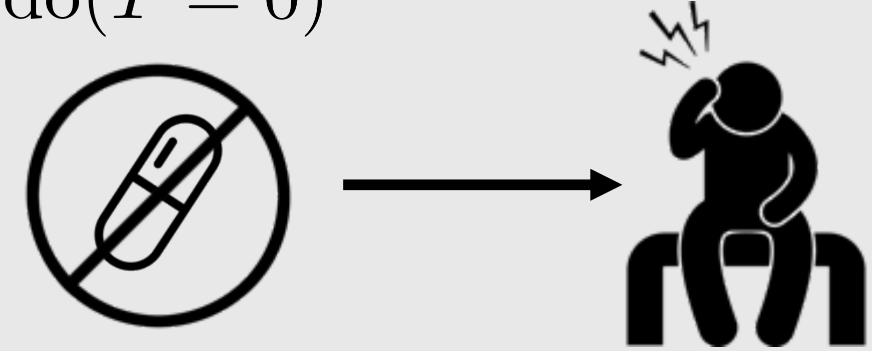
$\mathrm{do}(T = 1)$

$Y_i|_{\mathrm{do}(T=1)} \triangleq Y_i(1)$



$\mathrm{do}(T = 0)$

$Y_i|_{\mathrm{do}(T=0)} \triangleq Y_i(0)$

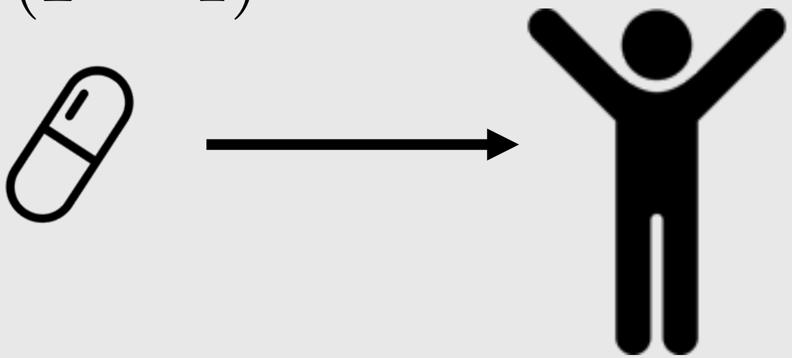


T	: observed treatment
Y	: observed outcome
i	: used in subscript to denote a specific unit/individual
$Y_i(1)$	: potential outcome under treatment
$Y_i(0)$	: potential outcome under no treatment

Potential outcomes: notation

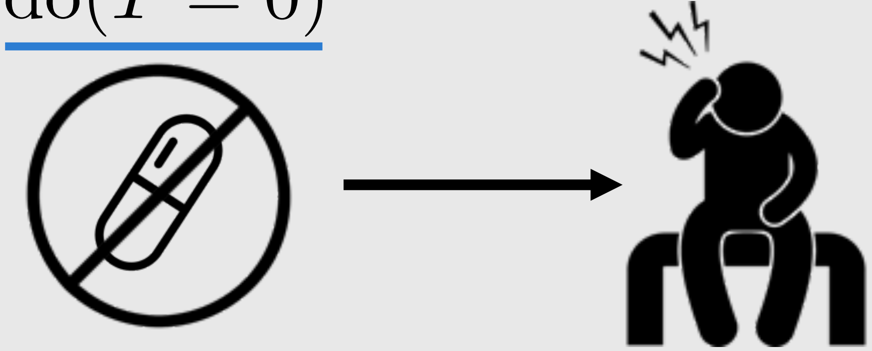
$\mathrm{do}(T = 1)$

$Y_i|_{\mathrm{do}(T=1)} \triangleq Y_i(1)$



$\mathrm{do}(T = 0)$

$Y_i|_{\mathrm{do}(T=0)} \triangleq Y_i(0)$

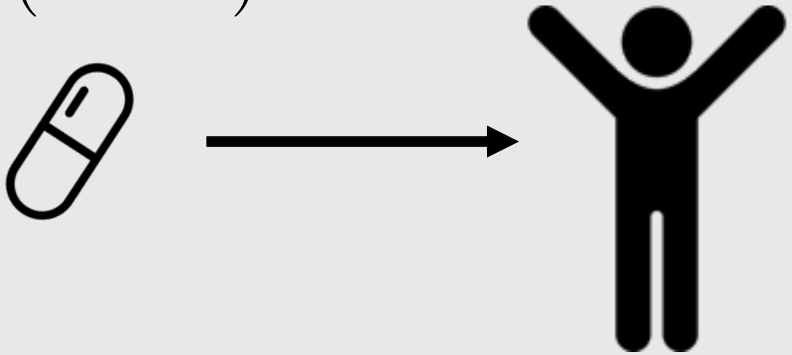


T	: observed treatment
Y	: observed outcome
i	: used in subscript to denote a specific unit/individual
$Y_i(1)$	: potential outcome under treatment
$Y_i(0)$	: potential outcome under no treatment

Potential outcomes: notation

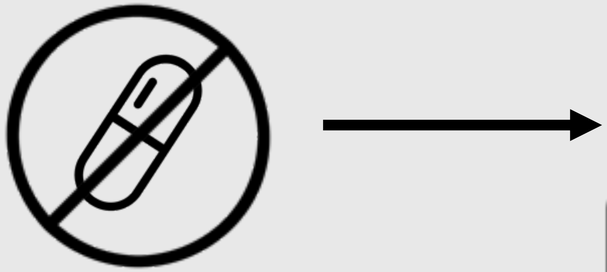
$\mathrm{do}(T = 1)$

$$Y_i|_{\mathrm{do}(T=1)} \triangleq Y_i(1)$$



$\mathrm{do}(T = 0)$

$$Y_i|_{\mathrm{do}(T=0)} \triangleq Y_i(0)$$



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

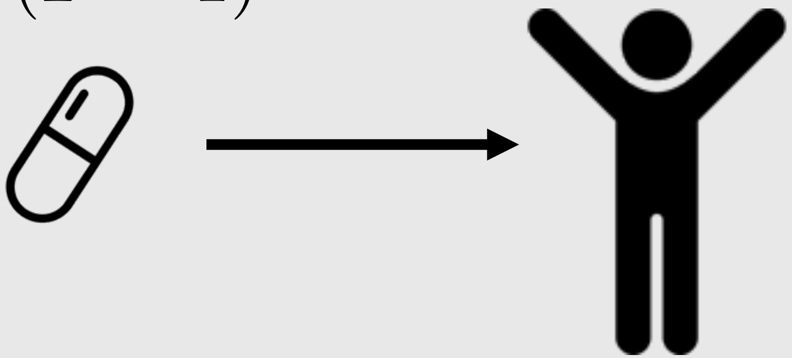
Causal effect

$$Y_i(1) - Y_i(0)$$

Potential outcomes: notation

$\text{do}(T = 1)$

$Y_i|_{\text{do}(T=1)} \triangleq Y_i(1)$



T : observed treatment

Y : observed outcome

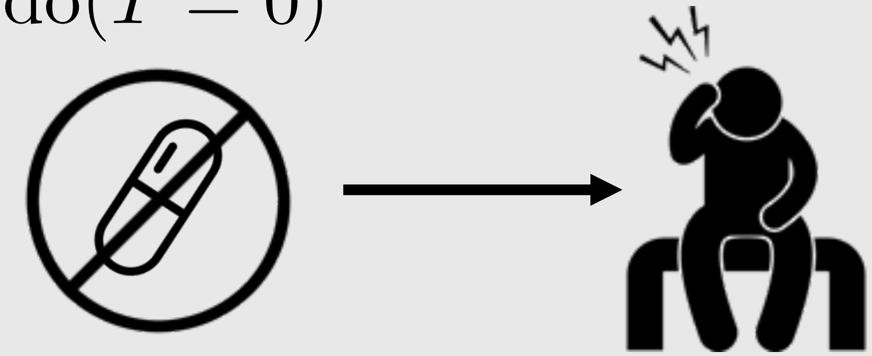
i : used in subscript to denote a specific unit/individual

$Y_i(1)$: potential outcome under treatment

$Y_i(0)$: potential outcome under no treatment

$Y_i(0) = 0$

$\text{do}(T = 0)$

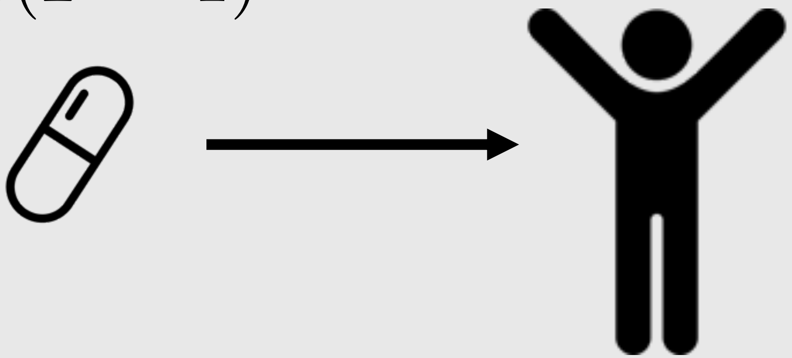


Causal effect

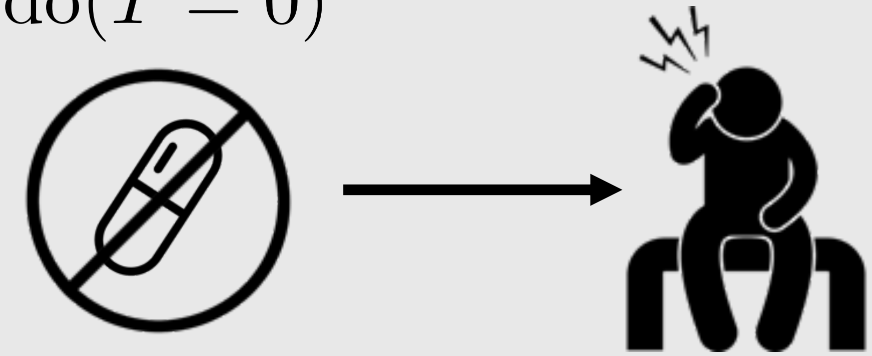
$Y_i(1) - Y_i(0)$

Potential outcomes: notation

$\mathrm{do}(T = 1)$ $\quad Y_i(1) = 1$



$\mathrm{do}(T = 0)$ $\quad Y_i(0) = 0$



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

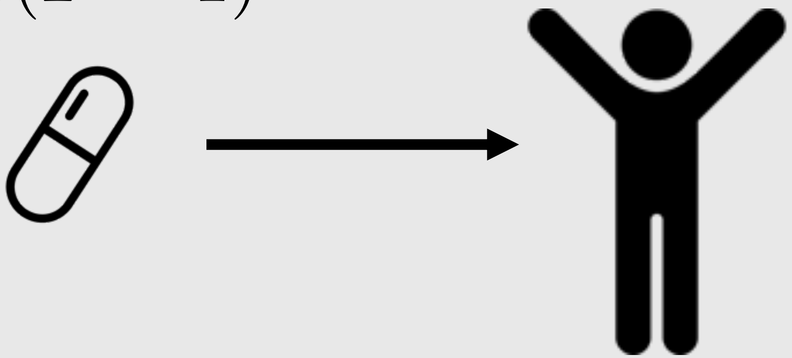
Causal effect

$$Y_i(1) - Y_i(0)$$

Potential outcomes: notation

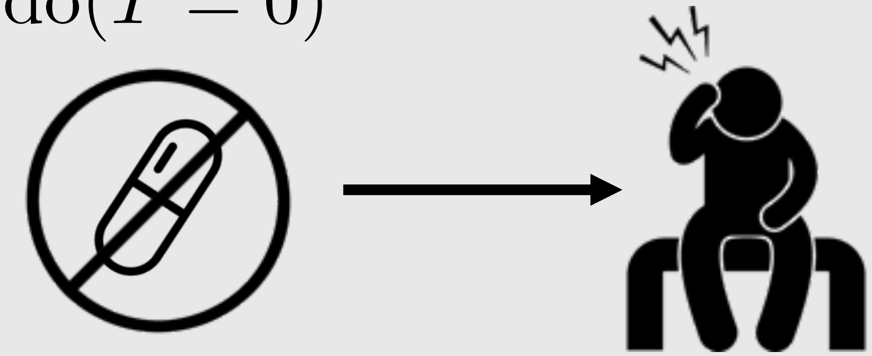
$Y_i(1) = 1$

$\text{do}(T = 1)$



$Y_i(0) = 0$

$\text{do}(T = 0)$



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

$$Y_i(1) - Y_i(0) = 1$$

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

$\mathrm{do}(T = 1)$

$Y_i(1) = 1$

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

$\mathrm{do}(T = 0)$

$Y_i(0) = 0$

Causal effect

$Y_i(1) - Y_i(0) = 1$

Fundamental problem of causal inference

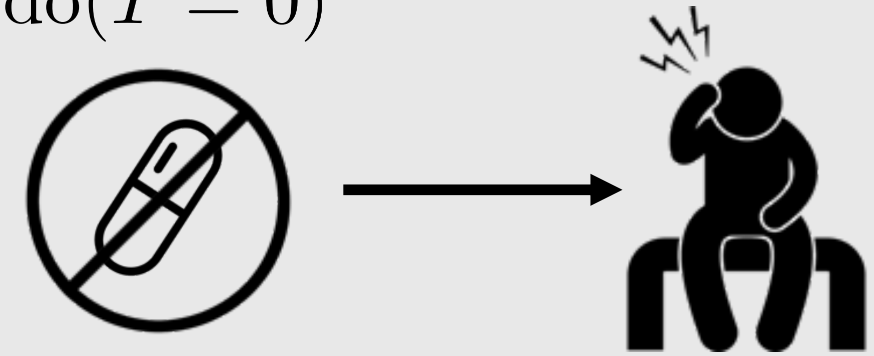
Counterfactual

$\mathrm{do}(T=1)$ $\qquad Y_i(1) = 1$

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

Factual

$\mathrm{do}(T=0)$ $\qquad Y_i(0) = 0$



Causal effect

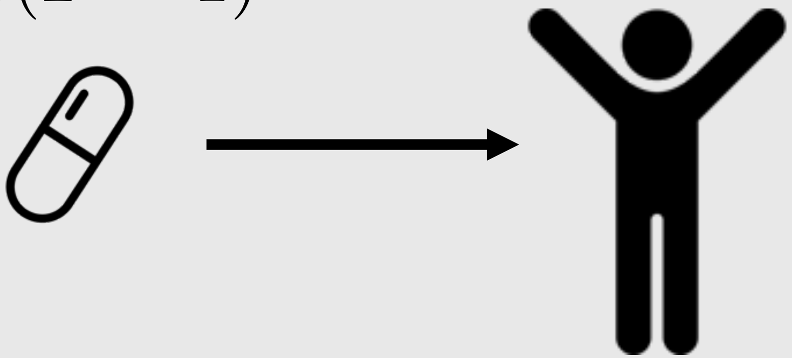
$$Y_i(1) - Y_i(0) = 1$$

Fundamental problem of causal inference

Factual

$\mathrm{do}(T = 1)$

$Y_i(1) = 1$



Counterfactual

$\mathrm{do}(T = 0)$

$Y_i(0) = 0$

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

$Y_i(1) - Y_i(0) = 1$

Missing data interpretation

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	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

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)

$$Y_i(1) - Y_i(0)$$

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	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

Average treatment effect (ATE)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	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

Average treatment effect (ATE)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	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

Average treatment effect (ATE)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	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

Average treatment effect (ATE)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	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

Average treatment effect (ATE)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	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

Average treatment effect (ATE)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?
			$\underline{2/3}$		

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)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?
			$2/3$	$1/3$	

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)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$$\underline{2/3} - \underline{1/3} = 1/3$$

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)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$$^2/_3 \quad - \quad ^1/_3 \quad = \quad ^1/_3$$

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)

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$$^2/_3 \quad - \quad ^1/_3 \quad = \quad ^1/_3$$

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)

$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \neq \underbrace{\mathbb{E}[Y \mid T = 1] - \mathbb{E}[Y \mid T = 0]}_{\text{associational difference}}$$

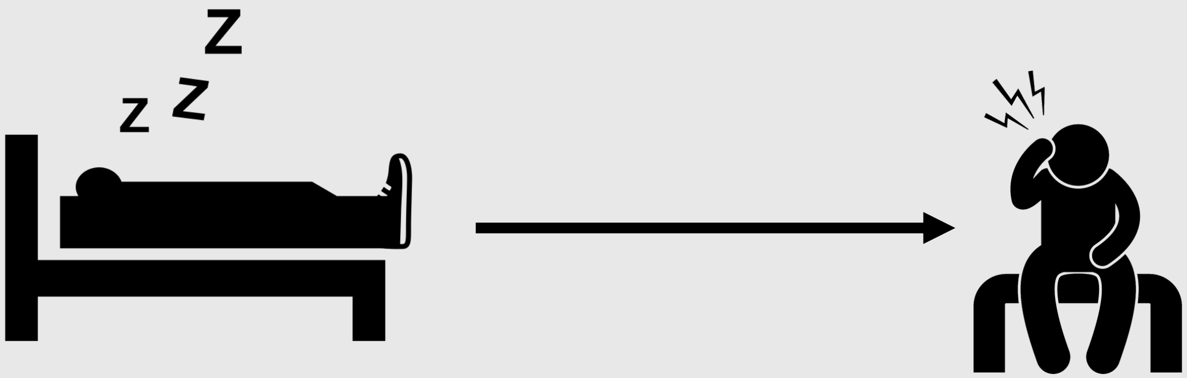
i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$$2/3 \quad - \quad 1/3 \quad = \quad 1/3$$

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

Association is not causation

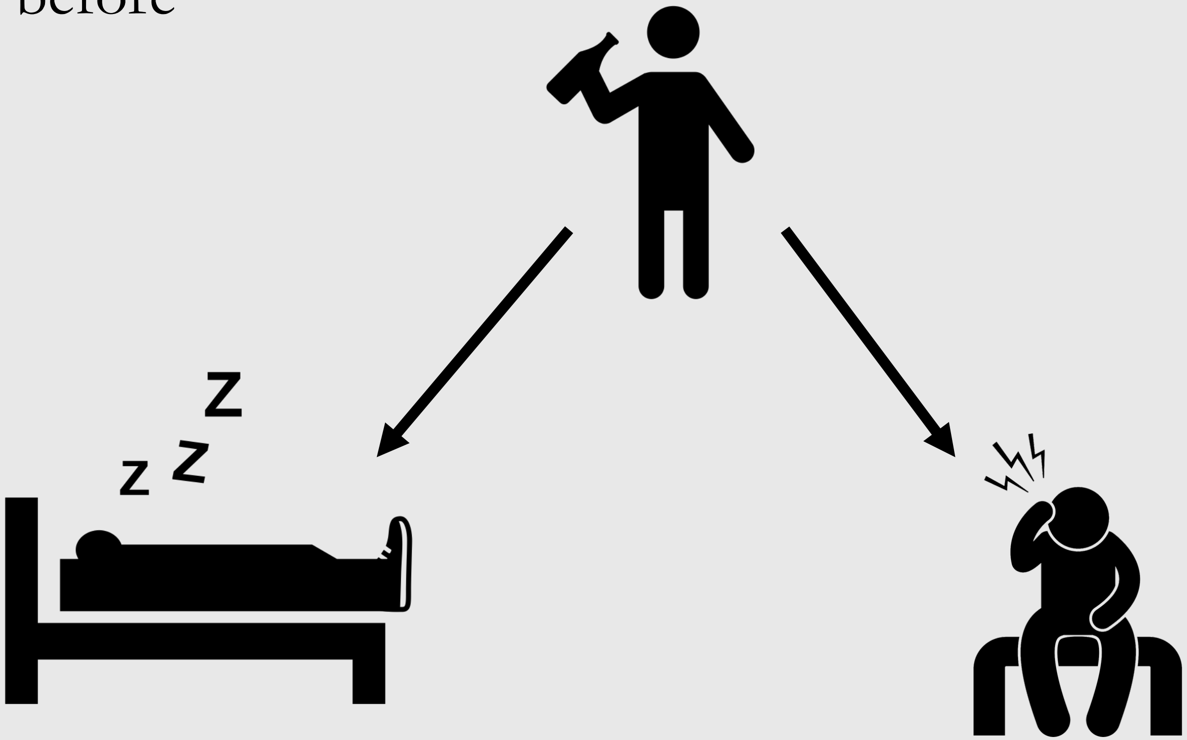
Sleeping with shoes on is strongly correlated with waking up with a headache



Association is not causation

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

Common cause: drinking the night before

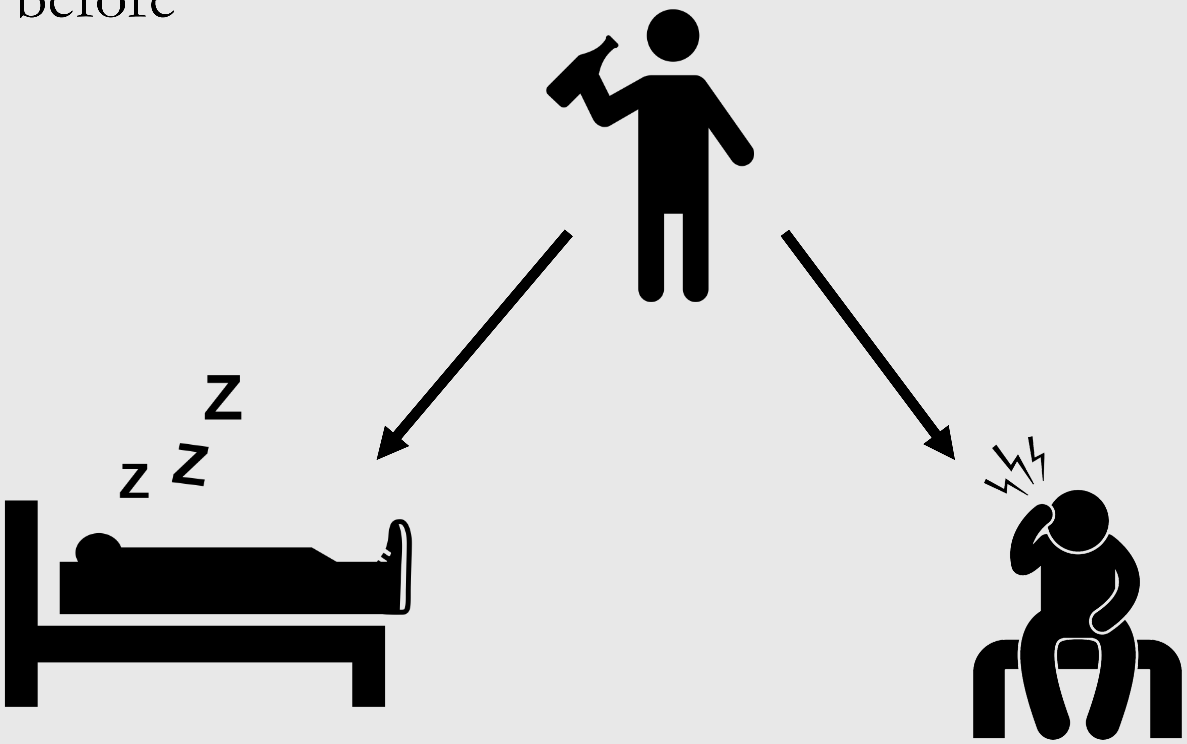


Association is not causation

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

Common cause: drinking the night before

1. Confounding

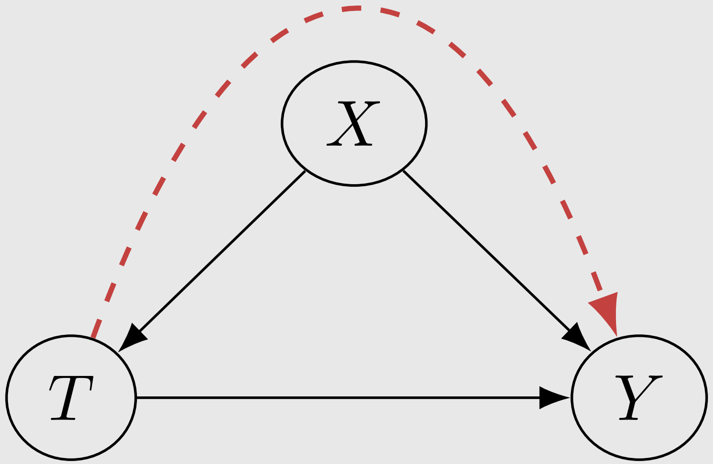


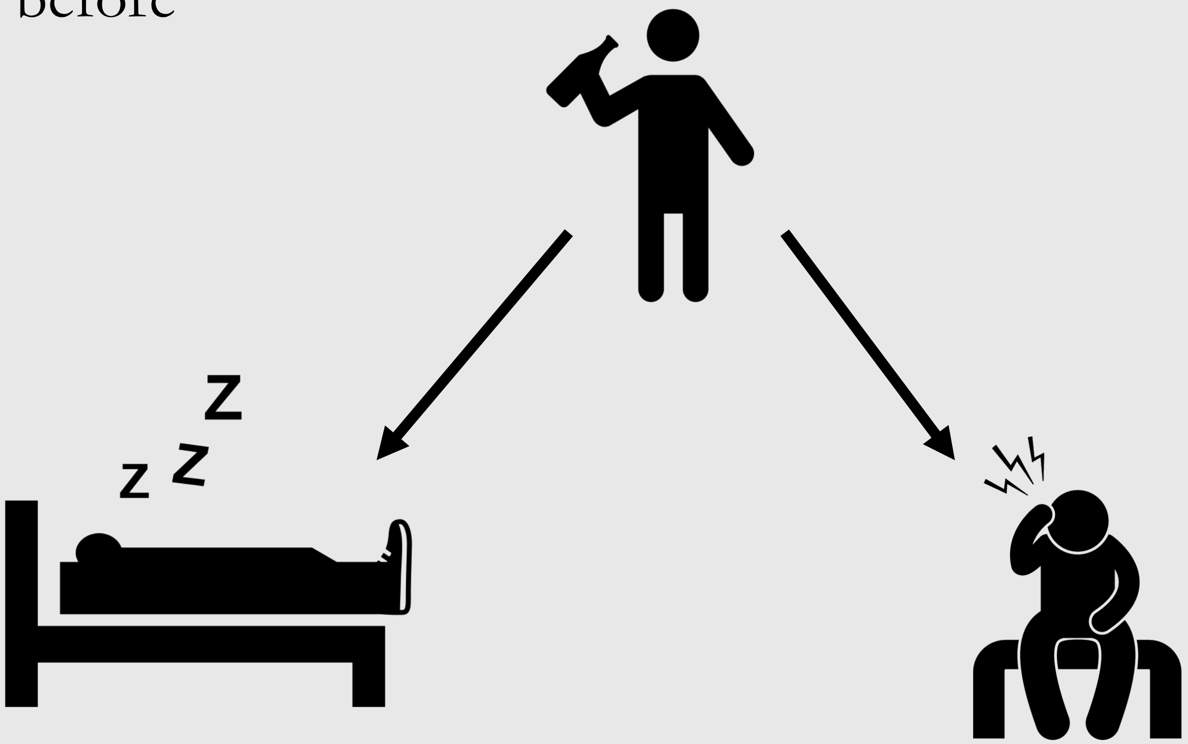
Association is not causation

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

Common cause: drinking the night before

1. Confounding



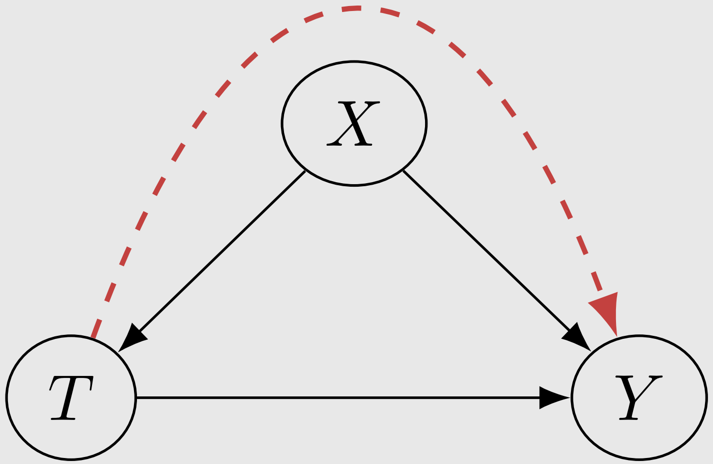


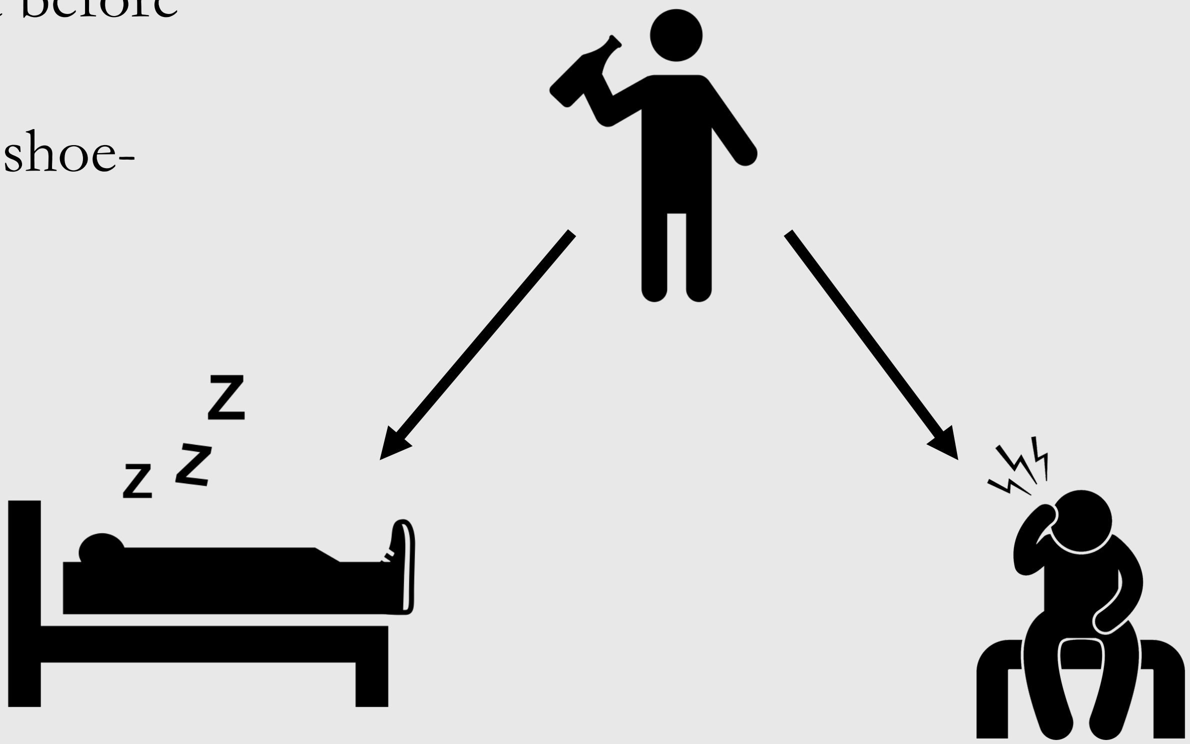
Association is not causation

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



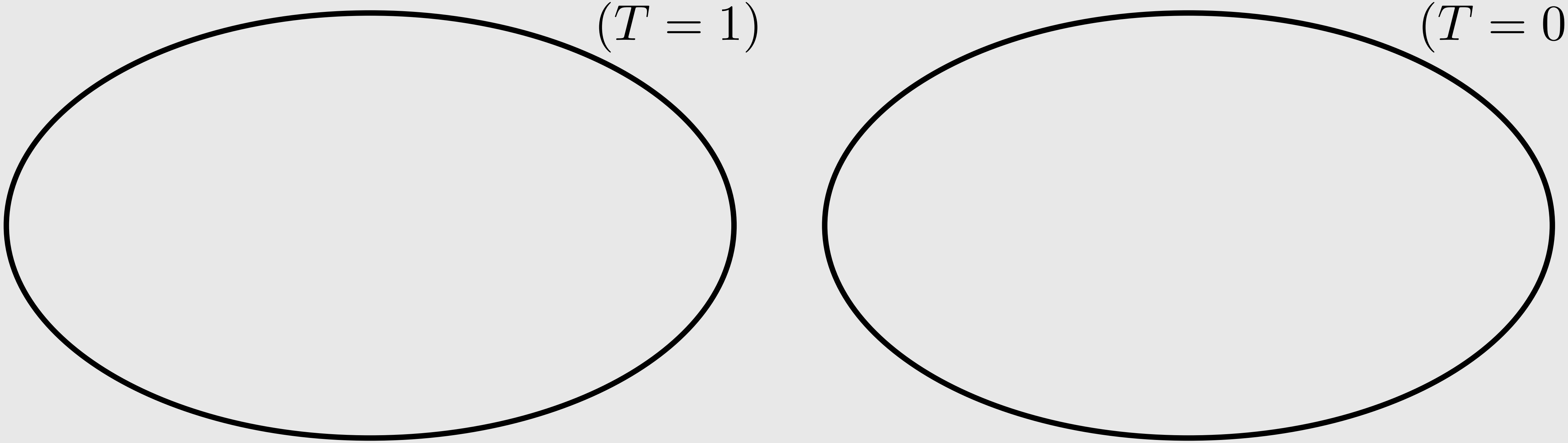


Why? Because the groups are not comparable

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

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

Went to sleep **without shoes on** $(T = 0)$

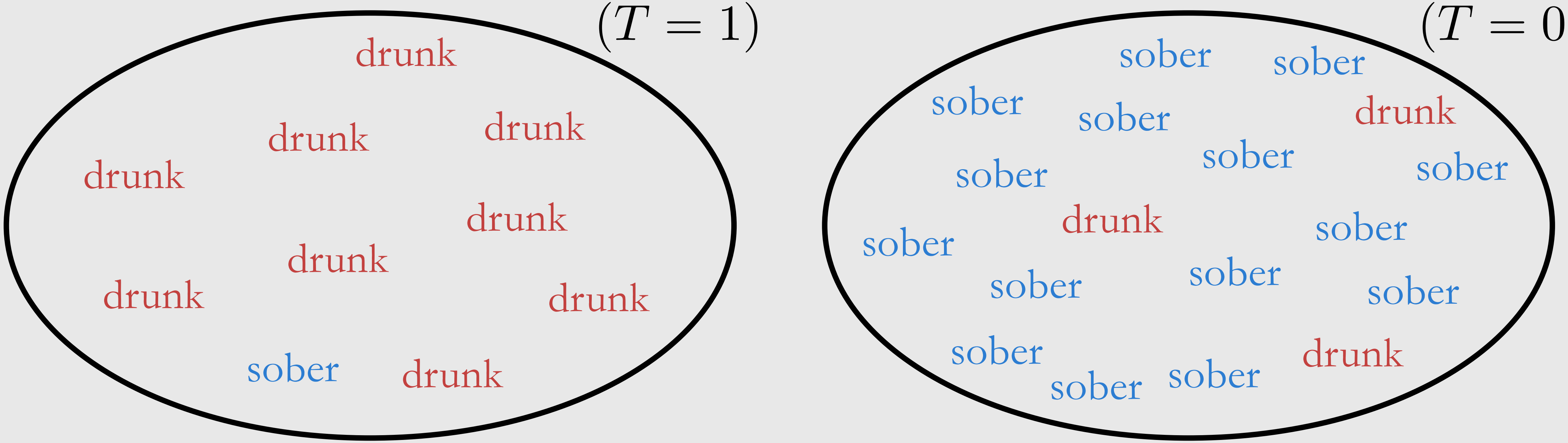


Why? Because the groups are not comparable

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

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

Went to sleep **without shoes on** $(T = 0)$

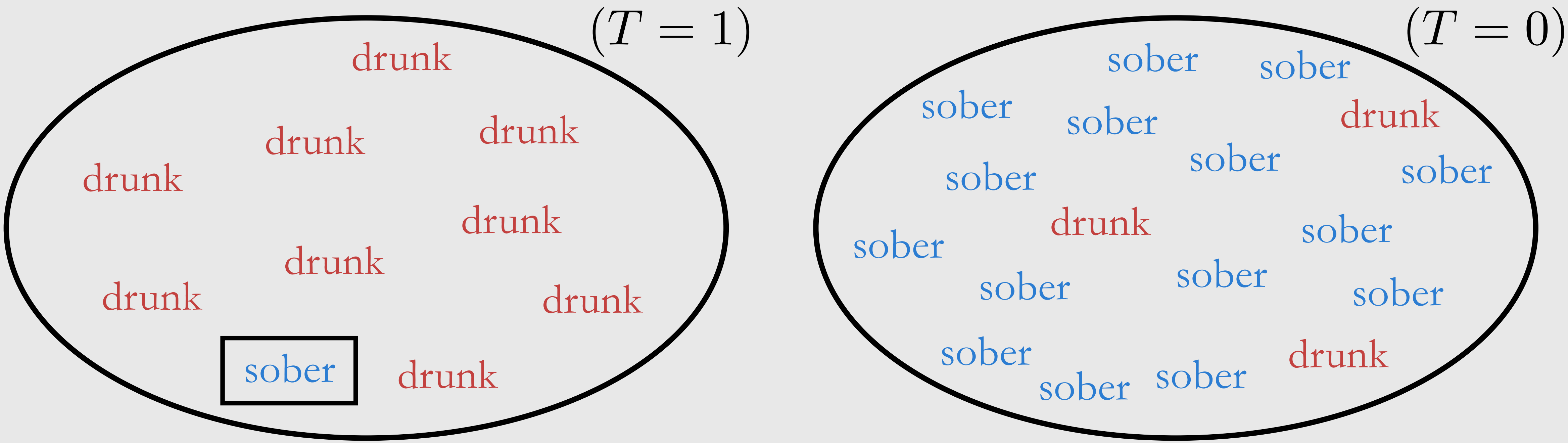


Why? Because the groups are not comparable

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

Went to sleep **with shoes** on

Went to sleep **without shoes** on

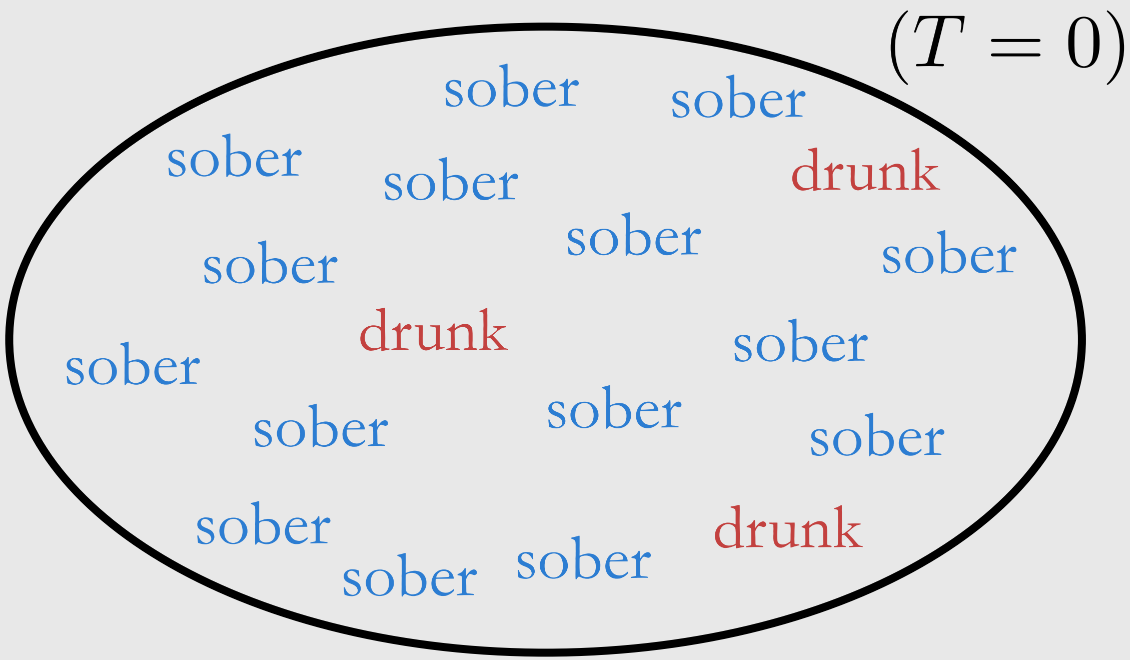


Why? Because the groups are not comparable

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

Went to sleep **with shoes** on

Went to sleep **without shoes** on

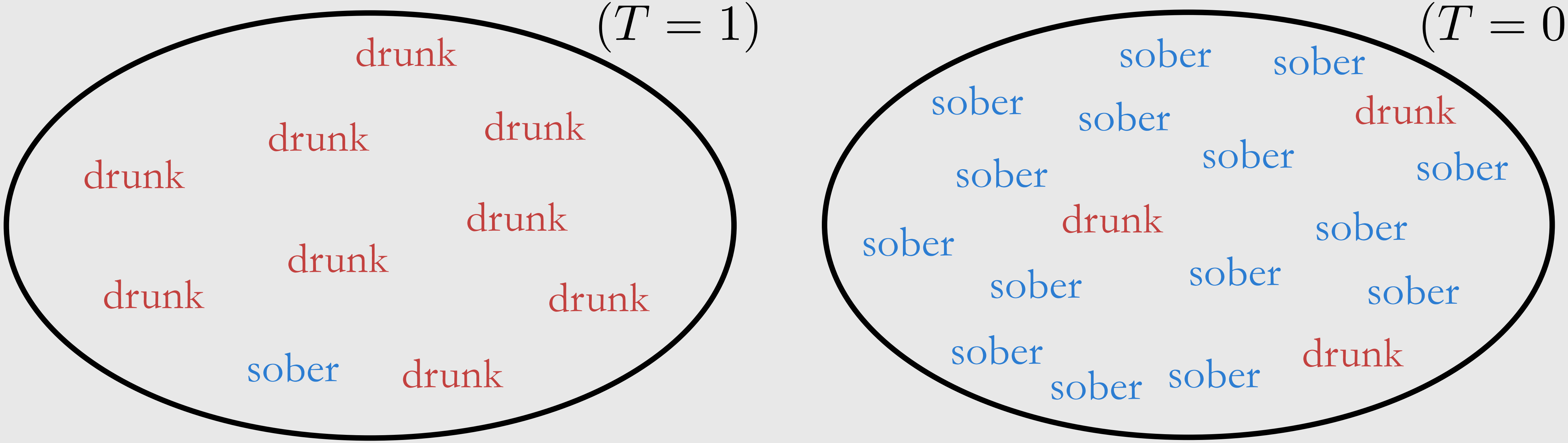


What would comparable groups look like?

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

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

Went to sleep **without shoes on** $(T = 0)$

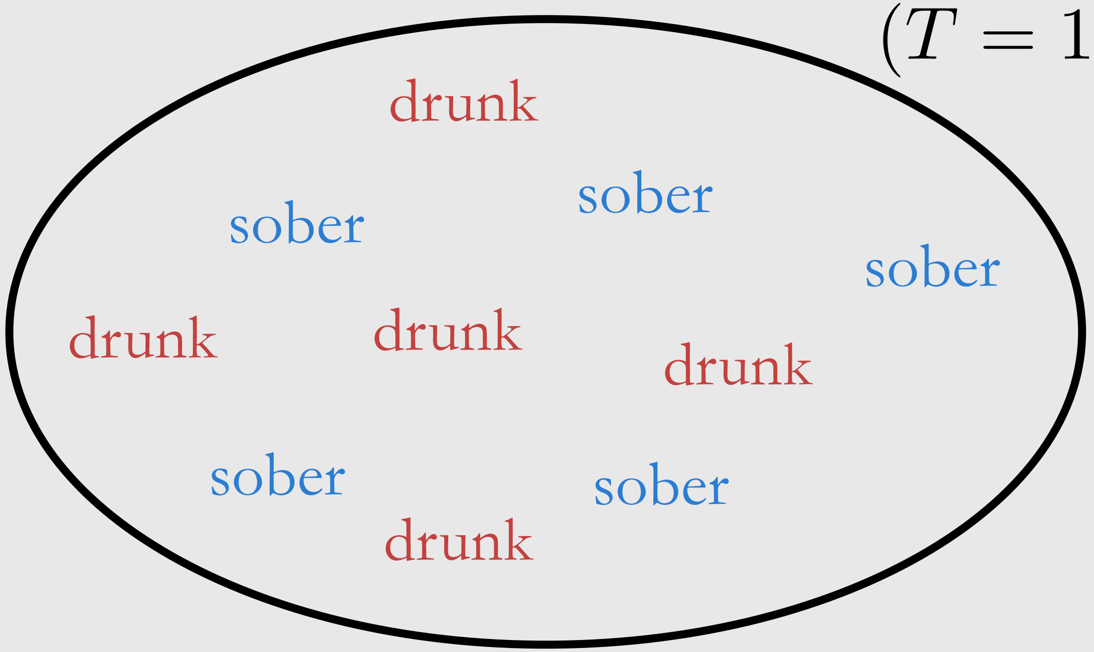


What would comparable groups look like?

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

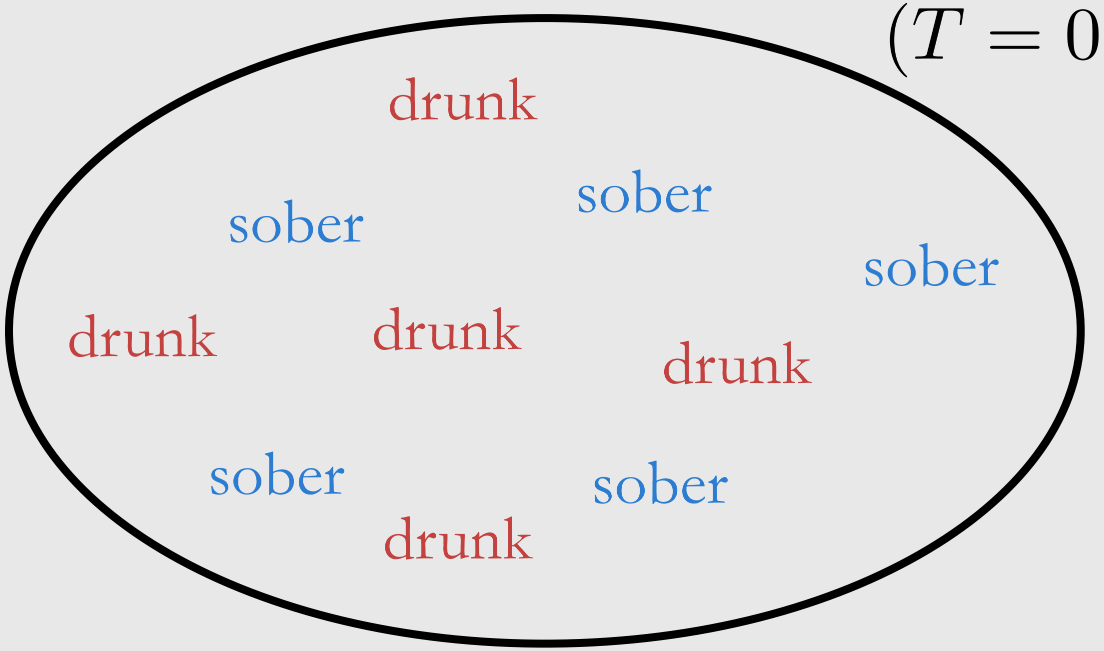
Went to sleep **with shoes on**

$(T = 1)$



Went to sleep **without shoes on**

$(T = 0)$

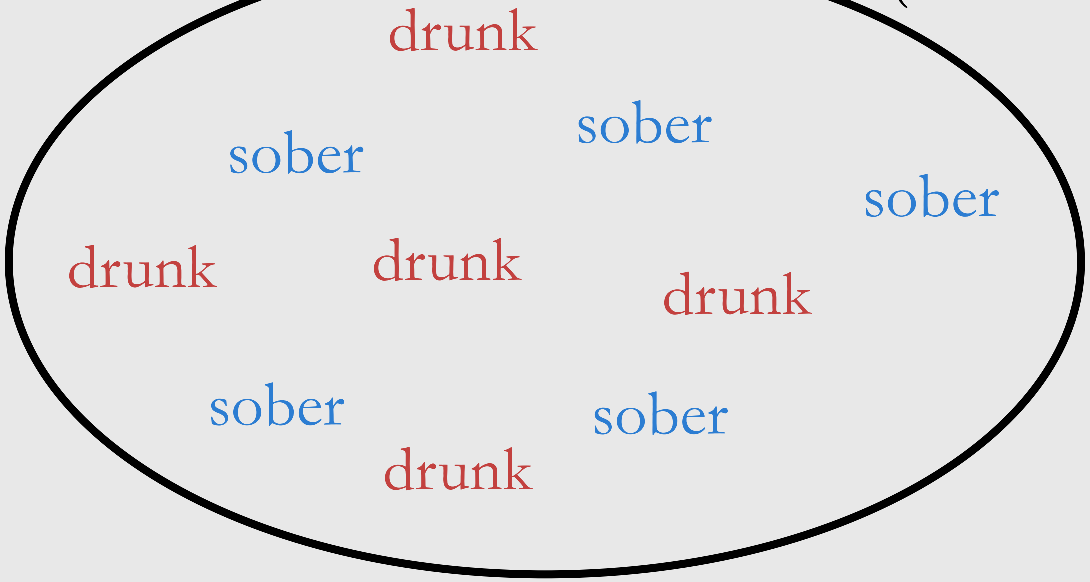


What would comparable groups look like?

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

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

Went to sleep **without shoes on** $(T = 0)$



Question:
Why is association not causation?

What assumptions would make the ATE equal to the associational difference?

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0	?	0	?
2	1	1	1	?	?
3	1	0	0	?	?
4	0	0	?	0	?
5	0	1	?	1	?
6	1	1	1	?	?

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?
			$^2/_3$		

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?
			$2/3$	$1/3$	

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$\underline{2/3} \quad - \quad \underline{1/3} \quad = \quad 1/3$

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

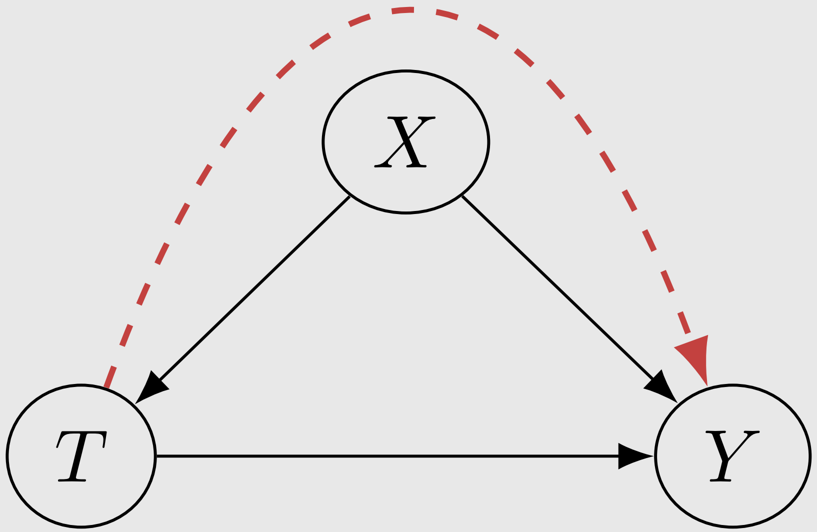
$$\mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] = \mathbb{E}[Y(1) \mid T=1] - \mathbb{E}[Y(0) \mid T=0] \quad \text{(ignorability)}$$
$$= \mathbb{E}[Y \mid T=1] - \mathbb{E}[Y \mid T=0]$$

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?
			$^2/_3$	$-\ ^1/_3$	$=\ ^1/_3$

Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?
			$2/3$ $-$	$1/3$ $=$	$1/3$

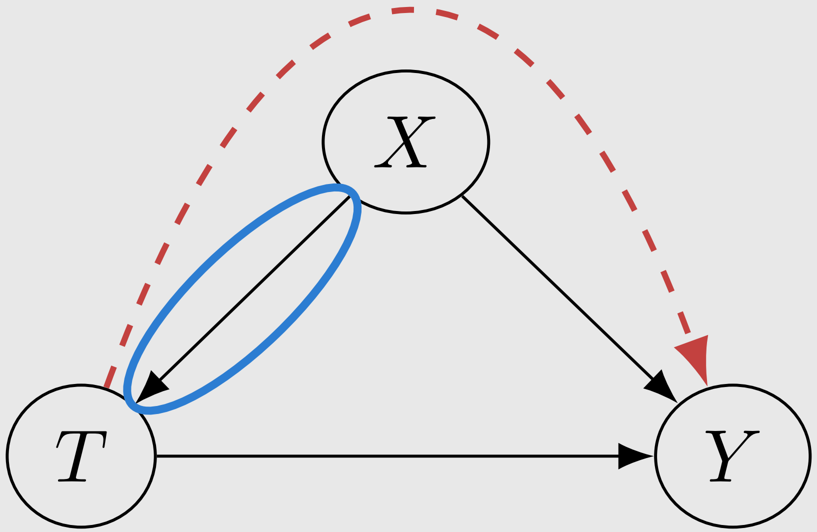


Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$^2/_3 \quad - \quad ^1/_3 \quad = \quad ^1/_3$

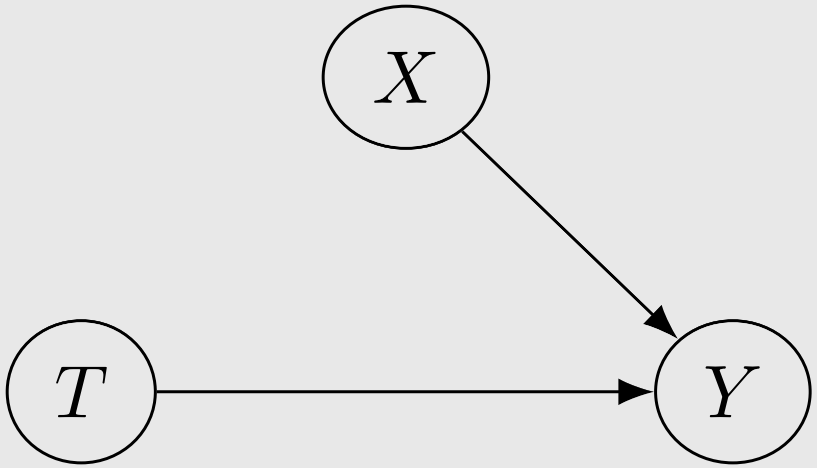


Ignorability: $(Y(1), Y(0)) \perp\!\!\!\perp T$

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

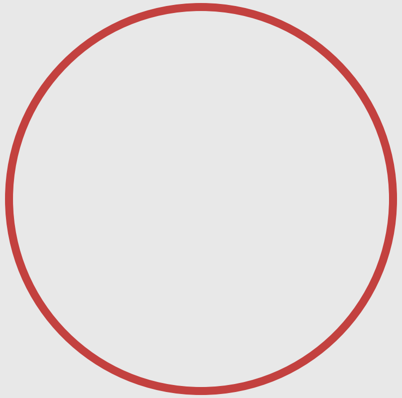
i	T	Y	$Y(1)$	$Y(0)$	$Y(1) - Y(0)$
1	0	0		0	?
2	1	1	1		?
3	1	0	0		?
4	0	0		0	?
5	0	1		1	?
6	1	1	1		?

$2/3 \quad - \quad 1/3 \quad = \quad 1/3$



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

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

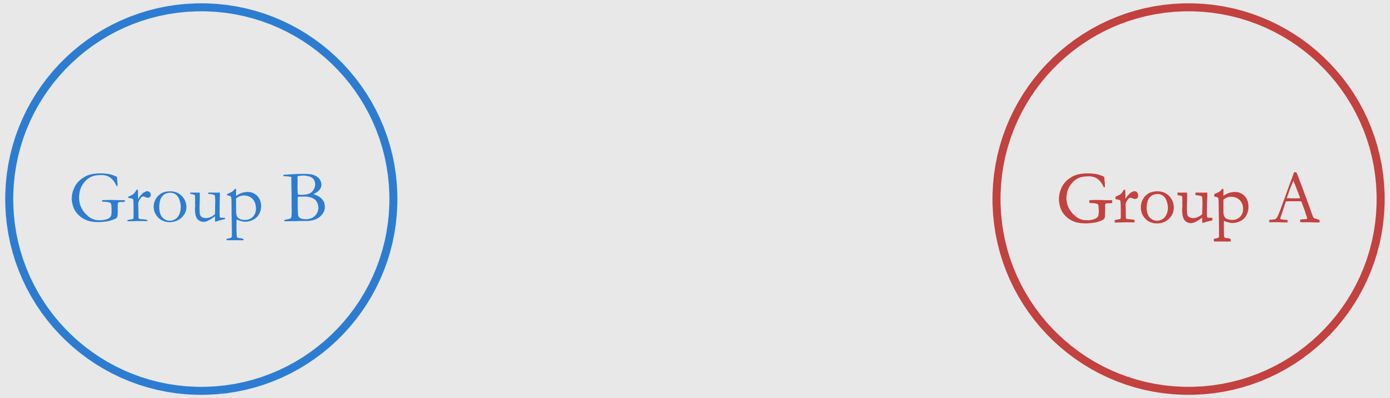
$T = 0$

Group B

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

Another perspective: exchangeability

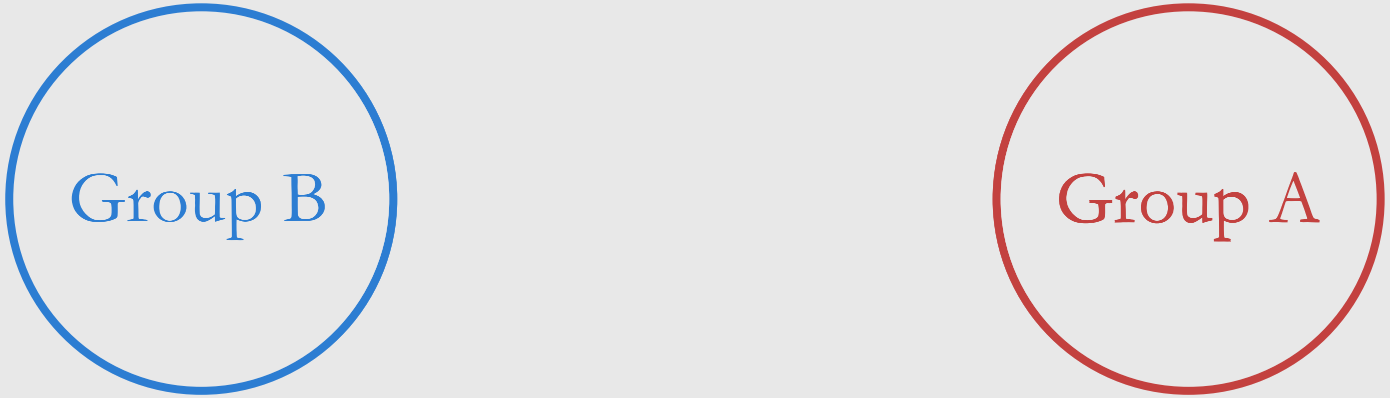
$T = 1$ $\qquad$ $T = 0$



$\mathbb{E}[Y \mid T = 1] = y_1$ $\qquad$ $\mathbb{E}[Y \mid T = 0] = y_0$

Another perspective: exchangeability

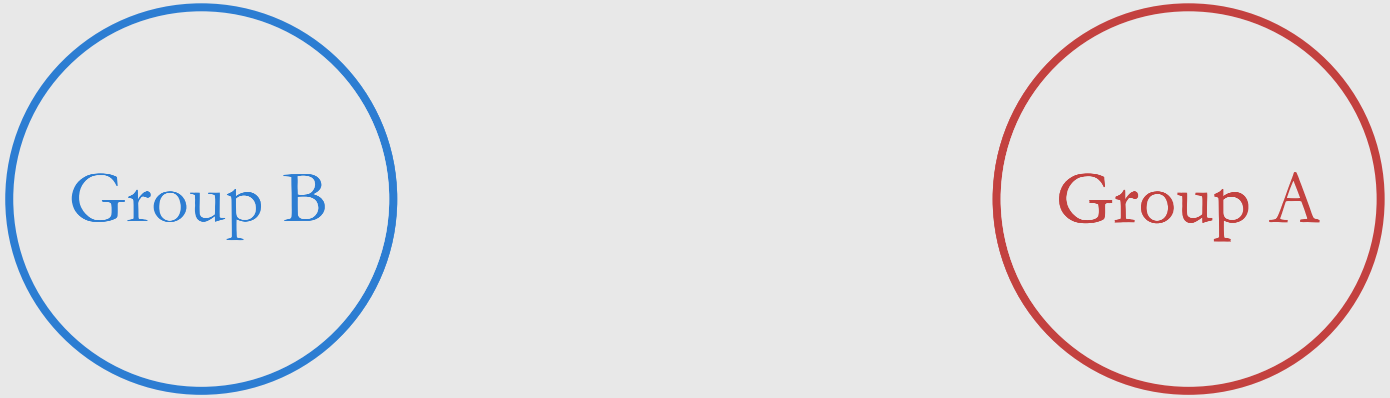
$T = 1$ $\qquad\qquad$ $T = 0$



$\mathbb{E}[Y \mid T = 1] = y_1$ $\qquad\qquad$ $\mathbb{E}[Y \mid T = 0] = y_0$

Another perspective: exchangeability

$T = 1$ $\qquad\qquad$ $T = 0$



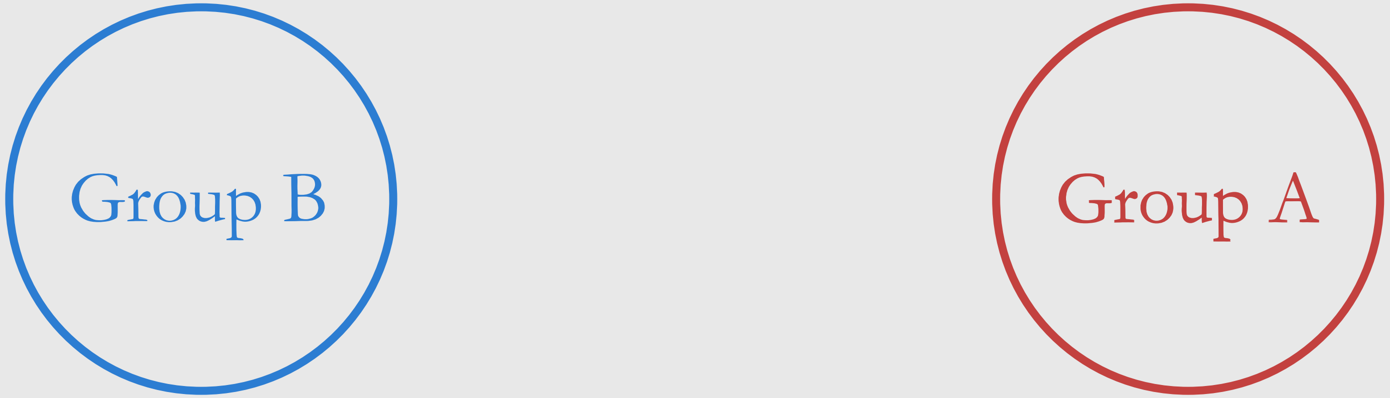
$\underline{\mathbb{E}[Y \mid T = 1] = y_1}$ $\qquad\qquad$ $\underline{\mathbb{E}[Y \mid T = 0] = y_0}$

Before switch $\qquad$ After switch

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

Another perspective: exchangeability

$T = 1$ $\qquad\qquad$ $T = 0$



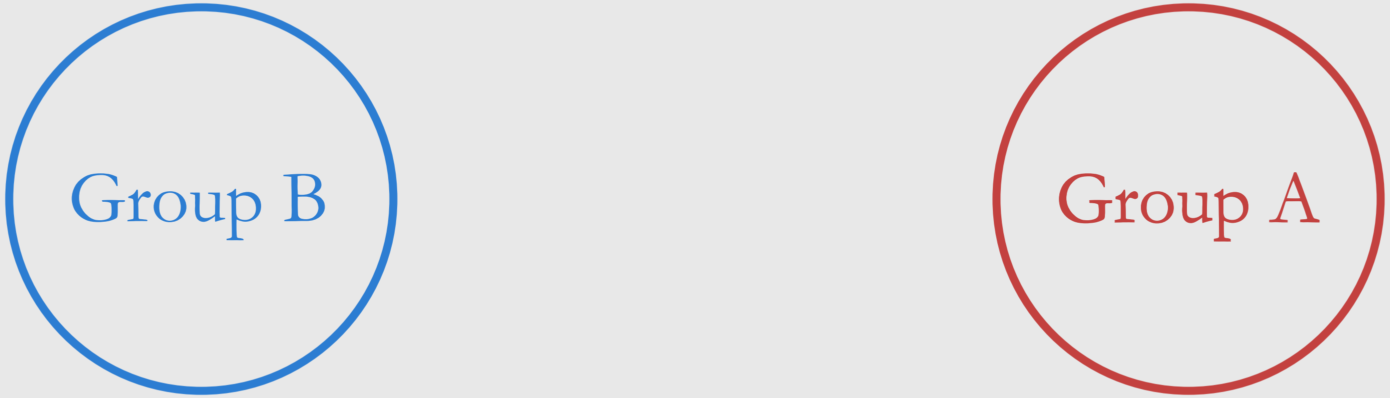
$\underline{\mathbb{E}[Y \mid T = 1] = y_1}$ $\qquad\qquad$ $\underline{\mathbb{E}[Y \mid T = 0] = y_0}$

Before switch $\qquad\qquad$ After switch

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

Another perspective: exchangeability

$T = 1$ $\qquad\qquad\qquad$ $T = 0$



$\mathbb{E}[Y \mid T = 1] = y_1$ $\qquad\qquad$ $\mathbb{E}[Y \mid T = 0] = y_0$

Before switch $\qquad\qquad$ After switch

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

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

Aside: identifiability

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

Aside: identifiability

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

Causal quantities

Aside: identifiability

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

Causal quantities

Statistical quantities

(accessible, since we have $P(x, t, y)$)

Aside: identifiability

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

Causal quantities

Statistical quantities

(accessible, since we have $P(x, t, y)$)

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

Randomized control trial (RCT)

 $\longrightarrow T$

Randomized control trial (RCT)

 $\longrightarrow T$

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

drunk
drunk drunk
drunk
drunk
drunk
drunk drunk
sober drunk

Went to sleep **without shoes** on $(T = 0)$

sober sober
sober sober drunk
sober sober sober
drunk sober
sober sober sober sober
sober drunk
sober sober

Randomized control trial (RCT)

 $\longrightarrow T$

Slept **with shoes** on $(T = 1)$

drunk
drunk drunk
drunk
drunk
drunk
drunk drunk
sober drunk

Went to sleep **without shoes** on $(T = 0)$

sober sober
sober sober drunk
sober sober sober
sober drunk sober
sober sober sober
sober
sober drunk
sober sober

Randomized control trial (RCT)

 $\longrightarrow T$

Slept **with shoes on** $(T = 1)$

drunk
drunk drunk
drunk
drunk
drunk
drunk drunk
sober drunk

Slept **without shoes on** $(T = 0)$

sober sober
sober sober drunk
sober sober sober
drunk sober
sober
sober sober sober
sober drunk
sober sober

Randomized control trial (RCT)

 $\longrightarrow T$

Slept **with shoes** on $(T = 1)$

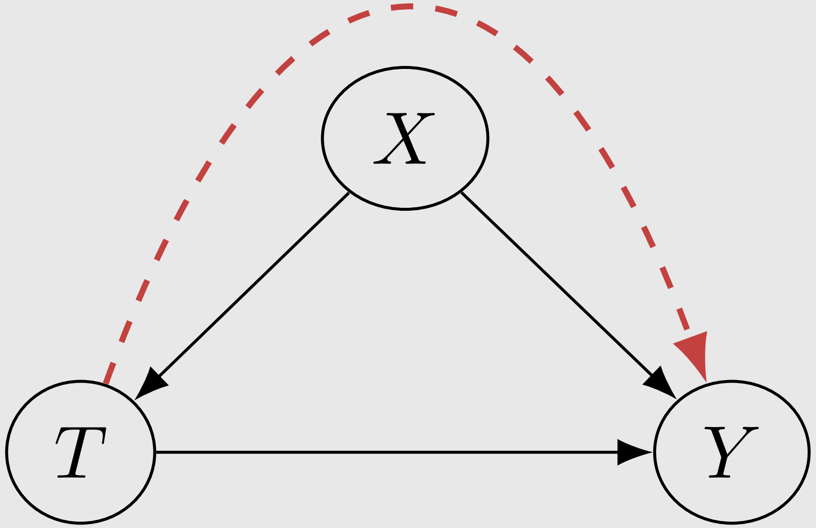
drunk
sober sober
sober
drunk drunk drunk
sober sober
drunk

Slept **without shoes** on $(T = 0)$

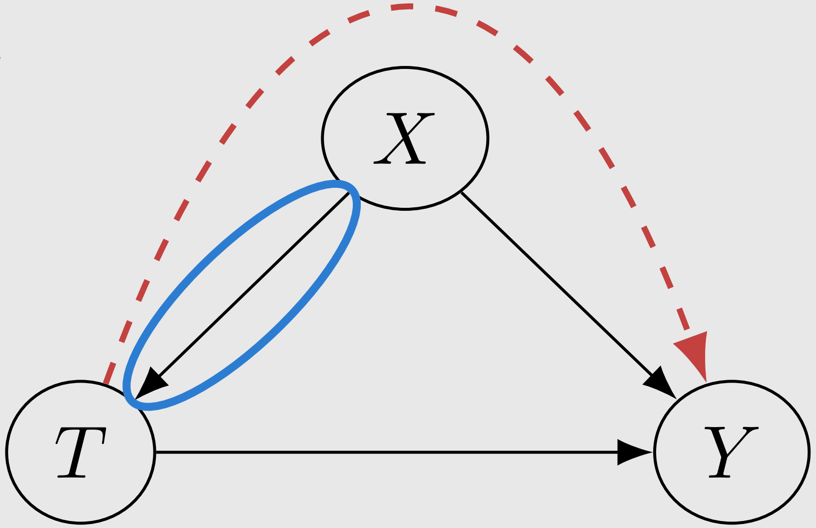
drunk
sober sober
sober
drunk drunk drunk
sober sober
drunk

Graphical interpretation of RCT

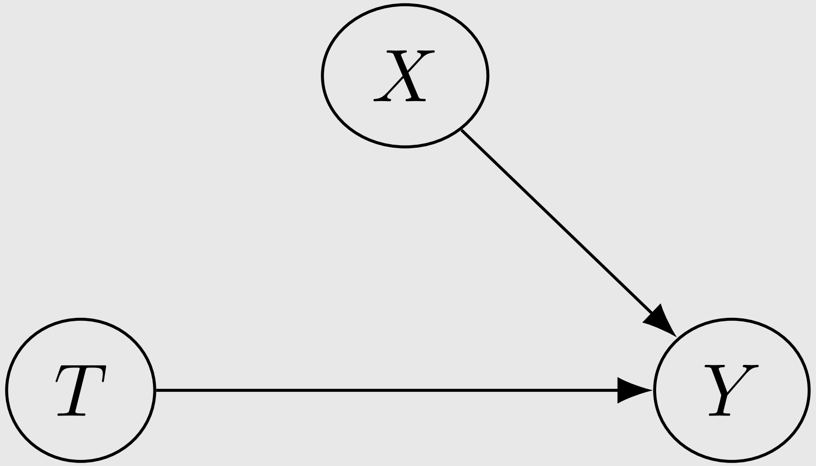
Graphical interpretation of RCT



Graphical interpretation of RCT



Graphical interpretation of RCT



Question:
What important property does an RCT give us?

Conditional exchangeability

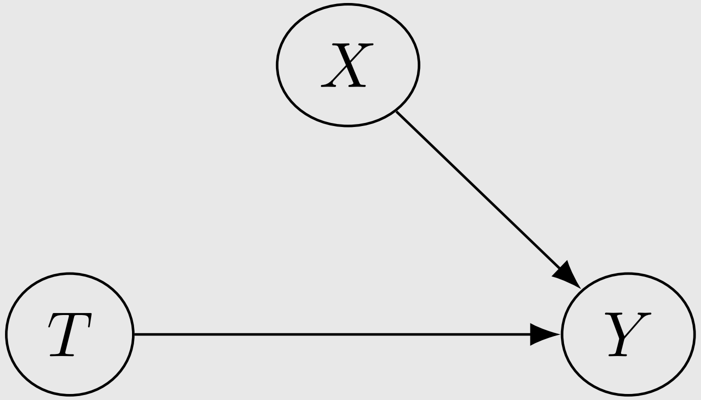
Exchangeability:

$$(Y(1), Y(0)) \perp\!\!\!\perp T$$

Conditional exchangeability

Exchangeability:

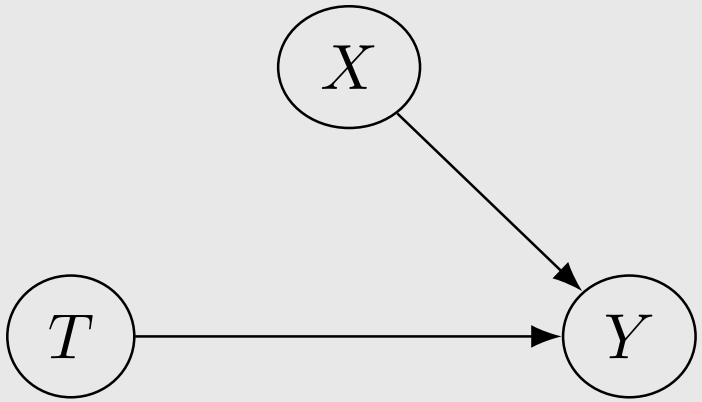
$$(Y(1), Y(0)) \perp\!\!\!\perp T$$

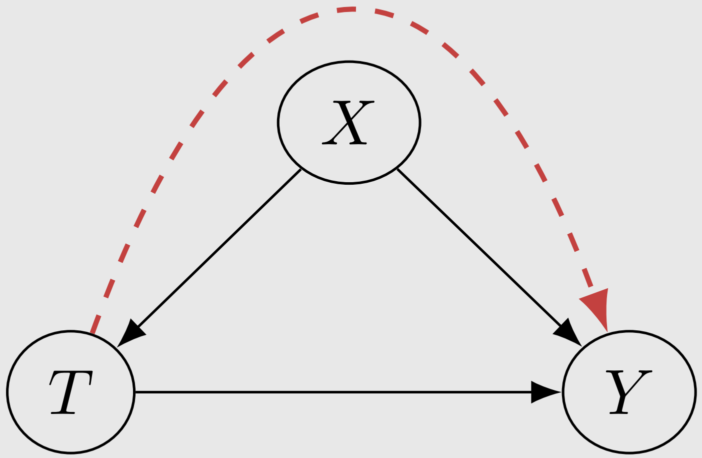


Conditional exchangeability

Exchangeability:

$$(Y(1), Y(0)) \perp\!\!\!\perp T$$

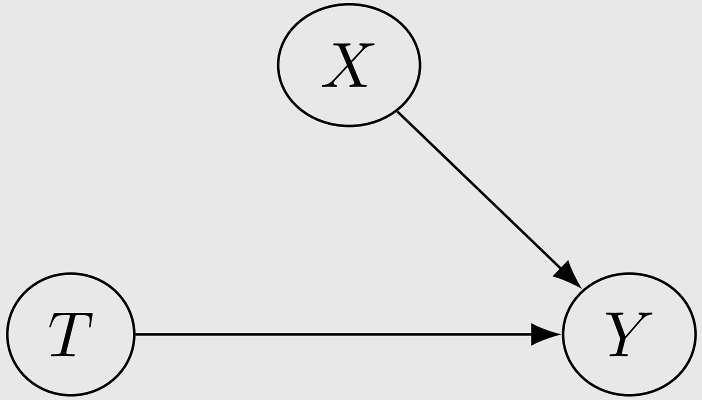




Conditional exchangeability

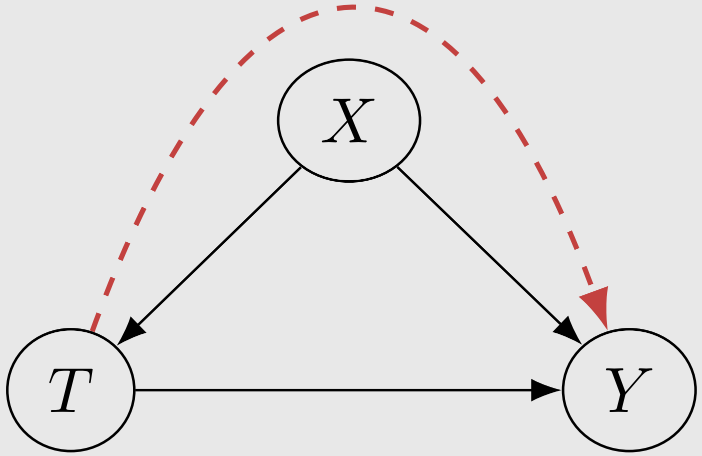
Exchangeability:

$(Y(1), Y(0)) \perp\!\!\!\perp T$



Conditional exchangeability:

$(Y(1), Y(0)) \perp\!\!\!\perp T \mid X$



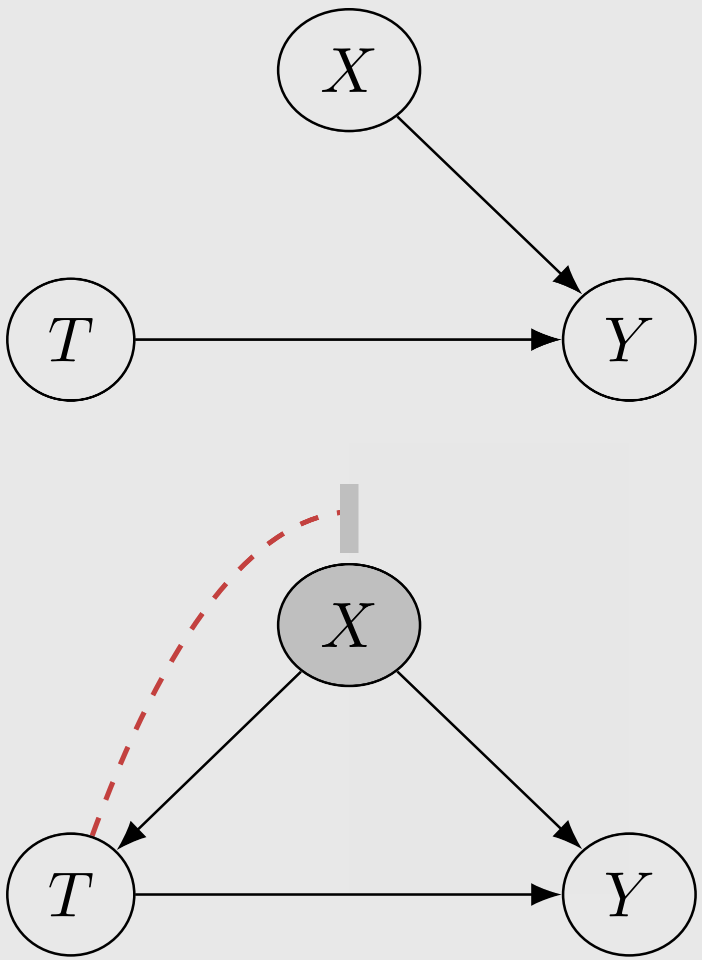
Conditional exchangeability

Exchangeability:

$$(Y(1), Y(0)) \perp\!\!\!\perp T$$

Conditional exchangeability:

$$(Y(1), Y(0)) \perp\!\!\!\perp T \mid X$$



Identification of conditional average treatment effect

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

Identification of conditional average treatment effect

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

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

Identification of conditional average treatment effect

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

$$
\begin{aligned}
\mathbb{E}[Y(1) - Y(0) \mid X] &= \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X] \\
&= \mathbb{E}[Y(1) \mid T = 1,\ X] - \mathbb{E}[Y(0) \mid T = 0,\ X]
\end{aligned}
$$

Identification of conditional average treatment effect

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

$$\begin{aligned}
\mathbb{E}[Y(1) - Y(0) \mid X] &= \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X] \\
&= \mathbb{E}[Y(1) \mid T = 1,\ X] - \mathbb{E}[Y(0) \mid T = 0,\ X] \\
&= \mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]
\end{aligned}$$

Identification of conditional average treatment effect

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

$$\begin{aligned}
\mathbb{E}[Y(1) - Y(0) \mid X] &= \mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X] \\
&= \mathbb{E}[Y(1) \mid T = 1,\ X] - \mathbb{E}[Y(0) \mid T = 0,\ X] \\
&= \mathbb{E}[Y \mid T = 1,\ X] - \mathbb{E}[Y \mid T = 0,\ X]
\end{aligned}$$

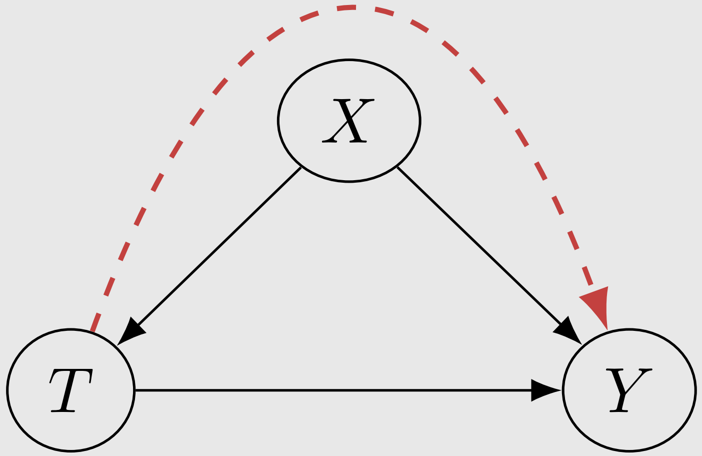
What about the ATE? $\mathbb{E}[Y(1) - Y(0)]$

The Adjustment Formula (identification of ATE)

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

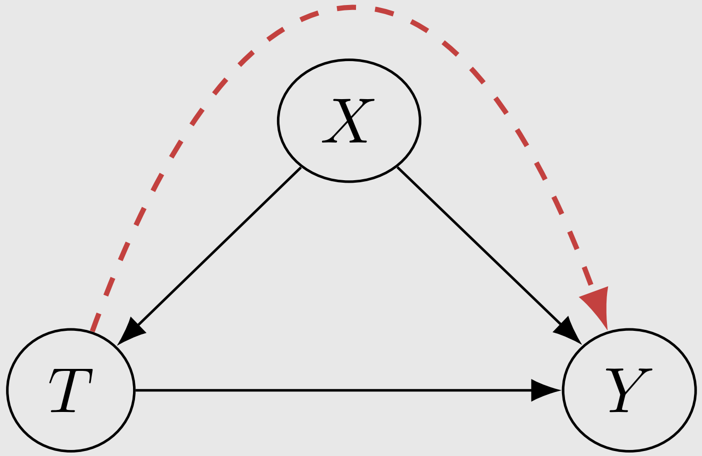
The Adjustment Formula (identification of ATE)

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



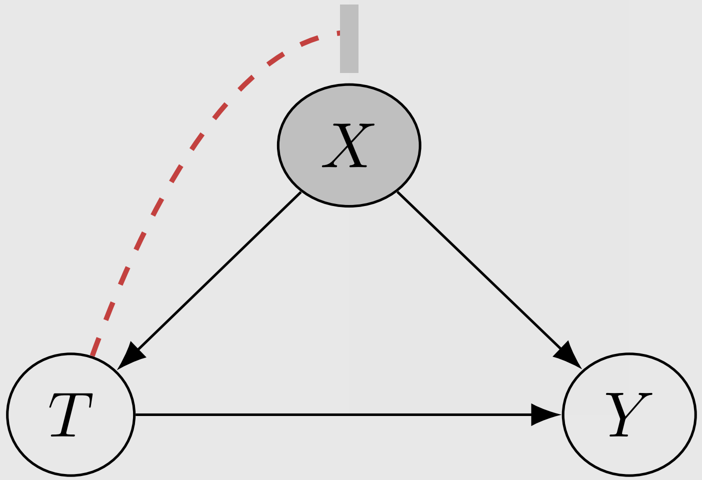
The Adjustment Formula (identification of ATE)

$$\begin{aligned}\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) \mid X] \\ &= \mathbb{E}_{\underline{X}} \left[\mathbb{E}[Y \mid T = 1, \underline{X}] - \mathbb{E}[Y \mid T = 0, \underline{X}]\right]\end{aligned}$$



The Adjustment Formula (identification of ATE)

$$\begin{aligned}\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}_X \mathbb{E}[Y(1) - Y(0) \mid X] \\ &= \mathbb{E}_{\underline{X}} \left[\mathbb{E}[Y \mid T = 1, \underline{X}] - \mathbb{E}[Y \mid T = 0, \underline{X}]\right]\end{aligned}$$



Unconfoundedness is an untestable assumption

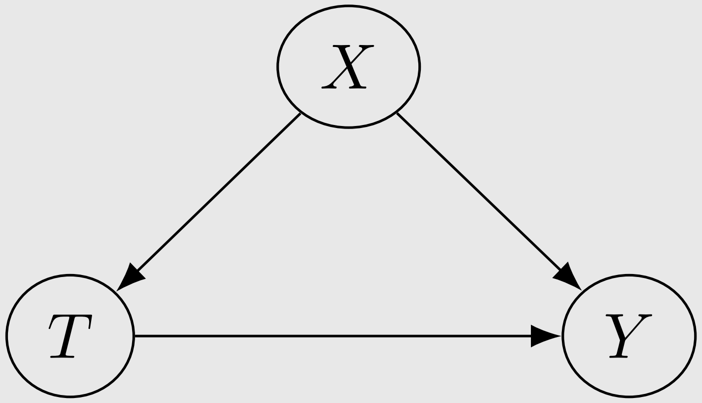
unconfoundedness = conditional ignorability = conditional exchangeability

Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

$$(Y(1), Y(0)) \perp\!\!\!\perp T \mid X$$

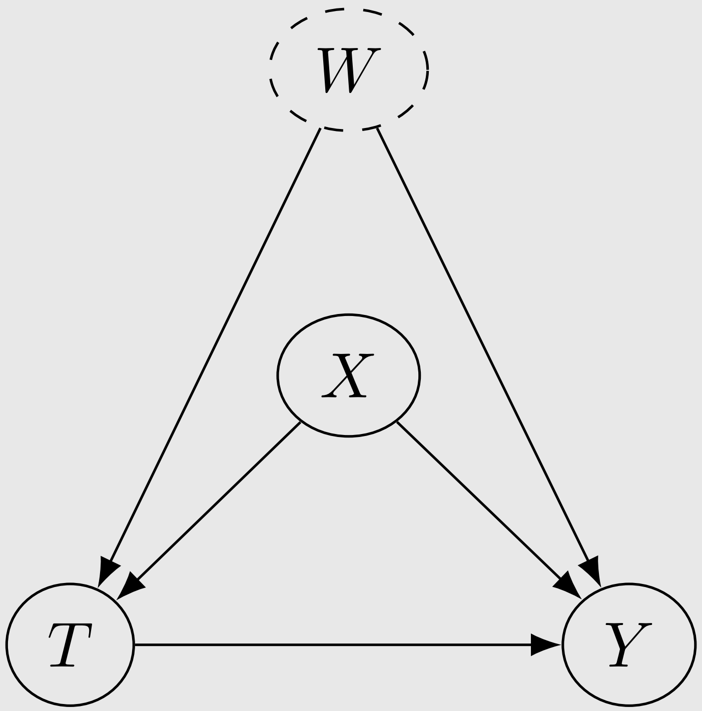


Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

$$(Y(1), Y(0)) \perp\!\!\!\perp T \mid X$$

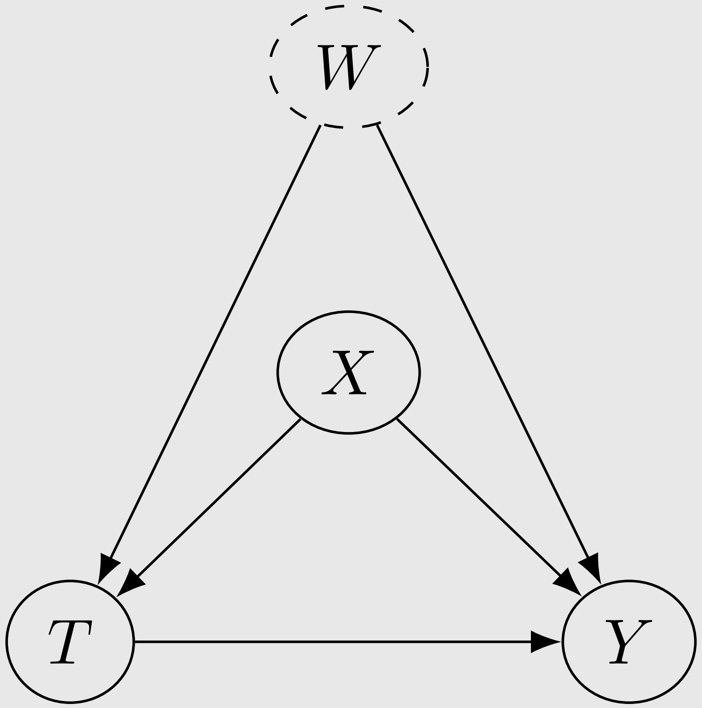


Unconfoundedness is an untestable assumption

unconfoundedness = conditional ignorability = conditional exchangeability

Conditional exchangeability:

$$(Y(1), Y(0)) \not\!\perp\!\!\!\perp T \mid X$$



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:

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

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:

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

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

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

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

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

Why? Recall the adjustment formula:

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

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

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

Positivity

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

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

Why? Recall the adjustment formula:

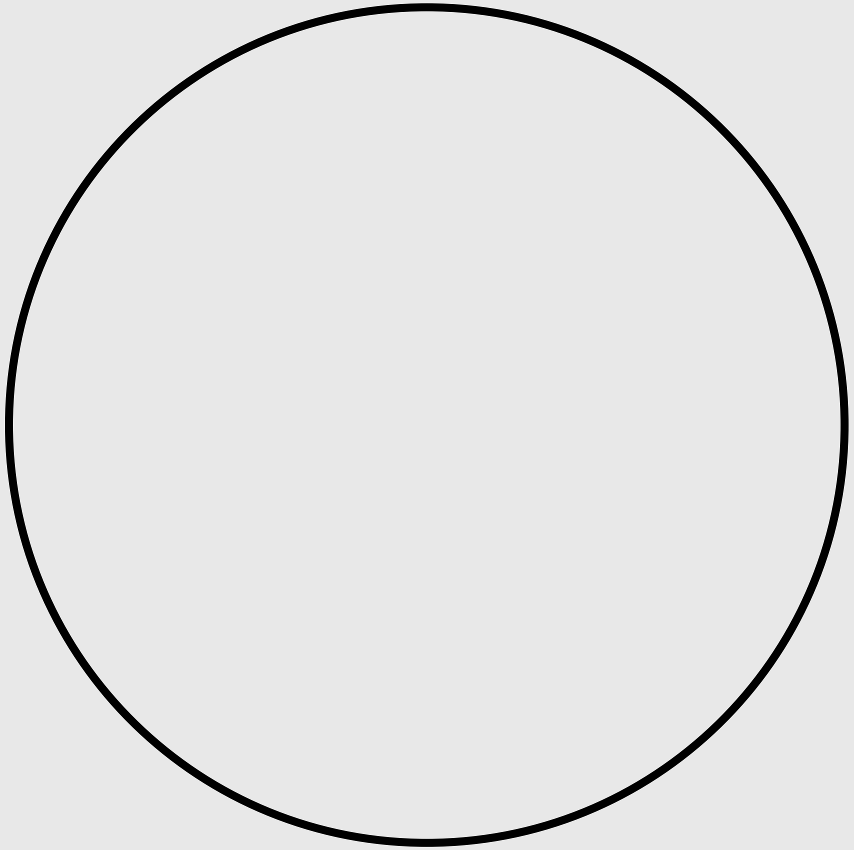
$$\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$$

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

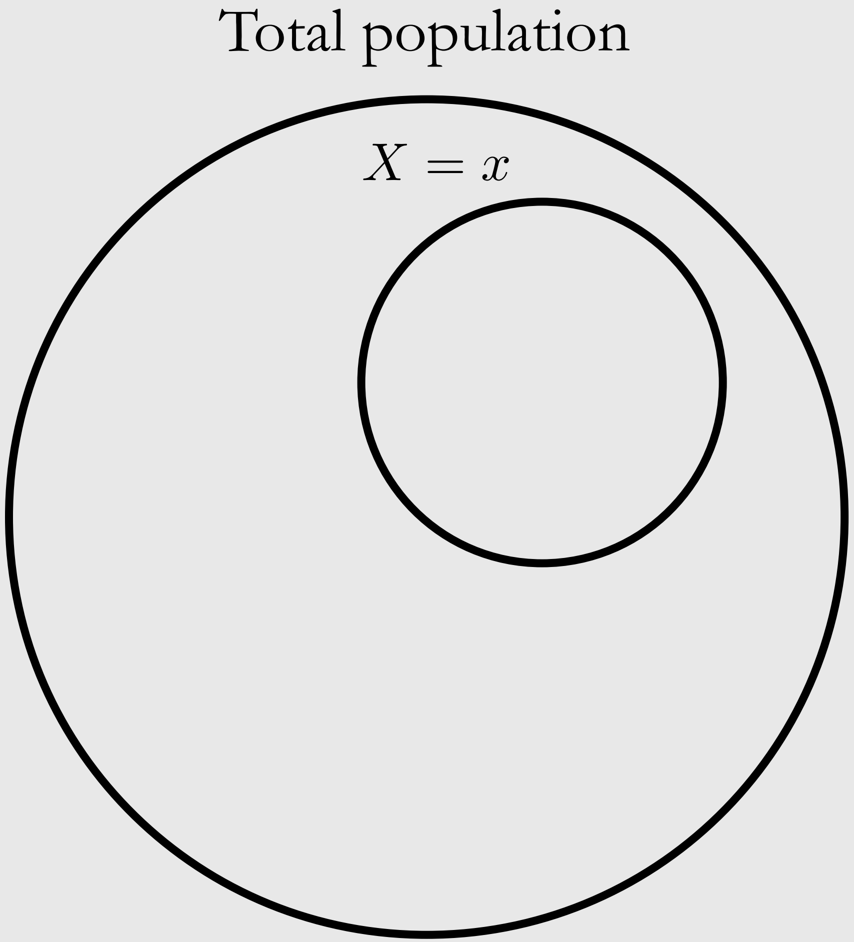
$$\sum_x \left(\sum_y y \, \frac{P(Y = y, T = 1, X = x)}{P(T = 1 \mid X = x) P(X = x)} - \sum_y y \, \frac{P(Y = y, T = 0, X = x)}{\underline{P(T = 0 \mid X = x)} P(X = x)} \right)$$

Positivity: intuition

Total population

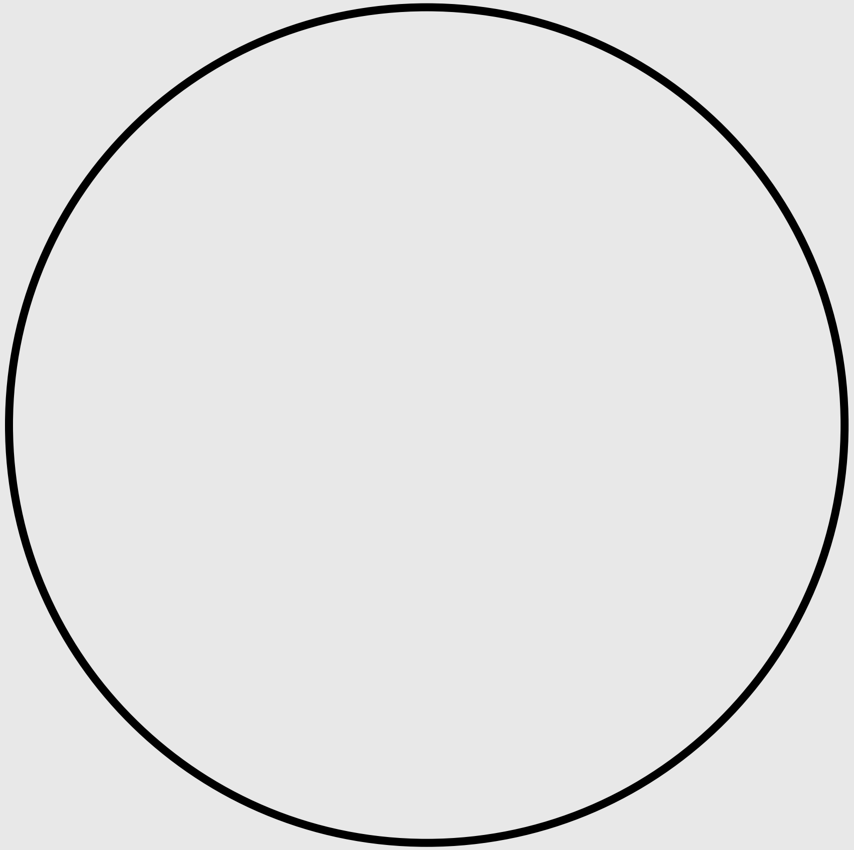


Positivity: intuition

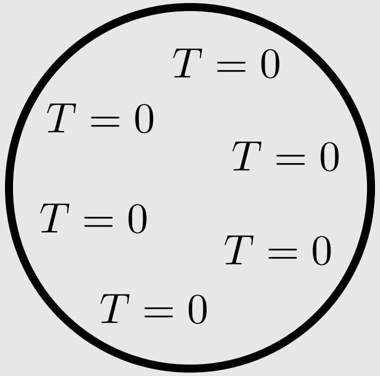


Positivity: intuition

Total population



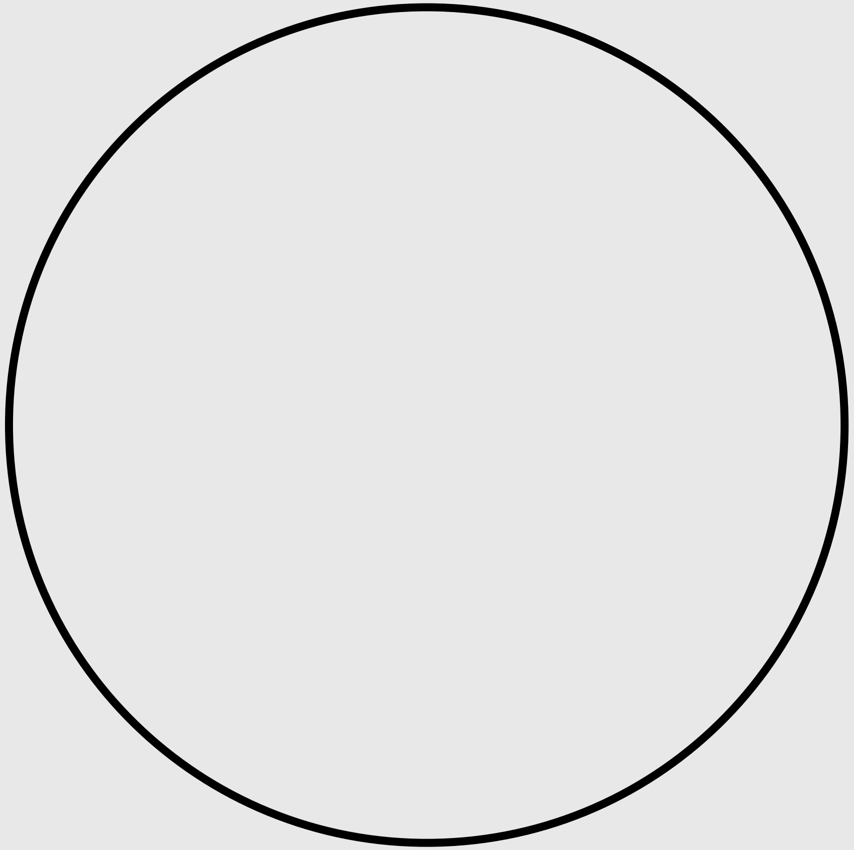
$X = x$



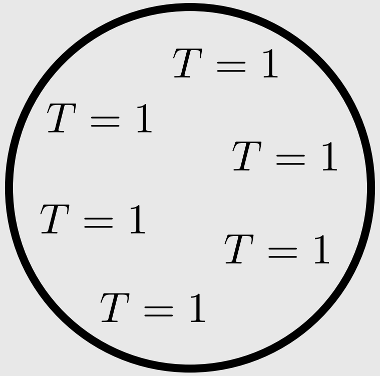
No one treated

Positivity: intuition

Total population



$X = x$

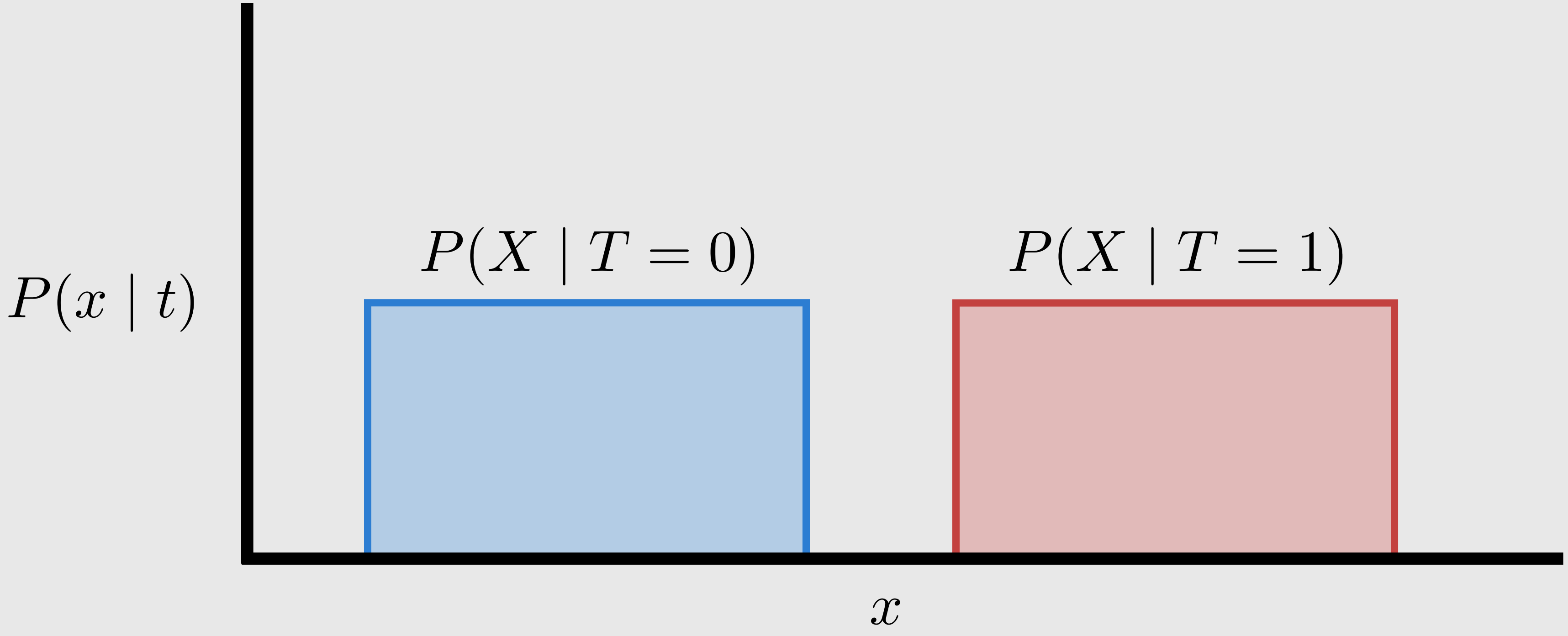


Everyone treated

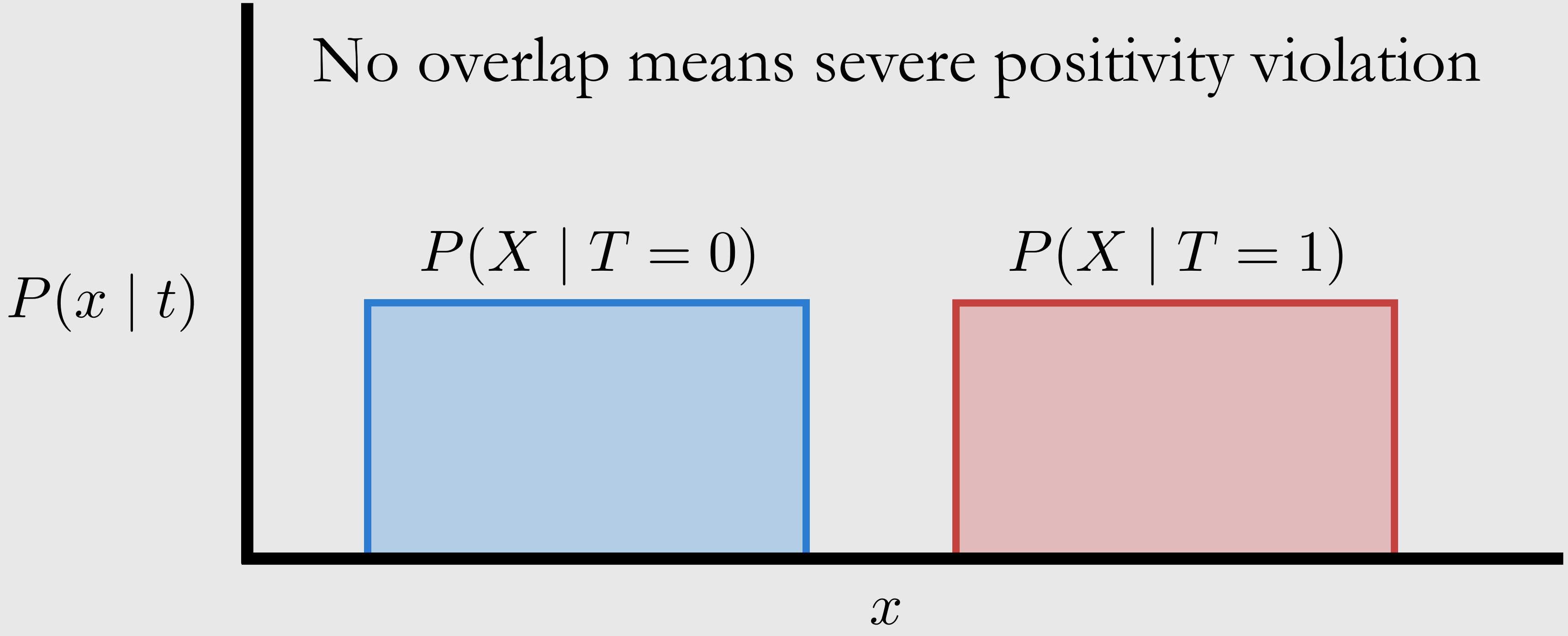
Another perspective: overlap

Overlap between $P(X \mid T = 0)$ and $P(X \mid T = 1)$

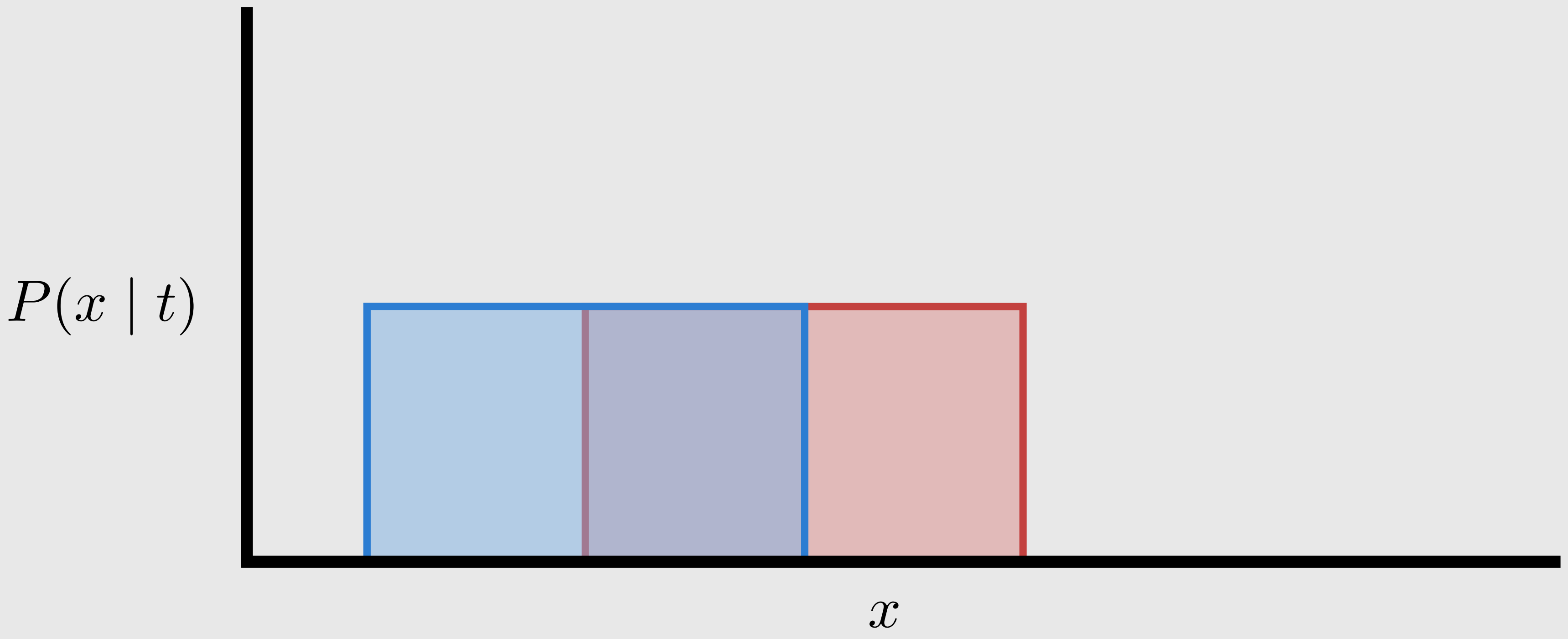
Another perspective: overlap



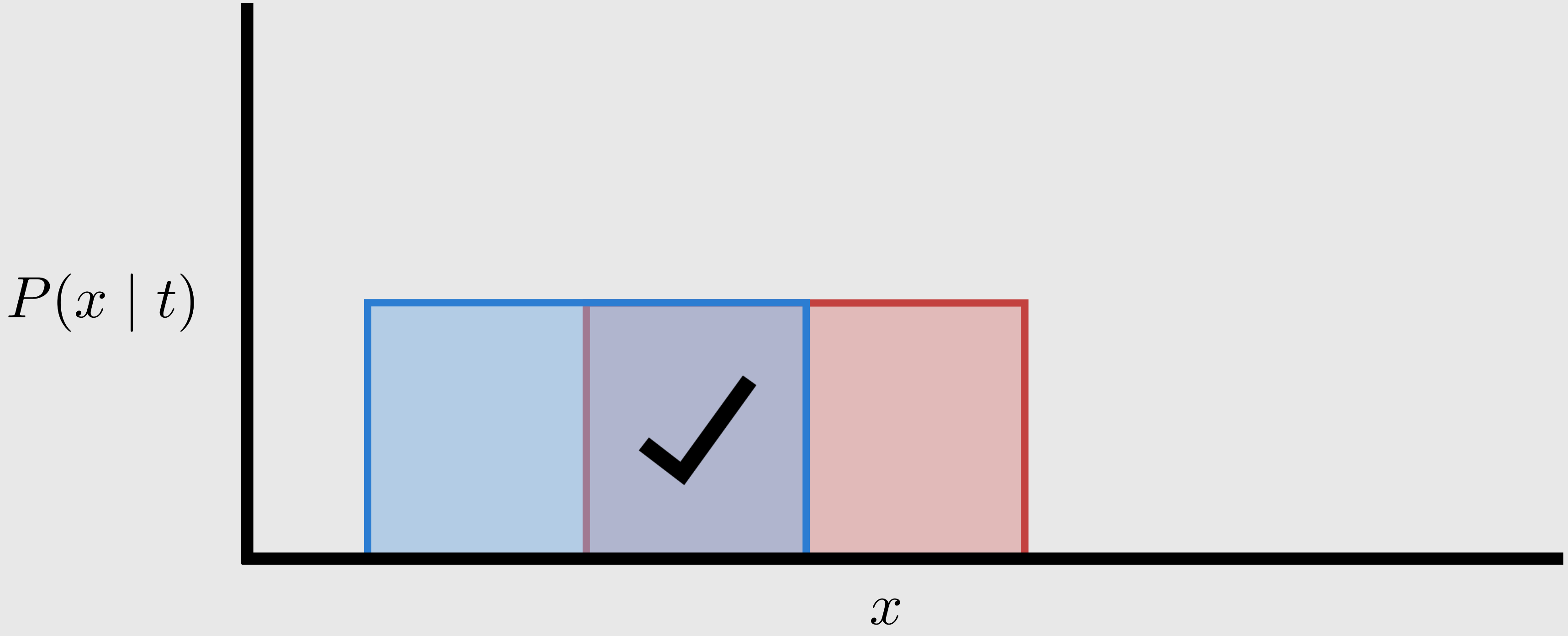
Another perspective: overlap



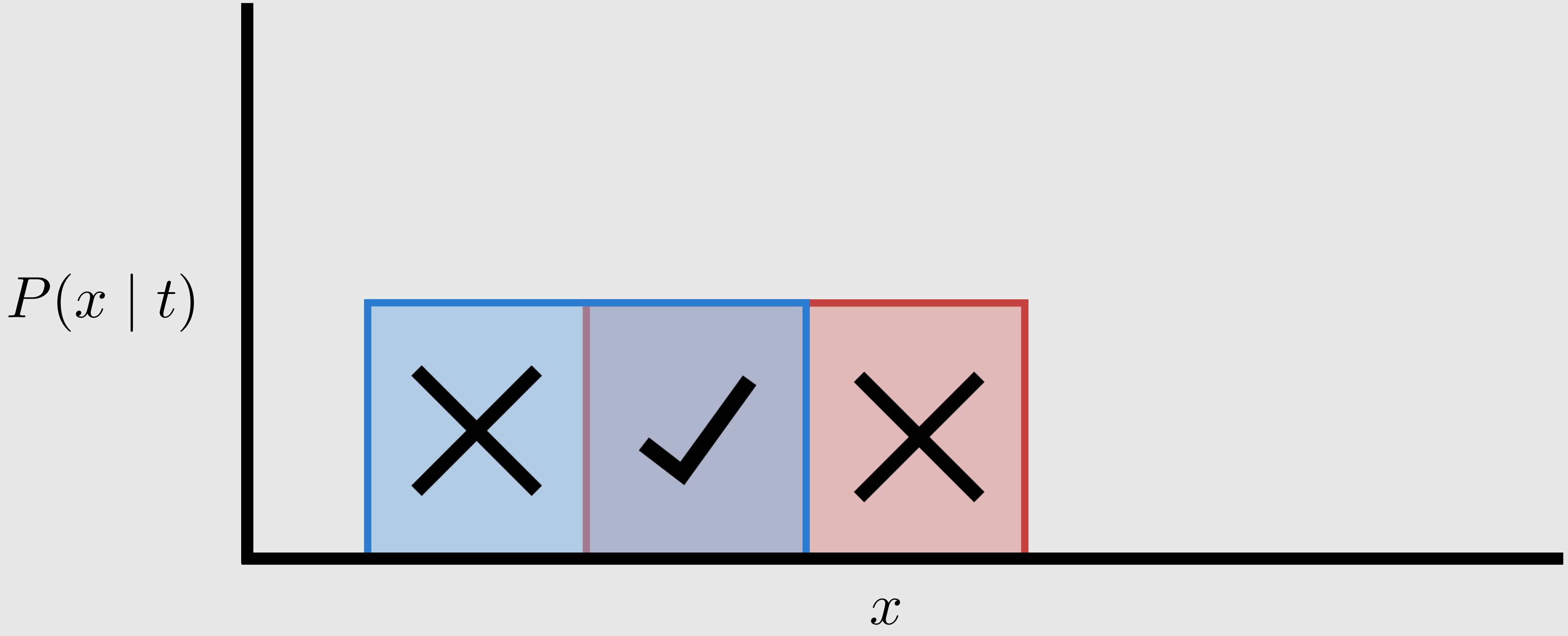
Another perspective: overlap



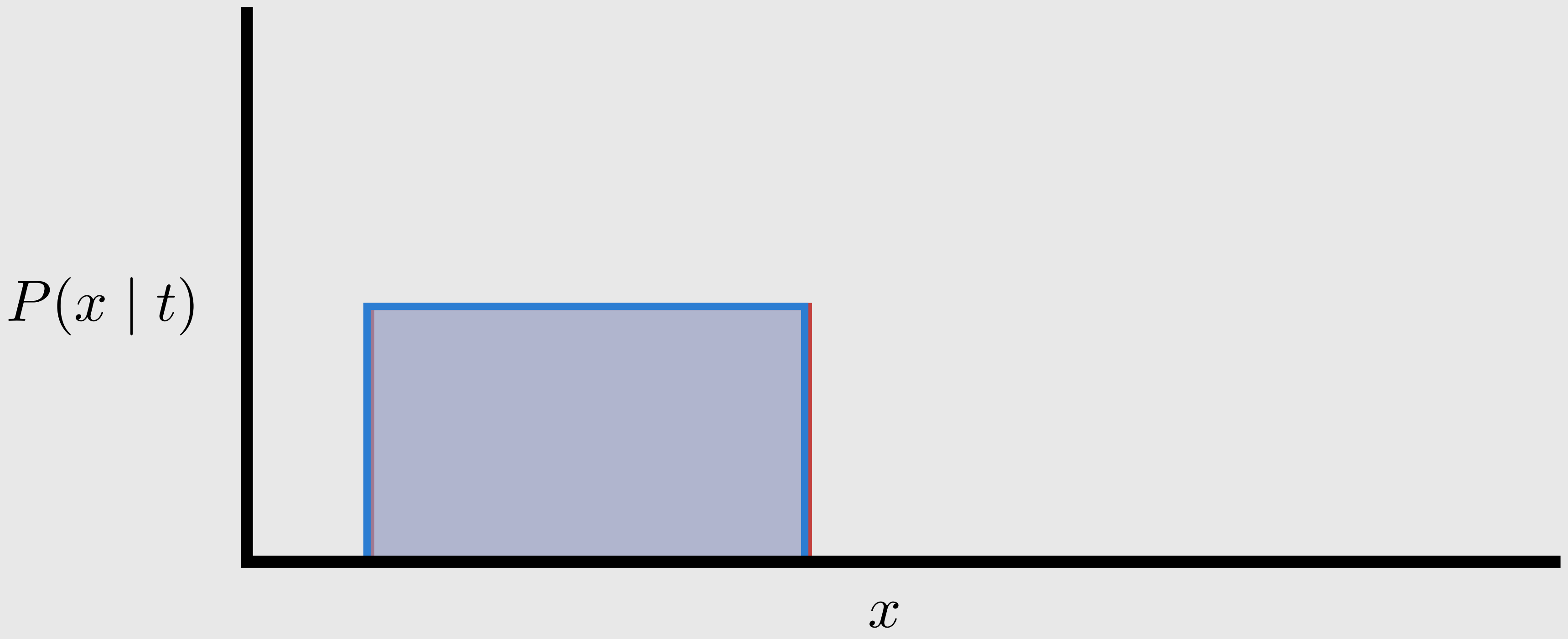
Another perspective: overlap



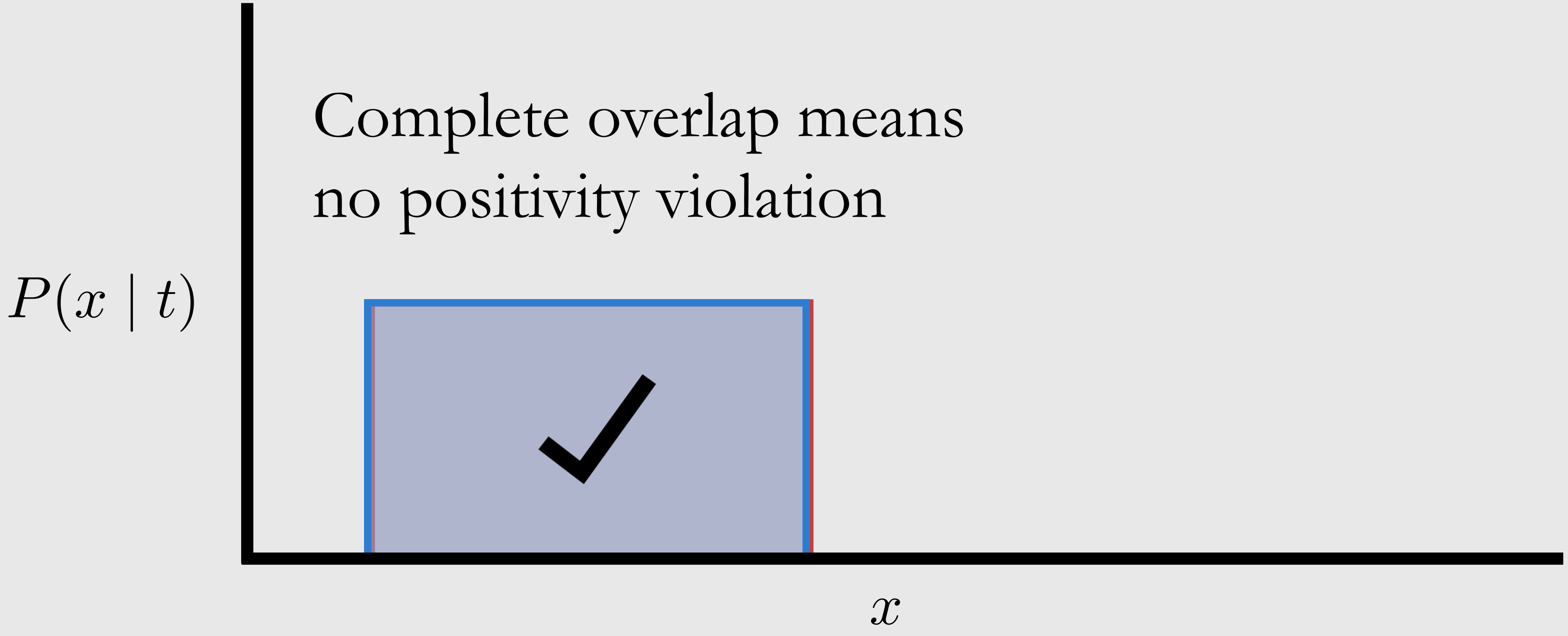
Another perspective: overlap



Another perspective: overlap

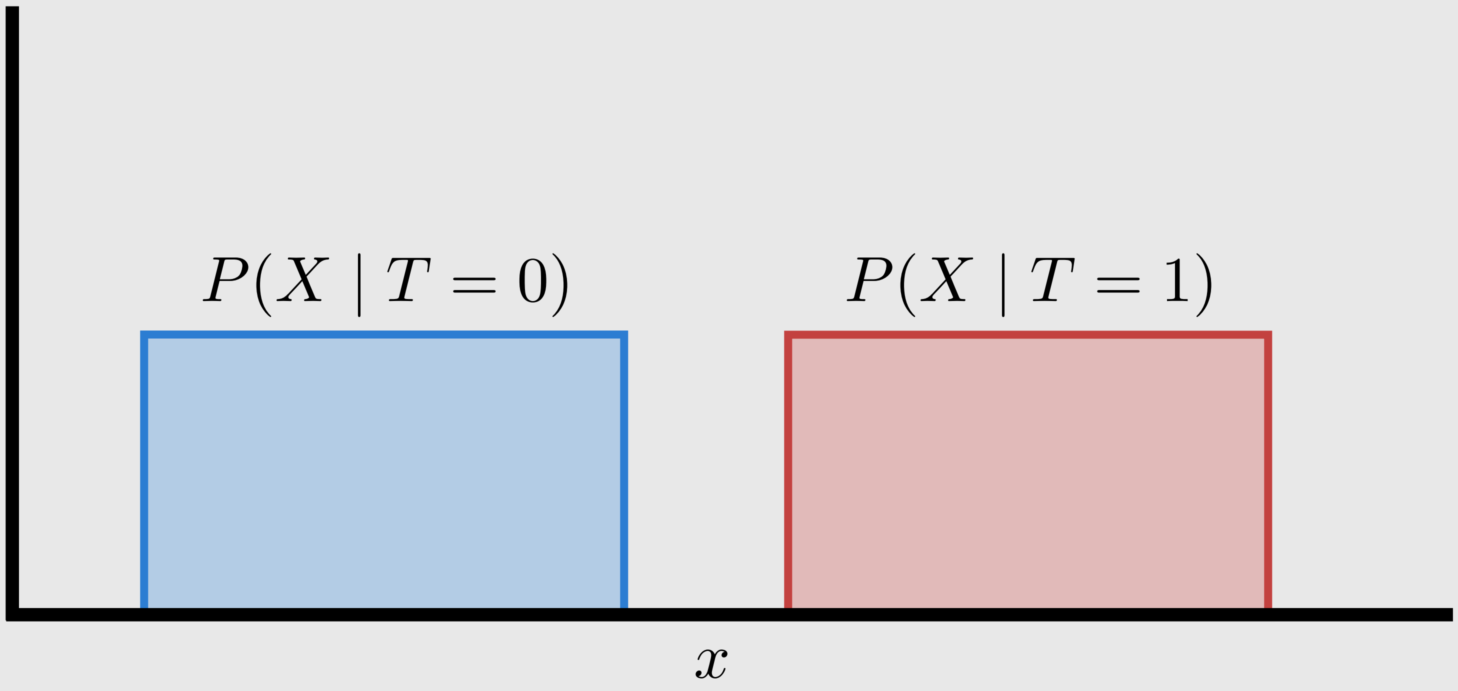


Another perspective: overlap



Question:
What goes wrong if we don’t have positivity?

The Positivity-Unconfoundedness Tradeoff



The Positivity-Unconfoundedness Tradeoff

The Positivity-Unconfoundedness Tradeoff

50% overlap

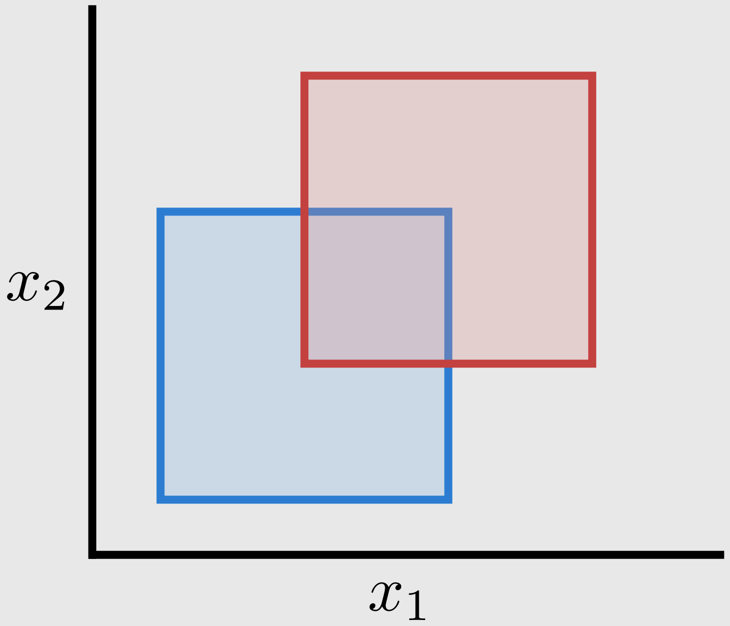
The Positivity-Unconfoundedness Tradeoff

1-dimensional

50% overlap

The Positivity-Unconfoundedness Tradeoff

2-dimensional



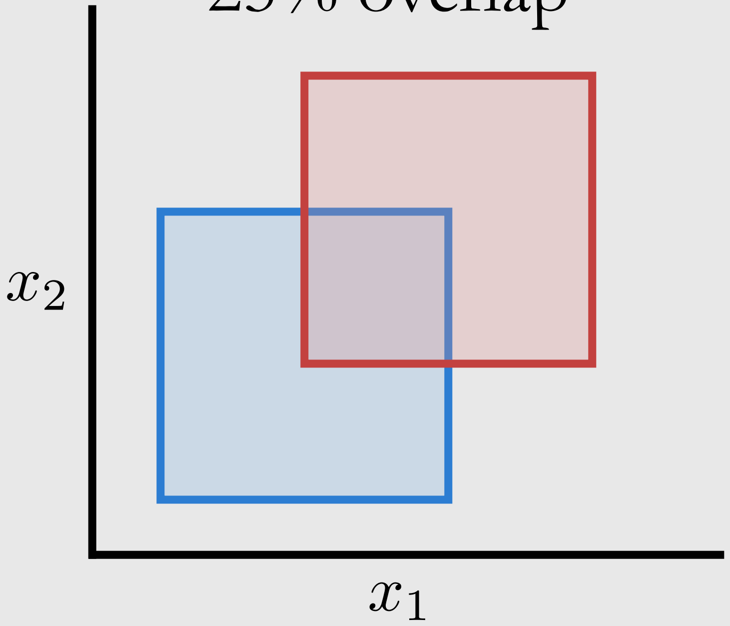
1-dimensional

50% overlap

The Positivity-Unconfoundedness Tradeoff

25% overlap

2-dimensional



1-dimensional

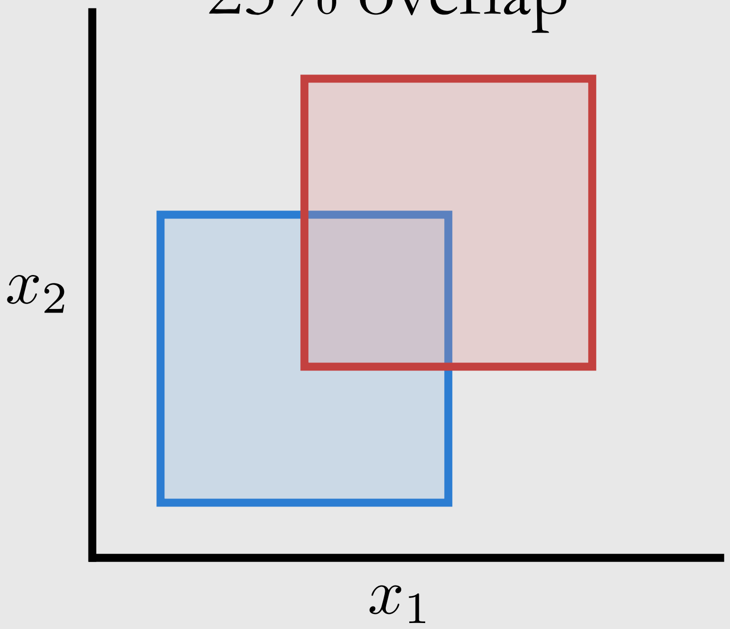
50% overlap

The Positivity-Unconfoundedness Tradeoff

25% overlap

3-dimensional – 12.5% overlap

2-dimensional



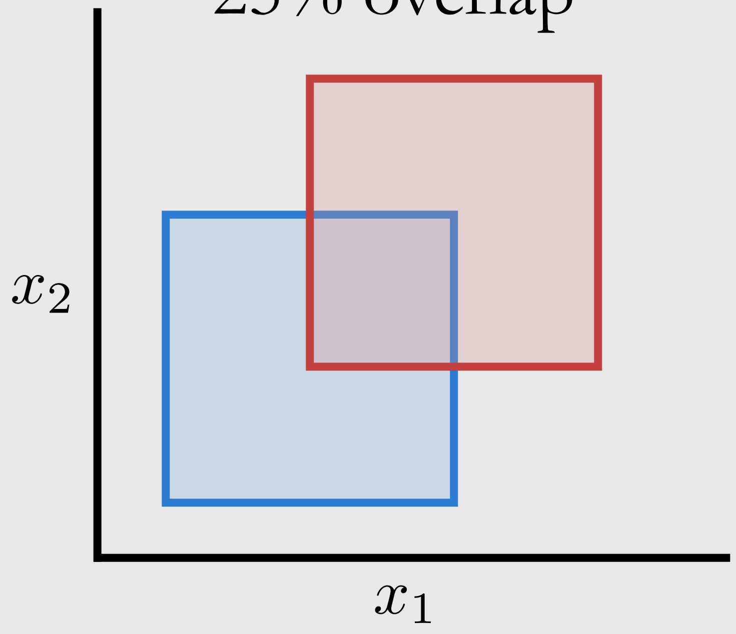
1-dimensional

50% overlap

The Positivity-Unconfoundedness Tradeoff

25% overlap

2-dimensional



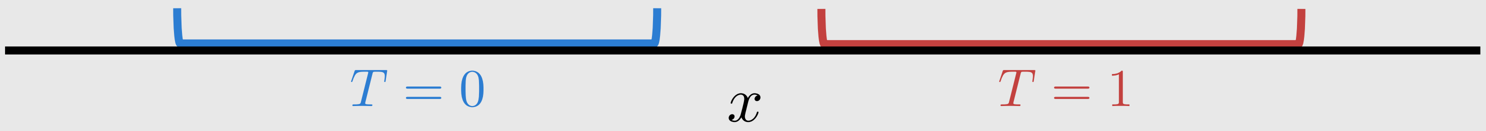
3-dimensional – 12.5% overlap

… and so on
(curse of dimensionality)

1-dimensional

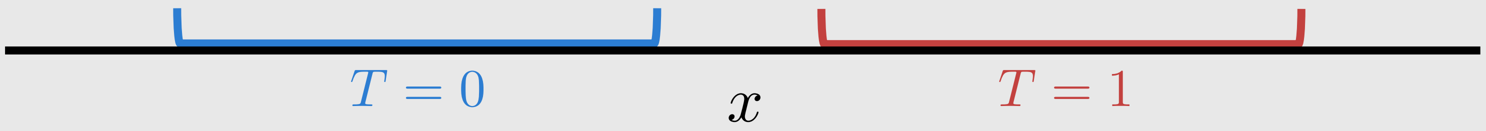
50% overlap

Extrapolation



Extrapolation

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



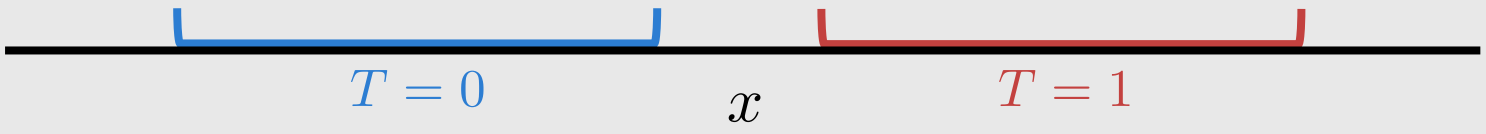
Extrapolation

Model

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

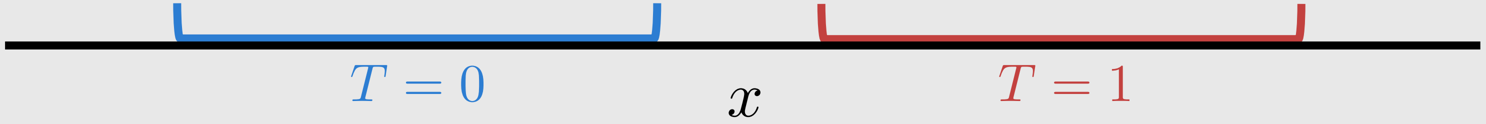
with

$f_1(x)$



Extrapolation

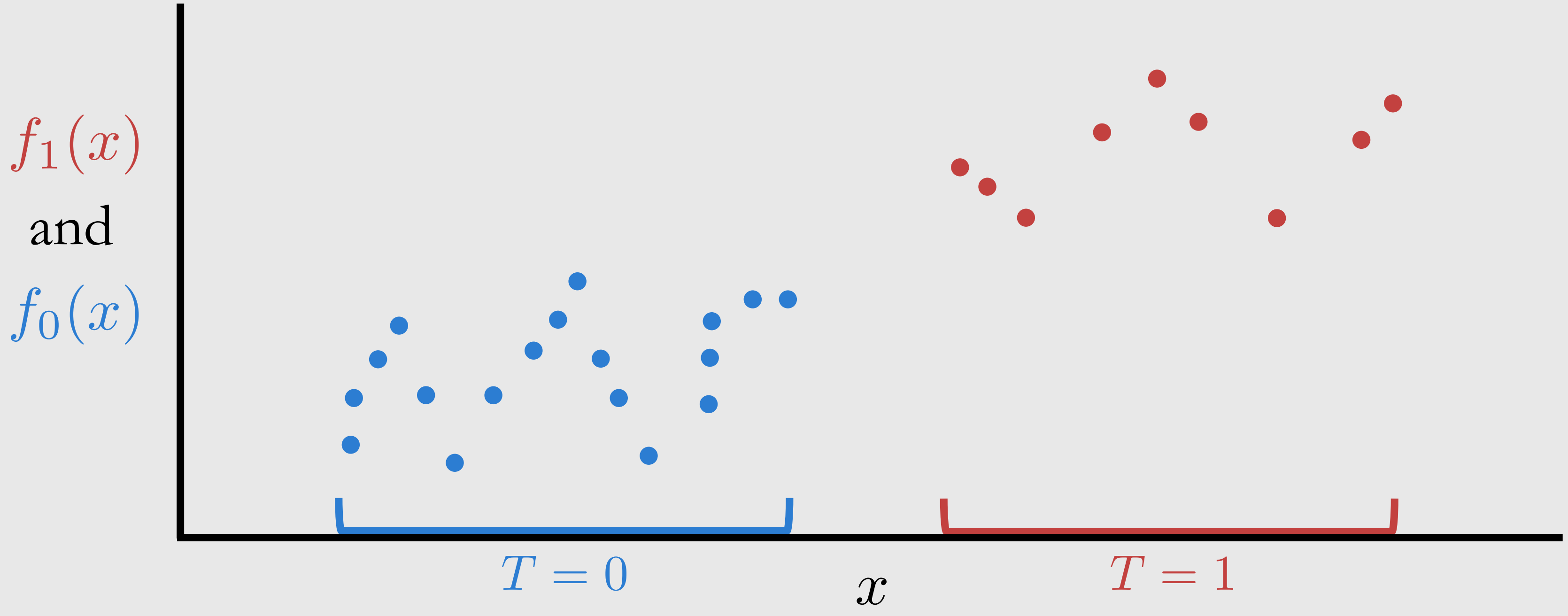
Adjustment formula: $$\sum_x \left(\overset{\text{Model}}{\underset{\text{with } f_1(x)}{\mathbb{E}[Y \mid T = 1, x]}} - \overset{\text{Model}}{\underset{\text{with } f_0(x)}{\mathbb{E}[Y \mid T = 0, x]}} \right)$$



Extrapolation

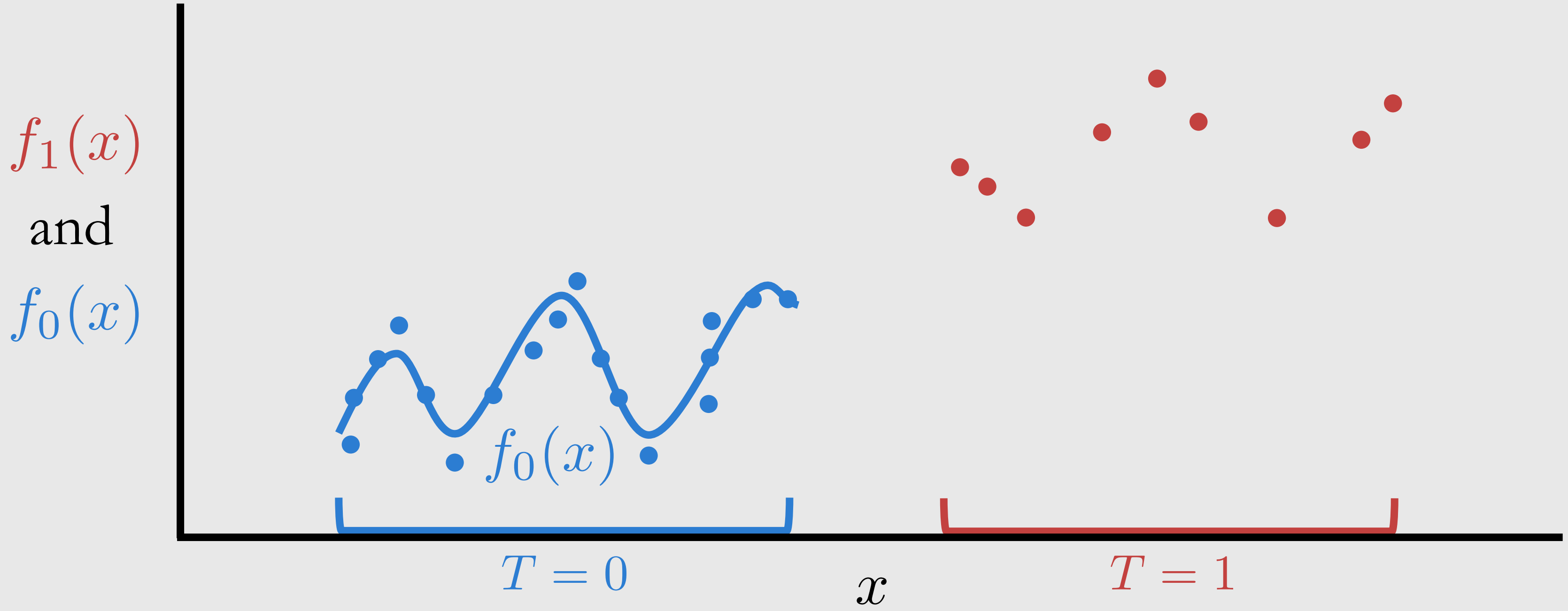
Adjustment formula: $\sum_x (\mathbb{E}[Y \mid T=1, x] - \mathbb{E}[Y \mid T=0, x])$

Model with $f_1(x)$ — Model with $f_0(x)$



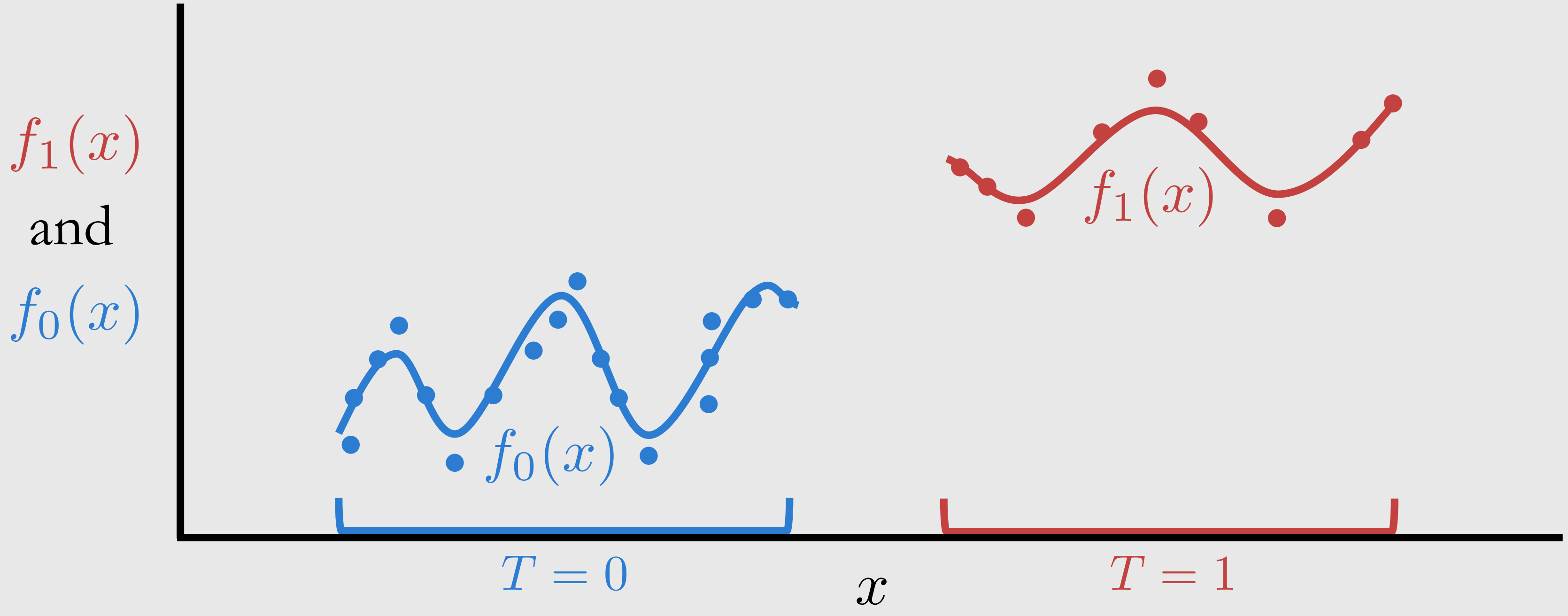
Extrapolation

Adjustment formula: $\sum_x \left(\underbrace{\mathbb{E}[Y \mid T = 1, x]}_{\text{Model with } f_1(x)} - \underbrace{\mathbb{E}[Y \mid T = 0, x]}_{\text{Model with } f_0(x)} \right)$



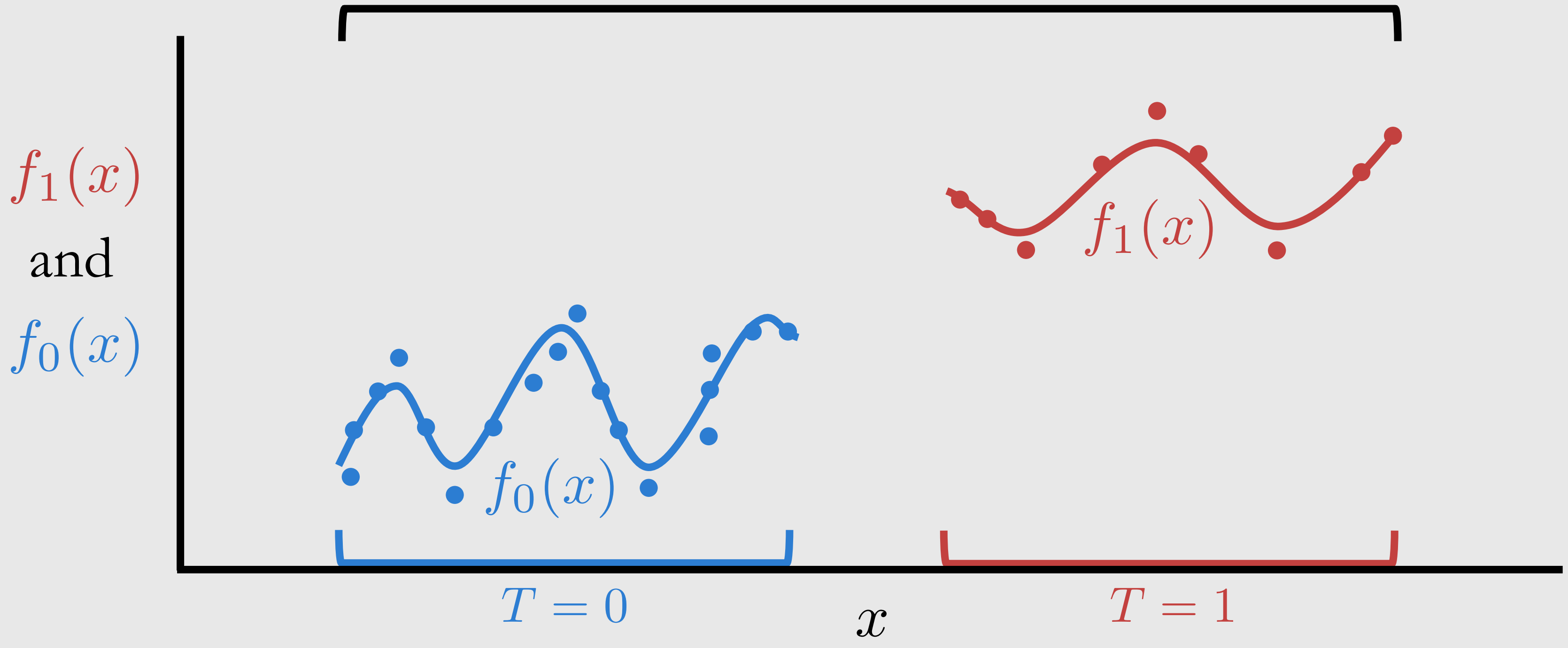
Extrapolation

$$\text{Adjustment formula:} \quad \sum_x \Big(\underset{\substack{\text{with} \\ f_1(x)}}{\overset{\text{Model}}{\mathbb{E}[Y \mid T = 1, x]}} - \underset{\substack{\text{with} \\ f_0(x)}}{\overset{\text{Model}}{\mathbb{E}[Y \mid T = 0, x]}} \Big)$$



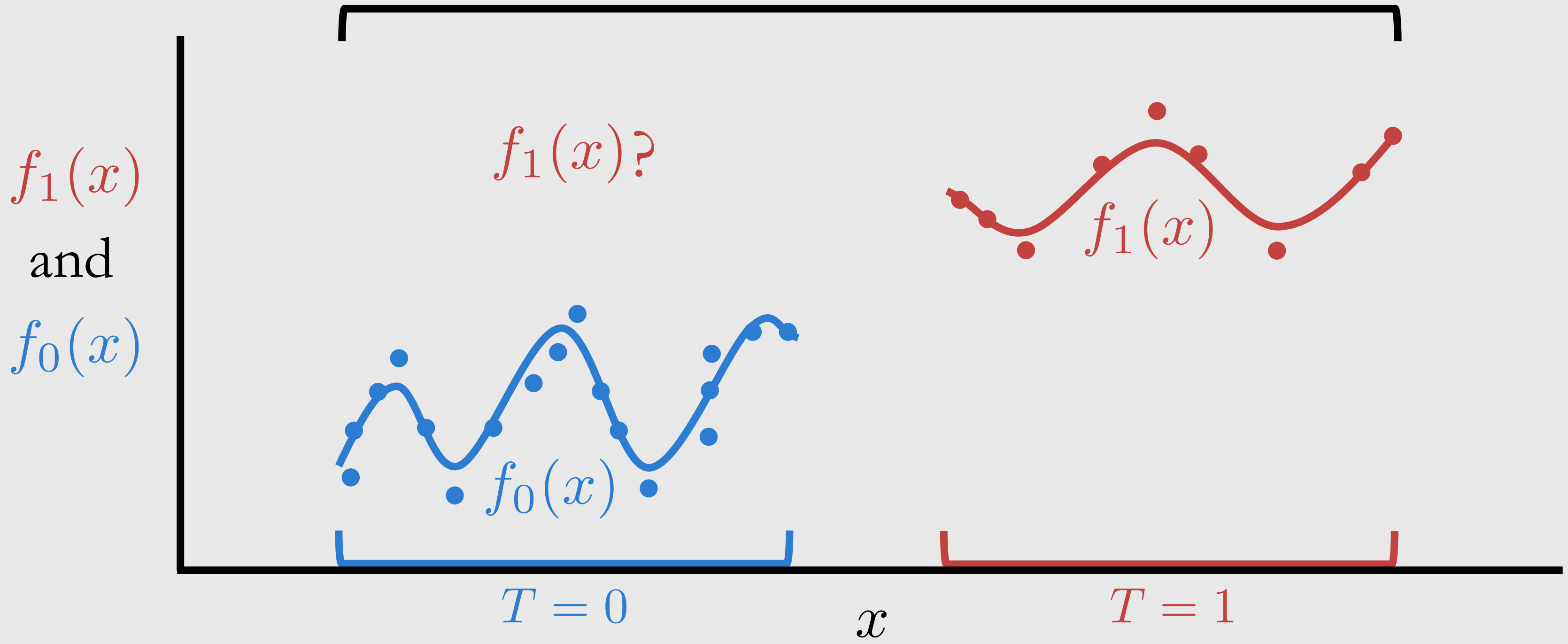
Extrapolation

Adjustment formula: $\sum_x \left(\underbrace{\overbrace{\mathbb{E}[Y \mid T=1, x]}^{\text{Model}}}_{\text{with } f_1(x)} - \underbrace{\overbrace{\mathbb{E}[Y \mid T=0, x]}^{\text{Model}}}_{\text{with } f_0(x)} \right)$



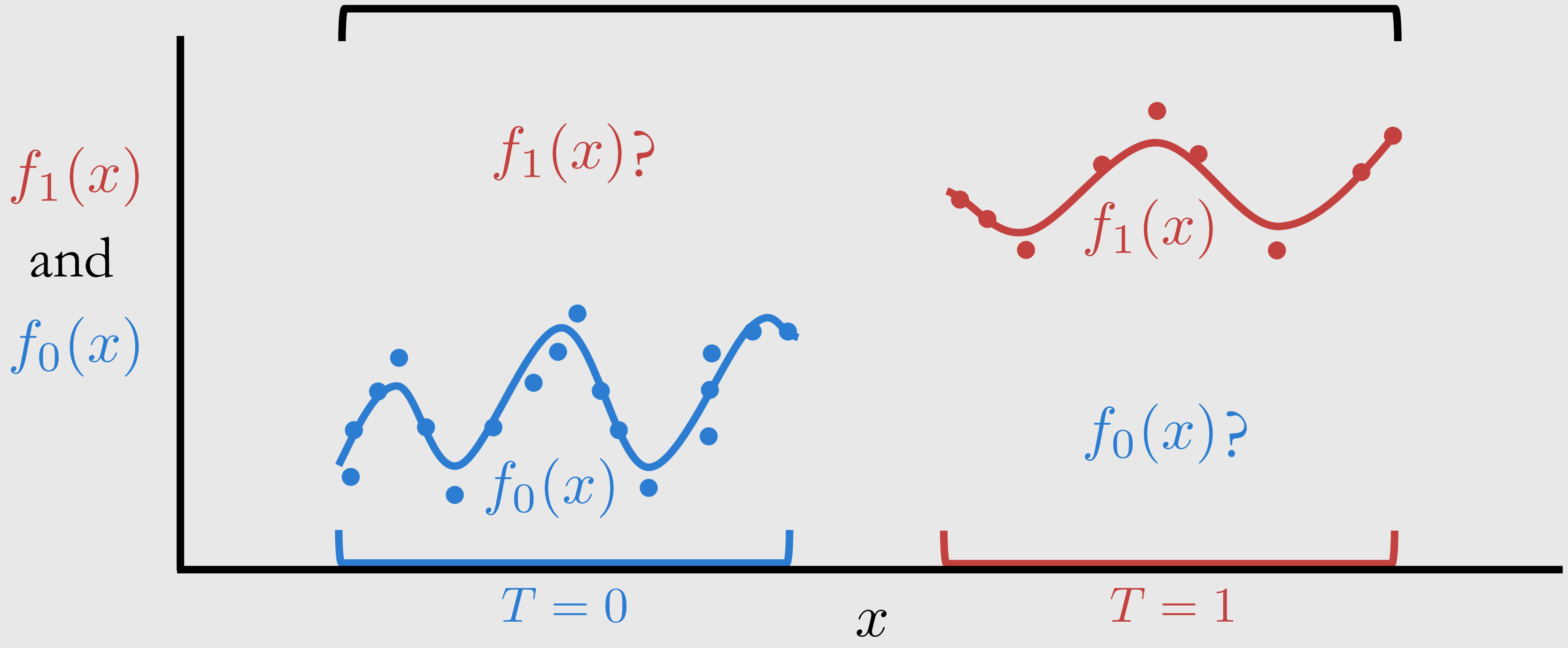
Extrapolation

$$\text{Adjustment formula: } \sum_x \big(\underbrace{\mathbb{E}[Y \mid T=1, x]}_{\text{Model with } f_1(x)}^{} - \underbrace{\mathbb{E}[Y \mid T=0, x]}_{\text{Model with } f_0(x)}\big)$$



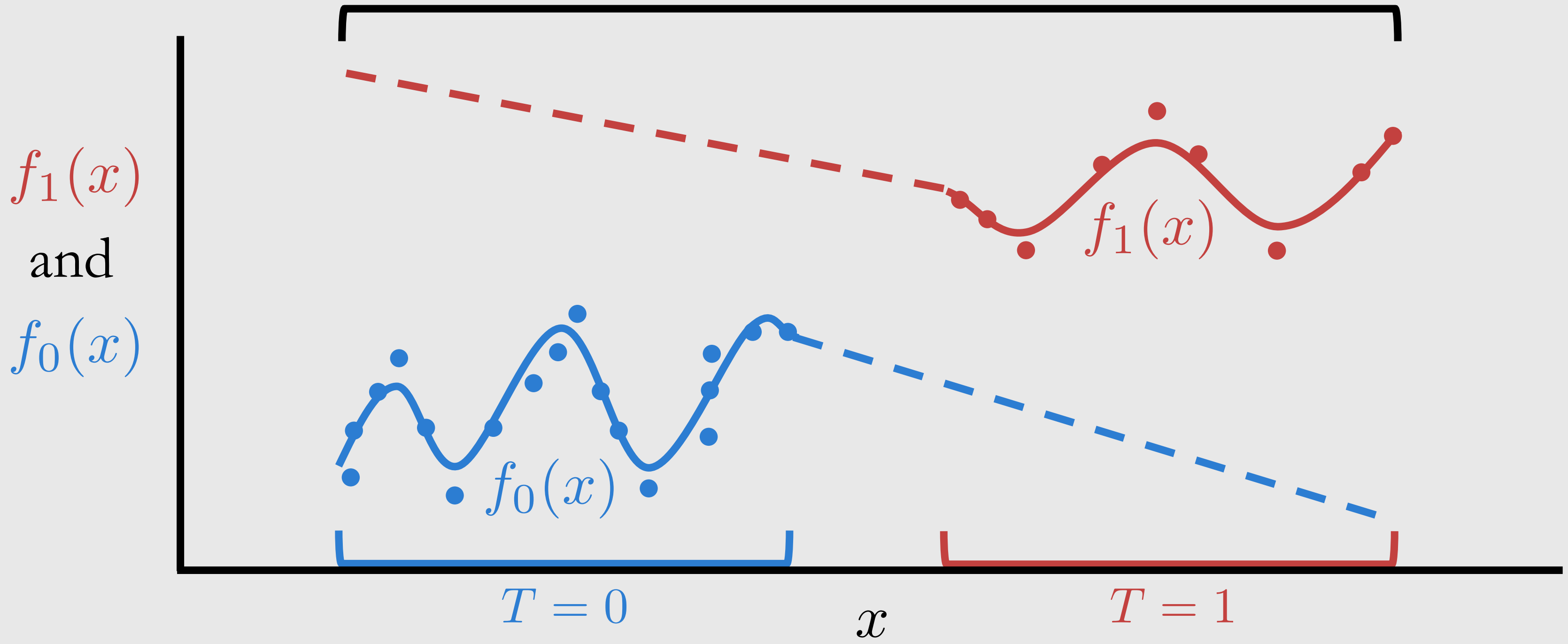
Extrapolation

Adjustment formula: $\sum_x \left(\underbrace{\mathbb{E}[Y \mid T=1, x]}_{\substack{\text{with} \\ f_1(x)}}^{\text{Model}} - \underbrace{\mathbb{E}[Y \mid T=0, x]}_{\substack{\text{with} \\ f_0(x)}}^{\text{Model}} \right)$



Extrapolation

Adjustment formula: $\sum_x \left(\underbrace{\overset{\text{Model}}{\mathbb{E}[Y \mid T=1, x]}}_{\text{with } f_1(x)} - \underbrace{\overset{\text{Model}}{\mathbb{E}[Y \mid T=0, x]}}_{\text{with } f_0(x)}\right)$

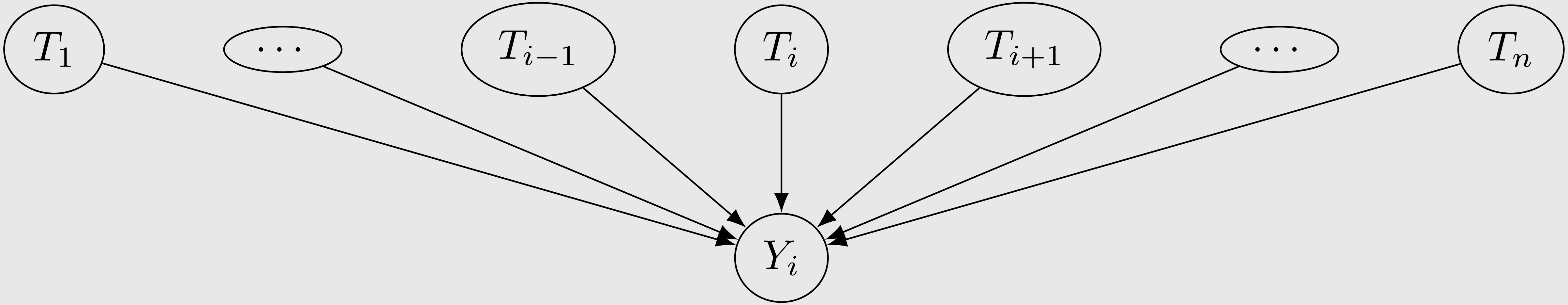


No interference

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

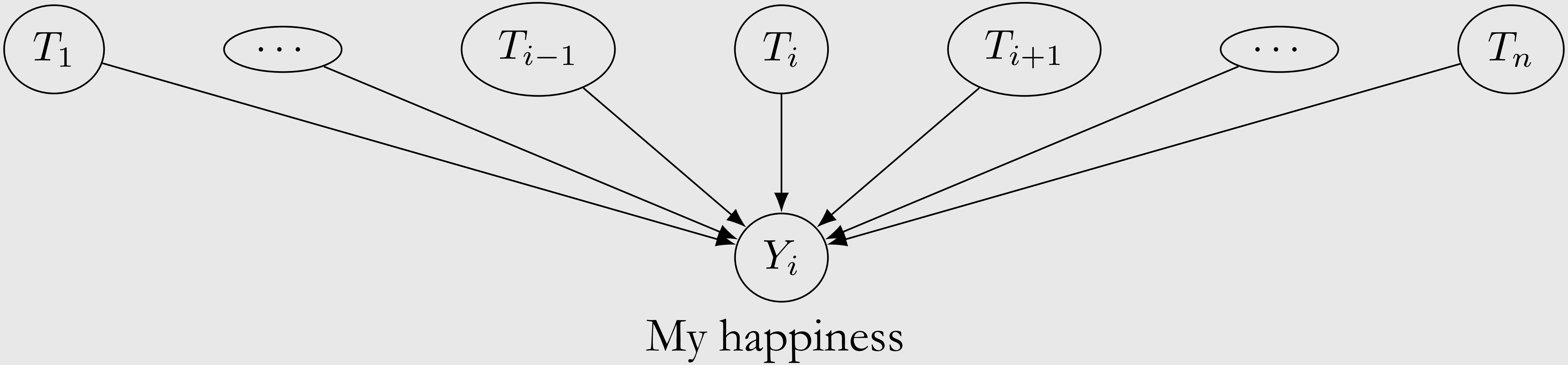
No interference

$$Y_i(t_1, \dots, t_{i-1}, t_i, t_{i+1}, \dots, t_n) = Y_i(t_i)$$



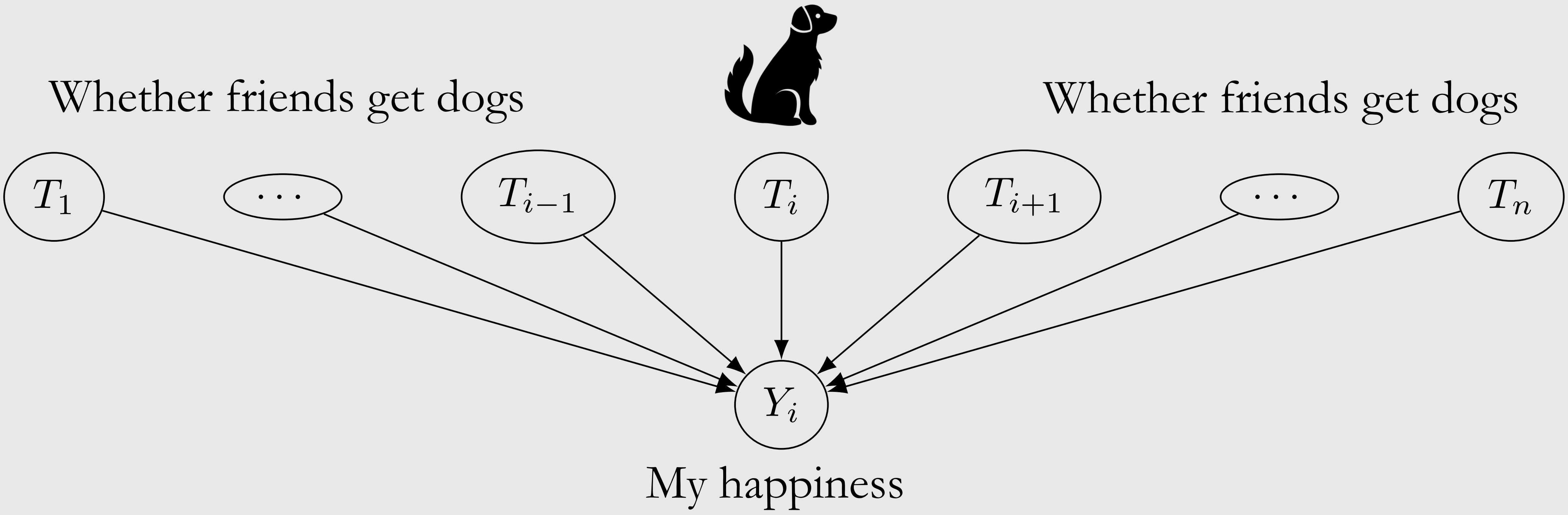
No interference

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



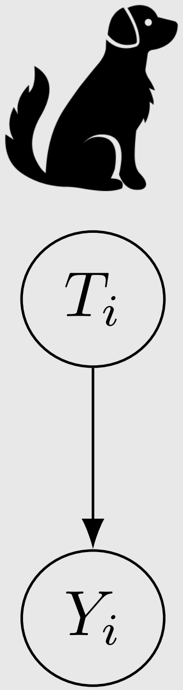
No interference

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



No interference

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



My happiness

Consistency: $T = t \implies Y = Y(t)$

Consistency: $T = t \implies Y = Y(t)$

$T = 1$
"I get a dog"

Consistency: $T = t \implies Y = Y(t)$

$T = 1$
"I get a dog"

$T = 0$
"I don't get a dog"

Consistency: $T = t \implies Y = Y(t)$

$T = 1$
"I get a dog"

$T = 0$
"I don't get a dog"

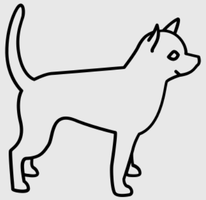
$(T = 1) \implies Y = 1$ (I'm happy)

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)

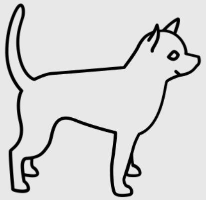


$(T = 1) \implies Y = 0$ (I'm not happy)

Consistency: $T = t \implies Y = Y(t)$

$T = 1$	$T = 0$
"I get a dog"	"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

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

Tying it all together

No interference

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

Tying it all together

No interference

$$
\begin{aligned}
\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] && \text{(linearity of expectation)} \\
&= \mathbb{E}_X\left[\mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]\right] \\
& && \text{(law of iterated expectations)}
\end{aligned}
$$

Tying it all together

No interference

$$
\begin{aligned}
\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] && \text{(linearity of expectation)} \\
&= \mathbb{E}_X\left[\mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]\right] \\
& && \text{(law of iterated expectations)} \\
&= \mathbb{E}_X\left[\mathbb{E}[Y(1) \mid T=1, X] - \mathbb{E}[Y(0) \mid T=0, X]\right] \\
& && \text{(unconfoundedness and positivity)}
\end{aligned}
$$

Tying it all together

No interference

$$\begin{aligned}
\mathbb{E}[Y(1) - Y(0)] &= \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] && \text{(linearity of expectation)} \\
&= \mathbb{E}_X\left[\mathbb{E}[Y(1) \mid X] - \mathbb{E}[Y(0) \mid X]\right] \\
& && \text{(law of iterated expectations)} \\
&= \mathbb{E}_X\left[\mathbb{E}[Y(1) \mid T = 1, X] - \mathbb{E}[Y(0) \mid T = 0, X]\right] \\
& && \text{(unconfoundedness and positivity)} \\
&= \mathbb{E}_X\left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right] && \text{(consistency)}
\end{aligned}$$

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

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. $\mathbb{E}_X\left[\mathbb{E}[Y \mid T=1, X] - \mathbb{E}[Y \mid T=0, X]\right]$)
- 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. $\mathbb{E}_X\left[\mathbb{E}[Y \mid T=1, X] - \mathbb{E}[Y \mid T=0, X]\right]$)
- 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. $\mathbb{E}_X \left[\mathbb{E}[Y \mid T = 1, X] - \mathbb{E}[Y \mid T = 0, X]\right]$)
- Estimate: approximation of some estimand, using data
- Estimation: process for getting from data + estimand to estimate

The Identification-Estimation Flowchart

$$\boxed{\text{Causal Estimand}} \xrightarrow{\text{Identification}} \boxed{\text{Statistical Estimand}} \xrightarrow{\text{Estimation}} \boxed{\text{Estimate}}$$

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

Estimation of ATE

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

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

Estimation:

Estimation of ATE

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

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

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

Estimation of ATE

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

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

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

Model (linear regression)

Estimation of ATE

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

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

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

Model (linear regression)

Estimate: 0.85

Estimation of ATE

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

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

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

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

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

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

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

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

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

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: $Y = \hat{\alpha} T + \hat{\beta} X$ $\qquad \hat{\alpha} = 0.85$

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

Using coefficient of linear regression

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

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

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

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

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

Using coefficient of linear regression

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

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

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

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: $\underline{Y = \alpha T + \beta X}$

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

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

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

$$Y_i(t) = \alpha t + \beta x_i$$

$$\begin{aligned} Y_i(1) - Y_i(0) = \alpha \cdot 1 + \beta x_i \\ -\alpha \cdot 0 - \beta x_i \end{aligned}$$

Using coefficient of linear regression

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

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

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

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

$$Y_i(t) = \alpha t + \beta x_i \qquad \begin{aligned} Y_i(1) - Y_i(0) &= \alpha \cdot 1 + \beta x_i \\ &\quad -\alpha \cdot 0 - \beta x_i = \alpha \end{aligned}$$

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

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

$$Y_i(t) = \alpha t + \beta x_i$$

$$\begin{aligned} Y_i(1) - Y_i(0) &= \alpha \cdot 1 + \beta x_i \\ &\quad -\alpha \cdot 0 - \beta x_i = \alpha \end{aligned}$$

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