# Potential Outcomes and RCTs

Alexander Quispe

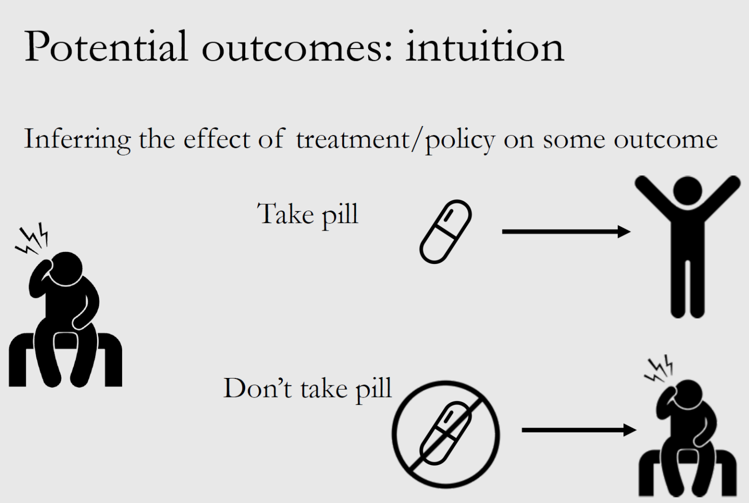
February 15, 2022

# What is the fundamental problem of causal inference?

This lecture notes are based on notes from Brady Neal ( Mila - Quebec AI Institute), Victor Chernozukhov (MIT) and Paul Goldsmith-Pinkham (Yale University)

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.

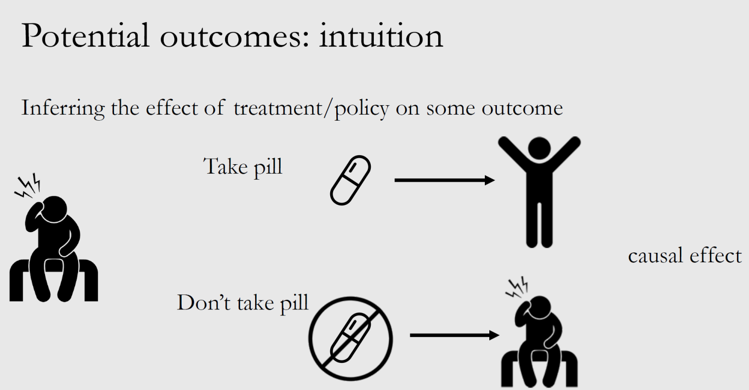


Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.

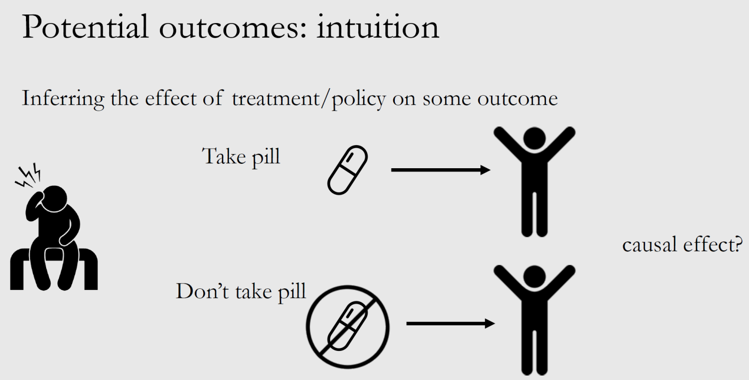


Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

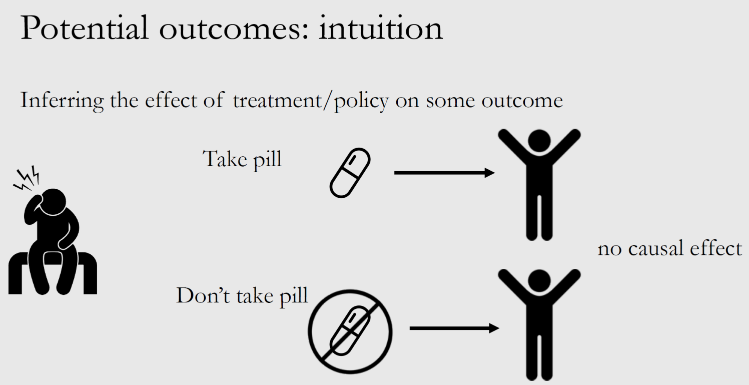
Take pill

Don’t take pill

causal effect?

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



Potential outcomes: intuition

Inferring the effect of treatment/policy on some outcome

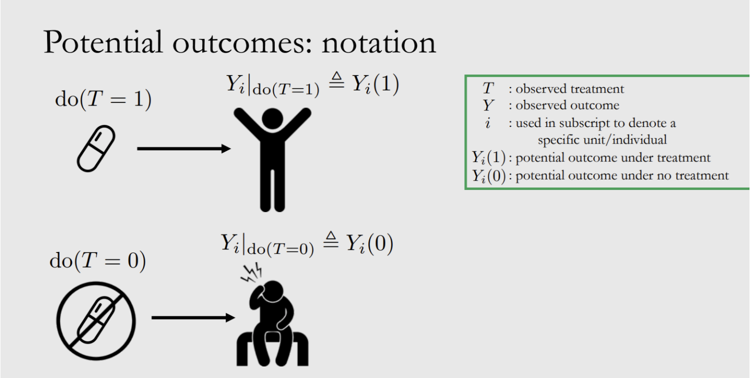
Take pill

Don't take pill

no causal effect

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## Potential outcomes: notation

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

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

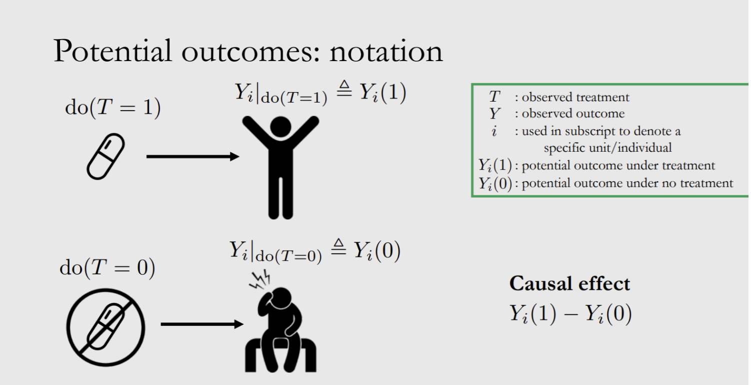
$\mathrm{do}(T = 0)$

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

| | |
|---|---|
| $T$ | : 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 |

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## Potential outcomes: notation

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

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

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

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

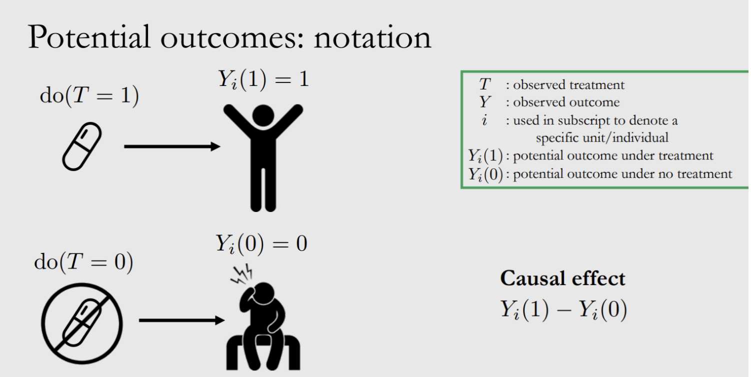
$T$ : observed treatment
$Y$ : observed outcome
$i$ : used in subscript to denote a specific unit/individual
$Y_i(1)$ : potential outcome under treatment
$Y_i(0)$ : potential outcome under no treatment

**Causal effect**

$Y_i(1) - Y_i(0)$

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## Potential outcomes: notation

do$(T = 1)$ $Y_i(1) = 1$

do$(T = 0)$ $Y_i(0) = 0$

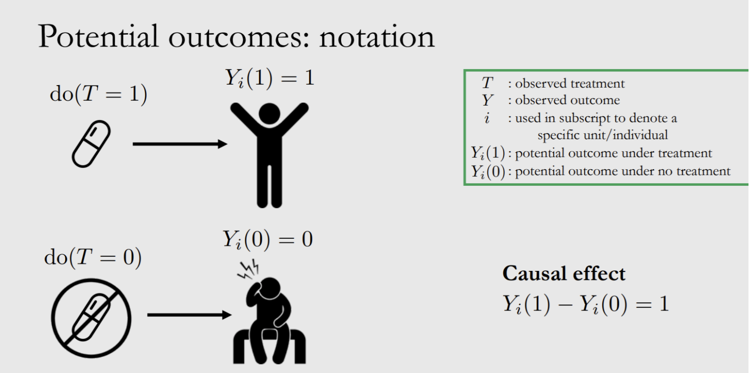
$T$ : observed treatment
$Y$ : observed outcome
$i$ : used in subscript to denote a specific unit/individual
$Y_i(1)$ : potential outcome under treatment
$Y_i(0)$ : potential outcome under no treatment

**Causal effect**

$Y_i(1) - Y_i(0)$

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## Potential outcomes: notation

do$(T = 1)$ $Y_i(1) = 1$

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

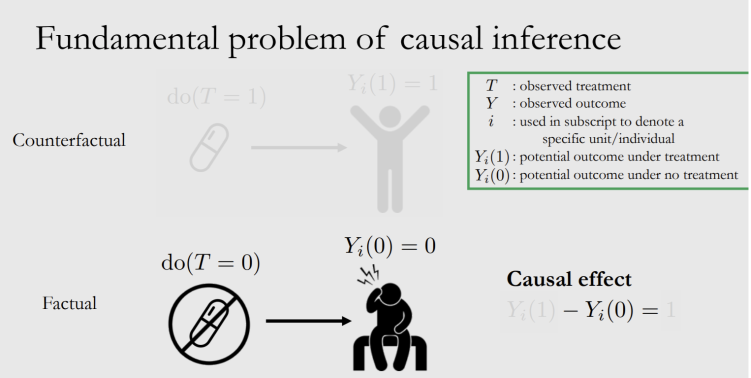
$Y_i(0)$ : potential outcome under no treatment

**Causal effect**

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

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



Fundamental problem of causal inference

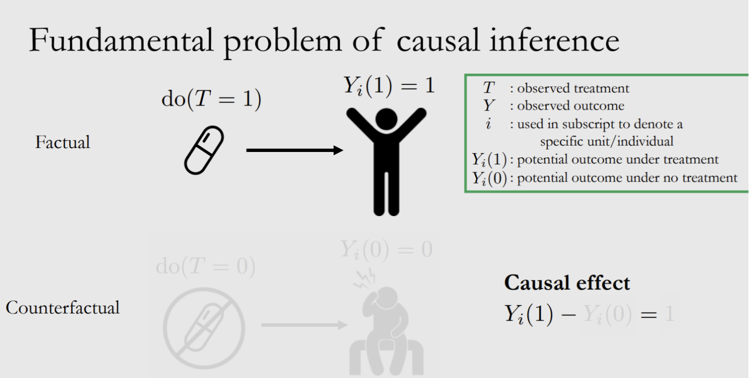
$T$ : observed treatment
$Y$ : observed outcome
$i$ : used in subscript to denote a specific unit/individual
$Y_i(1)$ : potential outcome under treatment
$Y_i(0)$ : potential outcome under no treatment

**Causal effect**

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

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



Fundamental problem of causal inference

Factual

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

$Y_i(1) = 1$

$T$ : 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

Counterfactual

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

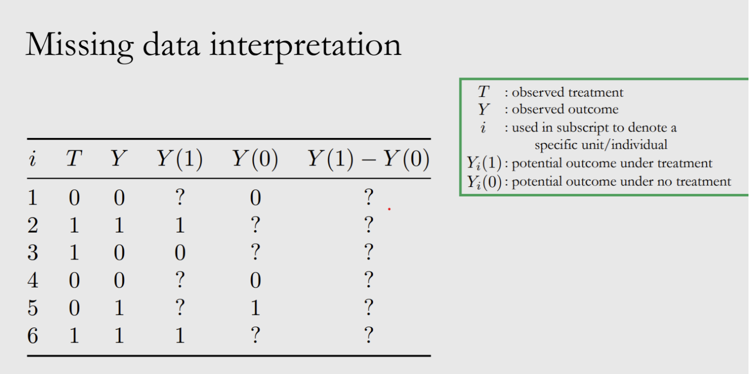
$Y_i(0) = 0$

**Causal effect**

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

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## 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

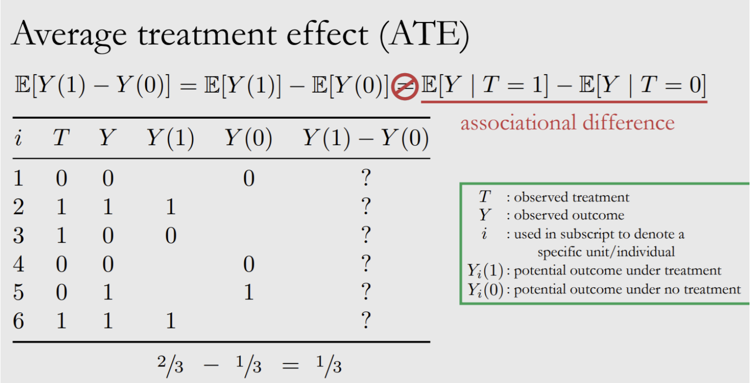
$i$ : used in subscript to denote a specific unit/individual

$Y_i(1)$ : potential outcome under treatment

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

# What is the fundamental problem of causal inference?

Figure: Lasso optimization with two coefficients.



## 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

# ATE IS NOT EQUAL TO AD

The problem with the observational study like the one in this contrived example is that the ”treatment status” is determined by the individual behavior which may be linked to the potential outcomes, causing the selection bias, namely the disagreement between AD and ATE.

# Causal estimands

- We will start with the Average Treatment Effect:
  - $\tau_{ATE} = \mathbb{E}(\tau_i) = \mathbb{E}(Y_i(1) - Y_i(0)) = \mathbb{E}(Y_i(1)) - \mathbb{E}(Y_i(0))$

- We **OBSERVE** the Asociation Difference:
  - $\tau_{AD} = \mathbb{E}(Y|D = 1) - \mathbb{E}(Y|D = 0)$

- This expression is defined over the full population, and includes individuals who may never recieve the treatment.
  - Average Treatment Effect on the Treated $\tau_{ATT} = \mathbb{E}(\tau_i|D_i = 1) = \mathbb{E}(Y_i(1) - Y_i(0)|D_i = 1) = \mathbb{E}(Y_i(1)|D_i = 1) - \mathbb{E}(Y_i(0)|D_i = 1)$
  - Estimated effect for individuals who *received* the treatment.
  - Note that one piece of this measure is purely observed data: $\mathbb{E}(Y_i(1)|D_i = 1)$

- Conditional Average Treatment Effect: $\tau_{CATE}(x) = \mathbb{E}(\tau_i|X_i = x) = \mathbb{E}(Y_i(1) - Y_i(0)|X_i = x)$ where $X_i$ is some additional characteristic.

# Under what conditions is the ATE identified?

**Strong Ignorability:** $D_i$ is *strongly ignorable* conditional on a vector $\mathbf{X}_i$ if

1. $(Y_i(0), Y_i(1)) \perp\!\!\!\perp D_i | \mathbf{X}_i$ - **Ignorability**
2. $0 < \Pr(D_i = 1 | X_i) < 1$ - **Positivity**

- The first condition asserts independence of the treatment from the “potential” outcomes
- The second condition asserts that there are both treated and untreated individuals for a given group.

# When could we not identify the ATE?

- Intuitively, we understand why we typically can't estimate a treatment effect

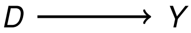
- Consider an unobservable variable, $U_i \in \{0, 1\}$ where $(Y_i(0), Y_i(1), D_i) \not\perp U_i$

- Simple example: when $E(D_i|U_i = 1) > E(D_i|U_i = 0)$ and $E(\tau_i|U_i = 1) > E(\tau_i|U_i = 0)$.

- In other word, there is a variable that influences both the potential outcomes and the choice of treatment.
  - In this case, estimating the counterfactual is contaminated by the variable $U_i$

# Identification through Directed Acylic Graphs (DAGs)

- We can encode the relationship between $D$ and $Y$ using an *arrow* in a graph. The direction emphasizes that $D$ causes $Y$, and not vice versa.



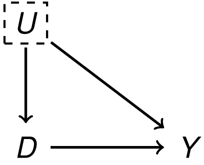
# Identification through Directed Acylic Graphs (DAGs)

- Now we can add the unobservable $U$, which drove identification concerns.

- In this case, $U$ is termed a *confounder*.

- What are the paths by which $D$ links to $Y$?
    - The standard direct effect $D \rightarrow Y$
    - The "Back-Door" path $D \leftarrow U \rightarrow Y$
    - Note that the back-door is *not* causal

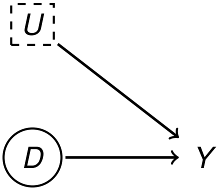
- **Effect of $D$ on $Y$ is not identified under this setup**



# The power of randomization

- The cleanest way to break the selection bias is through random assignment of treatment.

- Randomization is a powerful tool
  - E.g. a true randomized intervention such as randomly giving a treatment to half of a sample using a randomized process



# The power of randomization

Assumption (Random Assignment/Exogeneity)

$$D \perp Y(d) \textit{ and } 0 < P(D = 1) < 1 \tag{1}$$

Theorem (Randomization Removes Selection Bias)

$$E[Y \mid D = d] = E[Y(d) \mid D = d] = E[Y(d)] \tag{2}$$

$$\pi := E[Y \mid D = 1] - E[Y \mid D = 0] \tag{3}$$

$$\pi := E\{Y(1)\} - E\{Y(0)\} =: \delta \tag{4}$$

# The power of randomization

We can base our inference about ATE on linear regression.

$$Y(d) = d\alpha + X'\beta + \epsilon_d, \;\; E[\epsilon_d \mid X] = 0, \;\; \textit{where} \;\; X = (1, W) \tag{5}$$

We assume that covariates are recentered:

$$E\{W\} = 0, \;\; E[W \mid D = 1] = E[W \mid D = 0] \tag{6}$$

$$E\{Y(0)\} = \beta_1 \;\; E\{Y(1)\} = \beta_1 + \alpha, \;\; \epsilon := \epsilon_D \;\; E[\epsilon|D, X] = 0 \tag{7}$$

$$E[\epsilon(D, X)] = E[E[\epsilon_D \mid D, X](D, X)] = 0 \tag{8}$$

$$Y = \alpha D + X'\beta + \epsilon, \;\; \epsilon \perp (D, X) \tag{9}$$

$\alpha$ coincides with the regression/projection coefficient in the regression of $Y$ on $D$ and $X$.
We will be interested in ATE and relative ATE: $\alpha \;\; \alpha/\beta_1$