# Introduction, Potential Outcomes, and Assignments

Ryan T. Moore

American University

2024-08-20

# Table of contents I

# Welcome

# About Me

- Associate Prof of Government
  (American University)
- Associate Director, Center for Data Science
  (American University)
- Senior Social Scientist
  (The Lab @ DC)
- Fellow in Methodology
  (US Office of Evaluation Sciences: "OES")

# About Me

- Associate Prof of Government
  (American University)
- Associate Director, Center for Data Science
  (American University)
- Senior Social Scientist
  (The Lab @ DC)
- Fellow in Methodology
  (US Office of Evaluation Sciences: "OES")
- 2021-22: Visiting Fellow
  (Institut Barcelona d'Estudis Internacionals)

# About Me

- Associate Prof of Government (American University)
- Associate Director, Center for Data Science (American University)
- Senior Social Scientist (The Lab @ DC)
- Fellow in Methodology (US Office of Evaluation Sciences: "OES")
- 2021-22: Visiting Fellow (Institut Barcelona d'Estudis Internacionals)
- Research agenda: political methodology, causal inference, experimental design, experiments in public policy

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments
- Create and analyze variety of randomized designs, including for blocked, clustered, conjoint, list, and multiarm bandit experiments

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments
- Create and analyze variety of randomized designs, including for blocked, clustered, conjoint, list, and multiarm bandit experiments
- Relate experiments to regression quantities and estimate heterogeneous treatment effects

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments
- Create and analyze variety of randomized designs, including for blocked, clustered, conjoint, list, and multiarm bandit experiments
- Relate experiments to regression quantities and estimate heterogeneous treatment effects
- Estimate mediation effects and assess their sensitivity

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments
- Create and analyze variety of randomized designs, including for blocked, clustered, conjoint, list, and multiarm bandit experiments
- Relate experiments to regression quantities and estimate heterogeneous treatment effects
- Estimate mediation effects and assess their sensitivity

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments
- Create and analyze variety of randomized designs, including for blocked, clustered, conjoint, list, and multiarm bandit experiments
- Relate experiments to regression quantities and estimate heterogeneous treatment effects
- Estimate mediation effects and assess their sensitivity

- Materials at https://github.com/ryantmoore/causal-summer-waseda

# The Course: Learning Objectives

- Identify causal effects using the potential outcomes framework
- Perform design-based inference for randomized experiments
- Create and analyze variety of randomized designs, including for blocked, clustered, conjoint, list, and multiarm bandit experiments
- Relate experiments to regression quantities and estimate heterogeneous treatment effects
- Estimate mediation effects and assess their sensitivity

- Materials at https://github.com/ryantmoore/causal-summer-waseda
- (Walk through syllabus)

# The Course: Conduct

- Class discussion may be difficult!

# The Course: Conduct

- Class discussion may be difficult!
- I will ask you to quickly turn to someone nearby and *discuss* questions

# The Course: Conduct

- Class discussion may be difficult!
- I will ask you to quickly turn to someone nearby and *discuss* questions
- For example, soon, I will ask you whether a quantity has a valid *causal* interpretation

# The Course: Conduct

- Class discussion may be difficult!
- I will ask you to quickly turn to someone nearby and *discuss* questions
- For example, soon, I will ask you whether a quantity has a valid *causal* interpretation
- For example, in a day or two, I will ask you how to leave a casino with a lot of ¥!

# The Course: Conduct

- Class discussion may be difficult!
- I will ask you to quickly turn to someone nearby and *discuss* questions
- For example, soon, I will ask you whether a quantity has a valid *causal* interpretation
- For example, in a day or two, I will ask you how to leave a casino with a lot of ¥!
- Often $\approx$ 1-2 minutes

## The Course: Conduct

- Class discussion may be difficult!
- I will ask you to quickly turn to someone nearby and *discuss* questions
- For example, soon, I will ask you whether a quantity has a valid *causal* interpretation
- For example, in a day or two, I will ask you how to leave a casino with a lot of ¥!
- Often $\approx$ 1-2 minutes
- We learn what we don't quite understand when we process/talk about it

# What is Causal Inference?

# What is Causal Inference?

"What caused the terror attacks of 9/11?"

vs.

"What is the effect of foreign policy $X$ on domestic terror attacks?"

# What is Causal Inference?

"What caused the terror attacks of 9/11?"

vs.

"What is the effect of foreign policy $X$ on domestic terror attacks?"

"Causes of an effect"

vs.

"Effects of a cause"

# What is Causal Inference?

We will focus on "effects of a cause"

# What is Causal Inference?

We will focus on "effects of a cause", where the "cause" is well-defined.

# Example: Canvassing and Program Enrollment

## Example: Canvassing and Program Enrollment

"Canvassing": a systematic program where agents talk with residents (knock on doors) at the residents' home.

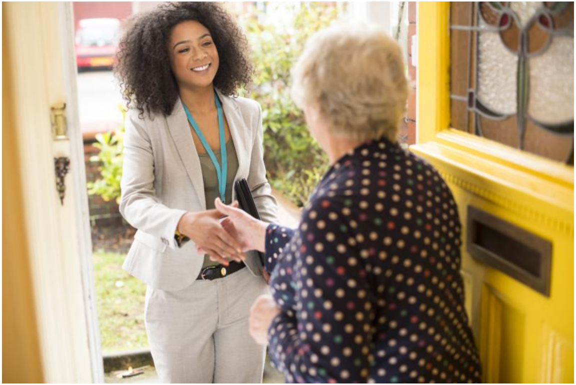

Figure 1: Credit: The Campaign Workshop

# Example: Canvassing and Program Enrollment

Suppose we ask, "Would a canvassing policy increase enrollment in a health insurance program?"

# Example: Canvassing and Program Enrollment

Suppose we ask, "Would a canvassing policy increase enrollment in a health insurance program?"

| Citizen | Canvassed? | Enrolled? |
|---|---|---|
| 1 | Yes | Yes |
| 2 | Yes | Yes |
| 3 | No | No |
| 4 | No | No |

# Example: Canvassing and Program Enrollment

Suppose we ask, "Would a canvassing policy increase enrollment in a health insurance program?"

| Citizen | Canvassed? | Enrolled? |
|---|---|---|
| 1 | Yes | Yes |
| 2 | Yes | Yes |
| 3 | No | No |
| 4 | No | No |

- What fraction enroll under canvassing vs. no canvassing?

# Example: Canvassing and Program Enrollment

Suppose we ask, "Would a canvassing policy increase enrollment in a health insurance program?"

| Citizen | Canvassed? | Enrolled? |
|---|---|---|
| 1 | Yes | Yes |
| 2 | Yes | Yes |
| 3 | No | No |
| 4 | No | No |

- What fraction enroll under canvassing vs. no canvassing?
- $\frac{2}{2} - \frac{0}{2} = 1$

# Example: Canvassing and Program Enrollment

Suppose we ask, “Would a canvassing policy increase enrollment in a health insurance program?”

| Citizen | Canvassed? | Enrolled? |
|---|---|---|
| 1 | Yes | Yes |
| 2 | Yes | Yes |
| 3 | No | No |
| 4 | No | No |

- What fraction enroll under canvassing vs. no canvassing?
- $\frac{2}{2} - \frac{0}{2} = 1$
- (For each person canvassed, expect 1 more enrollment.)

# Example: Canvassing and Program Enrollment

Suppose we ask, "Would a canvassing policy increase enrollment in a health insurance program?"

| Citizen | Canvassed? | Enrolled? |
|---|---|---|
| 1 | Yes | Yes |
| 2 | Yes | Yes |
| 3 | No | No |
| 4 | No | No |

- What fraction enroll under canvassing vs. no canvassing?
- $\frac{2}{2} - \frac{0}{2} = 1$
- (For each person canvassed, expect 1 more enrollment.)

- Did the policy "work" (cause more enrollment)?

But, is it causal?

# Motivating Example: Canvassing and Enrollment

But, is it causal?

What do we really want to know?

## Motivating Example: Canvassing and Enrollment

But, is it causal?

What do we really want to know?

Would canvassing actually *change* anyone's enrollment?

But, is it causal?

What do we really want to know?

Would canvassing actually *change* anyone's enrollment?

What would have happened under *other* conditions?

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | | | Yes |
| 2 | Yes | | | Yes |
| 3 | No | | | No |
| 4 | No | | | No |

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | | Yes |
| 2 | Yes | | | Yes |
| 3 | No | | | No |
| 4 | No | | | No |

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | | Yes |
| 2 | Yes | Yes | | Yes |
| 3 | No | | | No |
| 4 | No | | | No |

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | | Yes |
| 2 | Yes | Yes | | Yes |
| 3 | No | | No | No |
| 4 | No | | | No |

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | | Yes |
| 2 | Yes | Yes | | Yes |
| 3 | No | | No | No |
| 4 | No | | No | No |

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

What is the true causal effect of canvassing?

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

What is the true causal effect of canvassing?
What fraction enroll under canvass vs. no canvass?

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

What is the true causal effect of canvassing?
What fraction enroll under canvass vs. no canvass?
What fraction change enrollment due to canvass?

# Example: Canvassing and Enrollment

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

What is the true causal effect of canvassing?
What fraction enroll under canvass vs. no canvass?
What fraction change enrollment due to canvass?
$\frac{3}{4} - \frac{1}{4} = \frac{1}{2}$

# Example: Canvassing and Enrollment

Empirical data consistent with *different unobserved* outcomes:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (Yes) | No | No |

# Example: Canvassing and Enrollment

Empirical data consistent with *different unobserved* outcomes:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (Yes) | No | No |

What is the true causal effect of canvass?
What fraction enroll under canvass vs. no canvass?

# Example: Canvassing and Enrollment

Empirical data consistent with *different unobserved* outcomes:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (Yes) | No | No |

What is the true causal effect of canvass?
What fraction enroll under canvass vs. no canvass?
$\frac{4}{4} - \frac{1}{4} = \frac{3}{4}$

## Example: Canvassing and Enrollment

Well ... how do we know which?

# Example: Canvassing and Enrollment

Well ... how do we know which?

We can never know.

Can we know for one person?

Can we know for one person?

We can never know.

But I have some ideas.

But I have some ideas.

We could not canvass, then canvass later.

But I have some ideas.

We could not canvass, then canvass later.

We can never know.

We can never observe both "Canvassed" and "Not Canvassed" for a unit.

We can never observe both "Canvassed" and "Not Canvassed" for a unit.

We can never observe both *potential outcomes.*

We can never observe both "Canvassed" and "Not Canvassed" for a unit.

We can never observe both *potential outcomes.*

We can never observe both the factual and the counterfactual.

We can never observe both "Canvassed" and "Not Canvassed" for a unit.

We can never observe both *potential outcomes.*

We can never observe both the factual and the counterfactual.

We can never know.

We can never observe more than one *potential outcome* for a given unit.

So, how can we get a *causal* estimate?

So, how can we get a *causal* estimate?

We infer missing potential outcomes.

## Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group $\neq$ "No Canvass" group

# Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group $\neq$ "No Canvass" group
- The *potential outcomes* help predict whether Canvass!

# Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group $\neq$ "No Canvass" group
- The *potential outcomes* help predict whether Canvass!

# Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group $\neq$ "No Canvass" group
- The *potential outcomes* help predict whether Canvass!

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

# Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group $\neq$ "No Canvass" group
- The *potential outcomes* help predict whether Canvass!

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

Knowing whether would "enroll under canvass" *predicts* whether canvassed!

# Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group $\neq$ "No Canvass" group
- The *potential outcomes* help predict whether Canvass!

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

Knowing whether would "enroll under canvass" *predicts* whether canvassed!

$Pr(\text{Canvass} = \text{Yes}|\text{Would if Canvassed} = \text{Yes}) = \frac{2}{3}$

## Why didn't we recover truth?

The problem with our naive estimate of effect:

- "Canvass" group ≠ "No Canvass" group
- The *potential outcomes* help predict whether Canvass!

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | No | No |
| 4 | No | (No) | No | No |

Knowing whether would "enroll under canvass" *predicts* whether canvassed!

$Pr(\text{Canvass} = \text{Yes}|\text{Would if Canvassed} = \text{Yes}) = \frac{2}{3}$

But $Pr(\text{Canvass} = \text{Yes}|\text{Would if Canvassed} = \text{No}) = 0$

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

- What is the *true* causal effect of canvass?

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

- What is the *true* causal effect of canvass?
- $\frac{4}{4} - \frac{2}{4} = \frac{2}{4} = \frac{1}{2}$

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

- What is the *true* causal effect of canvass?
- $\frac{4}{4} - \frac{2}{4} = \frac{2}{4} = \frac{1}{2}$
- What would we *observe* as the effect of canvass?

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

- What is the *true* causal effect of canvass?
- $\frac{4}{4} - \frac{2}{4} = \frac{2}{4} = \frac{1}{2}$
- What would we *observe* as the effect of canvass?
- $\frac{2}{2} - \frac{1}{2} = \frac{1}{2}$

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

- What is the *true* causal effect of canvass?
- $\frac{4}{4} - \frac{2}{4} = \frac{2}{4} = \frac{1}{2}$
- What would we *observe* as the effect of canvass?
- $\frac{2}{2} - \frac{1}{2} = \frac{1}{2}$

# When comparing two groups **does** recover truth

Here, potential outcomes do **not** help predict Canvass:

| Citizen | Canvass? | Would Enroll if Canvass? | Would Enroll if No Canvass? | Enroll |
|---|---|---|---|---|
| 1 | Yes | Yes | (Yes) | Yes |
| 2 | Yes | Yes | (No) | Yes |
| 3 | No | (Yes) | Yes | Yes |
| 4 | No | (Yes) | No | No |

Knowing whether enroll if canvass not predictive.

- What is the *true* causal effect of canvass?
- $\frac{4}{4} - \frac{2}{4} = \frac{2}{4} = \frac{1}{2}$
- What would we *observe* as the effect of canvass?
- $\frac{2}{2} - \frac{1}{2} = \frac{1}{2}$

Good! ☺

# When does comparing groups recover truth?

Neither *potential outcome* should help predict treatment/intervention.

# When does comparing groups recover truth?

Neither *potential outcome* should help predict treatment/intervention.

True underlying responses in "Canvass" group =
True underlying responses in "No Canvass" group

(Note: *observed* outcomes can predict treatment.)

(Note: *observed* outcomes can predict treatment.)

That's the goal of an intervention!

(Note: *observed* outcomes can predict treatment.)

That's the goal of an intervention!

- Aspirin $\implies$ headache!
- Canvass $\implies$ turnout!
- Insurance $\implies$ health spending!
- CBT $\implies$ remission!

(Note: *observed* outcomes can predict treatment.)

That's the goal of an intervention!

- Aspirin $\implies$ headache!
- Canvass $\implies$ turnout!
- Insurance $\implies$ health spending!
- CBT $\implies$ remission!

But, full schedule of true underlying responses should **not** predict treatment.

# When does comparing groups recover truth?

How to ensure potential outcomes won't predict treatment?

# When does comparing groups recover truth?

How to ensure potential outcomes won't predict treatment?

How to *assign* treatment so it won't predict potential outcomes?

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)
- Let Party decide whom gets Canvass

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)
- Let Party decide whom gets Canvass
  (But, Party will only Canvass those it will affect)

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)
- Let Party decide whom gets Canvass
  (But, Party will only Canvass those it will affect)
  (We will estimate *too large* effect)

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)
- Let Party decide whom gets Canvass
  (But, Party will only Canvass those it will affect)
  (We will estimate *too large* effect)
- What if you *randomly* select whom gets Canvass?

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)
- Let Party decide whom gets Canvass
  (But, Party will only Canvass those it will affect)
  (We will estimate *too large* effect)
- What if you *randomly* select whom gets Canvass?
  ("Citizen was Canvassed" won't help guess pot. out.)

# When does comparing groups recover truth?

Possible *assignment mechanisms*:

- Let Citizens decide whether to get Canvass
  (But, those who choose Canvass will Enroll anyway)
  (We will estimate *too large* effect)
- Let Party decide whom gets Canvass
  (But, Party will only Canvass those it will affect)
  (We will estimate *too large* effect)
- What if you *randomly* select whom gets Canvass?
  ("Citizen was Canvassed" won't help guess pot. out.)
  (We will estimate *correct* effect)

## The Potential Outcomes Model

# What is Causal Inference?

- "Effects of causes"

# What is Causal Inference?

- "Effects of causes"

# What is Causal Inference?

- "Effects of causes"

  *The objective is to determine for some population of units ...the 'typical' causal effect of the [treatment vs. control conditions] on a dependent variable* $Y$*.*
  *—Rubin (1974)*

# What is Causal Inference?

- "Effects of causes"

  *The objective is to determine for some population of units …the 'typical' causal effect of the [treatment vs. control conditions] on a dependent variable* $Y$*.*
  *—Rubin (1974)*

- A "causal effect" is a comparative statement

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"

## What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"
- Timing of treatment: outcomes vs. covariates

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"
- Timing of treatment: outcomes vs. covariates
- Exclusivity of treatment (to unit)

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"
- Timing of treatment: outcomes vs. covariates
- Exclusivity of treatment (to unit)
- Excludability (assignment effect only via Tr)

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"
- Timing of treatment: outcomes vs. covariates
- Exclusivity of treatment (to unit)
- Excludability (assignment effect only via Tr)
- One study, one causal effect

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"
- Timing of treatment: outcomes vs. covariates
- Exclusivity of treatment (to unit)
- Excludability (assignment effect only via Tr)
- One study, one causal effect

# What is Causal Inference?

- Central definition for causal inference:
  "a *well-defined* treatment"
- Units must be "potentially exposable" to treatment
- "No causation without manipulation"
- Timing of treatment: outcomes vs. covariates
- Exclusivity of treatment (to unit)
- Excludability (assignment effect only via Tr)
- One study, one causal effect (roughly)

# The Potential Outcomes Model: Ideas

- *Unit*: a particular case at a point in time

## The Potential Outcomes Model: Ideas

- *Unit*: a particular case at a point in time
- *Treatment*: (putatively) causal variable of interest

# The Potential Outcomes Model: Ideas

- *Unit*: a particular case at a point in time
- *Treatment*: (putatively) causal variable of interest
- *Potential outcome*: outcome that would obtain if unit were to receive tr condition

# The Potential Outcomes Model: Ideas

- *Unit*: a particular case at a point in time
- *Treatment*: (putatively) causal variable of interest
- *Potential outcome*: outcome that would obtain if unit were to receive tr condition
- *Assignment mechanism*: means by which units come to be sorted into conditions

Stable Unit Treatment Value Assumption (SUTVA)

- No versions of the treatment, varying in effectiveness

# The Potential Outcomes Model: Ideas

Stable Unit Treatment Value Assumption (SUTVA)

- No versions of the treatment, varying in effectiveness
- No interference between units

# The Potential Outcomes Model: Notation

- Units: index $i \in \{1, \ldots, 2n\}$

# The Potential Outcomes Model: Notation

- Units: index $i \in \{1, \dots, 2n\}$
- Binary treatment: $T_i \in \{0, 1\}$

# The Potential Outcomes Model: Notation

- Units: index $i \in \{1, \dots, 2n\}$
- Binary treatment: $T_i \in \{0, 1\}$

- Pot outcome for $i$ under $T_i = 1$: $Y_{i1}$ or $Y_i(1)$

# The Potential Outcomes Model: Notation

- Units: index $i \in \{1, \dots, 2n\}$
- Binary treatment: $T_i \in \{0, 1\}$

- Pot outcome for $i$ under $T_i = 1$: $Y_{i1}$ or $Y_i(1)$
- Pot outcome for $i$ under $T_i = 0$: $Y_{i0}$ or $Y_i(0)$

## The Potential Outcomes Model: Notation

- Units: index $i \in \{1, \ldots, 2n\}$
- Binary treatment: $T_i \in \{0, 1\}$

- Pot outcome for $i$ under $T_i = 1$: $Y_{i1}$ or $Y_i(1)$
- Pot outcome for $i$ under $T_i = 0$: $Y_{i0}$ or $Y_i(0)$

- For vector of $Y_{i1} \quad \forall i$, write $Y_1$

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

(sometimes written $Y_i^{obs}$)

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

  (sometimes written $Y_i^{obs}$)
- If $T_i = 1$,

$$Y_i = Y_i(1) \cdot 1 + Y_i(0)(1 - 1) = Y_i(1)$$

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

  (sometimes written $Y_i^{obs}$)
- If $T_i = 1$,

$$Y_i = Y_i(1) \cdot 1 + Y_i(0)(1 - 1) = Y_i(1)$$

- If $T_i = 0$,

$$Y_i = Y_i(1) \cdot 0 + Y_i(0)(1 - 0) = Y_i(0)$$

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

  (sometimes written $Y_i^{obs}$)
- If $T_i = 1$,

$$Y_i = Y_i(1) \cdot 1 + Y_i(0)(1 - 1) = Y_i(1)$$

- If $T_i = 0$,

$$Y_i = Y_i(1) \cdot 0 + Y_i(0)(1 - 0) = Y_i(0)$$

# The Potential Outcomes Model: Notation

Observed outcomes

- The observed outcome:

$$Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$$

  (sometimes written $Y_i^{obs}$)
- If $T_i = 1$,

$$Y_i = Y_i(1) \cdot 1 + Y_i(0)(1 - 1) = Y_i(1)$$

- If $T_i = 0$,

$$Y_i = Y_i(1) \cdot 0 + Y_i(0)(1 - 0) = Y_i(0)$$

The assignment mechanism *selects* which potential outcome we observe.

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)
- *estimator*: statistic used to approximate/guess parameter ($\hat{\theta}$, $\hat{\beta}$)

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)
- *estimator*: statistic used to approximate/guess parameter ($\hat{\theta}$, $\hat{\beta}$)
- *estimand*: the parameter an estimator attempts to estimate

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)
- *estimator*: statistic used to approximate/guess parameter ($\hat{\theta}$, $\hat{\beta}$)
- *estimand*: the parameter an estimator attempts to estimate
- *estimate*: application of an estimator func to some obs data

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)
- *estimator*: statistic used to approximate/guess parameter ($\hat{\theta}$, $\hat{\beta}$)
- *estimand*: the parameter an estimator attempts to estimate
- *estimate*: application of an estimator func to some obs data

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)
- *estimator*: statistic used to approximate/guess parameter ($\hat{\theta}$, $\hat{\beta}$)
- *estimand*: the parameter an estimator attempts to estimate
- *estimate*: application of an estimator func to some obs data

"The sample statistic $\bar{x}$ is an estimator of true mean param $\mu$".

# The Potential Outcomes Model: Estimands

Statistical language

- *parameter*: unknown numeric value characterizing feature of prob model (Greek; $\theta$, $\beta$)
- *statistic*: quantity calculable from observed data. A function. (Roman)
- *estimator*: statistic used to approximate/guess parameter ($\hat{\theta}$, $\hat{\beta}$)
- *estimand*: the parameter an estimator attempts to estimate
- *estimate*: application of an estimator func to some obs data

"The sample statistic $\bar{x}$ is an estimator of true mean param $\mu$".

$\mu$ is my estimand. 5.1 is my estimate.

# The Potential Outcomes Model: Estimands

For $i$, individual treatment effect

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

- True treatment effect?

# The Potential Outcomes Model: Estimands

For $i$, individual treatment effect

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

- True treatment effect?
- Estimate?

# The Potential Outcomes Model: Estimands

For $i$, individual treatment effect

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

- True treatment effect?
- Estimate?
- Observable?

# The Potential Outcomes Model: Estimands

For $i$, individual treatment effect

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

- True treatment effect?
- Estimate?
- Observable?

# The Potential Outcomes Model: Estimands

For $i$, individual treatment effect

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

- True treatment effect?
- Estimate?
- Observable?

We can never know.

# The Potential Outcomes Model: Estimands

- Average treatment effect

$$ATE = \overline{Y_1 - Y_0} = \overline{Y_1} - \overline{Y_0}$$

# The Potential Outcomes Model: Estimands

- Average treatment effect

$$ATE = \overline{Y_1 - Y_0} = \overline{Y_1} - \overline{Y_0}$$

- Average treatment effect for the treated

$$ATT = (\overline{Y_1}|T = 1) - (\overline{Y_0}|T = 1)$$

# The Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
\overline{\tau} = \overline{TE} = ATE \quad &\equiv \quad \overline{Y_1 - Y_0} \\
&= \quad \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i1} - Y_{i0})
\end{aligned}
$$

# The Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
\overline{\tau} = \overline{TE} = ATE &\equiv \overline{Y_1 - Y_0} \\
&= \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i1} - Y_{i0}) \\
&= \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i1}) - \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i0})
\end{aligned}
$$

# The Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
\overline{\tau} = \overline{TE} = ATE &\equiv \overline{Y_1 - Y_0} \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1} - Y_{i0}) \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1}) - \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i0}) \\
&= \overline{Y_1} - \overline{Y_0}
\end{aligned}
$$

# The Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
\overline{\tau} = \overline{TE} = ATE &\equiv \overline{Y_1 - Y_0} \\
&= \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i1} - Y_{i0}) \\
&= \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i1}) - \frac{1}{2n} \sum_{i=1}^{2n} (Y_{i0}) \\
&= \overline{Y_1} - \overline{Y_0}
\end{aligned}
$$

- True?

# The Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
\overline{\tau} = \overline{TE} = ATE &\equiv \overline{Y_1 - Y_0} \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1} - Y_{i0}) \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1}) - \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i0}) \\
&= \overline{Y_1} - \overline{Y_0}
\end{aligned}
$$

- True?
- Estimated?

# The Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
\overline{\tau} = \overline{TE} = ATE &\equiv \overline{Y_1 - Y_0} \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1} - Y_{i0}) \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1}) - \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i0}) \\
&= \overline{Y_1} - \overline{Y_0}
\end{aligned}
$$

- True?
- Estimated?
- Observable?

# The Potential Outcomes Model: Estimands

Statistical Independence

## The Potential Outcomes Model: Estimands

Statistical Independence

Define

$$Y \perp\!\!\!\perp X$$

as

$$Pr(Y|X) = Pr(Y)$$

# The Potential Outcomes Model: Estimands

Statistical Independence

Define

$$Y \perp\!\!\!\perp X$$

as

$$Pr(Y|X) = Pr(Y)$$

(Knowing $X$ doesn't change probability of $Y$)

# The Potential Outcomes Model: Estimands

Something we **can** calculate:

If we know $(Y_1, Y_0)$ indep of $T$

# The Potential Outcomes Model: Estimands

Something we **can** calculate:

If we know $(Y_1, Y_0)$ indep of $T$
(i.e., knowing true underlying response does **not** indicate $T$)

# The Potential Outcomes Model: Estimands

Something we **can** calculate:

If we know $(Y_1, Y_0)$ indep of $T$
(i.e., knowing true underlying response does **not** indicate $T$)

Then,

# The Potential Outcomes Model: Estimands

Something we **can** calculate:

If we know $(Y_1, Y_0)$ indep of $T$
(i.e., knowing true underlying response does **not** indicate $T$)

Then,

$$\begin{aligned} \overline{Y_1} &= (\overline{Y_1}|T=1) \\ \overline{Y_0} &= (\overline{Y_0}|T=0) \end{aligned}$$

# The Potential Outcomes Model: Estimands

Something we **can** calculate:

If we know $(Y_1, Y_0)$ indep of $T$
(i.e., knowing true underlying response does **not** indicate $T$)

Then,

$$\begin{aligned} \overline{Y_1} &= (\overline{Y_1}|T=1) \\ \overline{Y_0} &= (\overline{Y_0}|T=0) \end{aligned}$$

Then, can substitute

$$\begin{aligned} ATE &= \overline{Y_1} - \overline{Y_0} \\ &= (\overline{Y_1}|T=1) - (\overline{Y_0}|T=0) \end{aligned}$$

# The Potential Outcomes Model: Estimands

Something we **can** calculate:

If we know $(Y_1, Y_0)$ indep of $T$
(i.e., knowing true underlying response does **not** indicate $T$)

Then,

$$
\begin{aligned}
\overline{Y_1} &= (\overline{Y_1} | T = 1) \\
\overline{Y_0} &= (\overline{Y_0} | T = 0)
\end{aligned}
$$

Then, can substitute

$$
\begin{aligned}
ATE &= \overline{Y_1} - \overline{Y_0} \\
&= (\overline{Y_1} | T = 1) - (\overline{Y_0} | T = 0)
\end{aligned}
$$

Observed diff in Tr and Co group means gives $\overline{TE}$!

# Why Experiments?

Random assignment of treatment promotes

$$(Y_1, Y_0) \perp\!\!\!\perp T$$

# Why Experiments?

Random assignment of treatment promotes

$$(Y_1, Y_0) \perp\!\!\!\perp T$$

When $(Y_1, Y_0)$ and $T$ independent, we can substitute things we know

$$(\overline{Y_1}|T = 1) \text{ and } (\overline{Y_0}|T = 0)$$

# Why Experiments?

Random assignment of treatment promotes

$$(Y_1, Y_0) \perp\!\!\!\perp T$$

When $(Y_1, Y_0)$ and $T$ independent, we can substitute things we know

$$(\overline{Y_1}|T = 1) \text{ and } (\overline{Y_0}|T = 0)$$

for things we can *never* know

$$\overline{Y_1} \text{ and } \overline{Y_0}$$

## Common Assumptions, Null Hyp's in Causal Inference

- Constant effect:

$$\tau_i = Y_{i1} - Y_{i0} = \tau \quad \forall i$$

# Common Assumptions, Null Hyp's in Causal Inference

- Constant effect:

$$\tau_i = Y_{i1} - Y_{i0} = \tau \quad \forall i$$

- Null hypothesis of no average effect:

$$ATE = \overline{\tau} = 0$$

# Common Assumptions, Null Hyp's in Causal Inference

- Constant effect:

$$\tau_i = Y_{i1} - Y_{i0} = \tau \quad \forall i$$

- Null hypothesis of no average effect:

$$ATE = \overline{\tau} = 0$$

- Sharp null hypothesis of no effect:

$$\tau_i = 0$$

# Compliance with Treatment Assignment

Sometimes, units don't follow assignment!

# Compliance with Treatment Assignment

Sometimes, units don't follow assignment!

| Assigned Tr | Assigned Co | Type |
|---|---|---|
| Tr | Co | Complier |

# Compliance with Treatment Assignment

Sometimes, units don't follow assignment!

| Assigned Tr | Assigned Co | Type |
|---|---|---|
| Tr | Co | Complier |
| Tr | Tr | Always-taker |

# Compliance with Treatment Assignment

Sometimes, units don't follow assignment!

| Assigned Tr | Assigned Co | Type |
|---|---|---|
| Tr | Co | Complier |
| Tr | Tr | Always-taker |
| Co | Co | Never-taker |

# Compliance with Treatment Assignment

Sometimes, units don't follow assignment!

| Assigned Tr | Assigned Co | Type |
|---|---|---|
| Tr | Co | Complier |
| Tr | Tr | Always-taker |
| Co | Co | Never-taker |
| Co | Tr | Defier |

# Compliance with Treatment Assignment

Of course, we cannot observe $i$ assigned both Tr and Co.

# Compliance with Treatment Assignment

Of course, we cannot observe $i$ assigned both Tr and Co.

| | | Treatment Taken | |
|---|---|---|---|
| | | Treatment | Control |
| Tr Assigned | Tr | Complier/Always | Never/Defier |
| | Co | Always/Defier | Complier/Never |

# Common Estimates under Noncompliance

Let $T_i$ be treatment *assigned*, $D_i$ be treatment *received*.

# Common Estimates under Noncompliance

Let $T_i$ be treatment *assigned*, $D_i$ be treatment *received*.

- Intent-to-treat effect (as-assigned)

$$
\begin{aligned}
ITT &= (\overline{Y_1}|T=1) - (\overline{Y_0}|T=0) \\
&= (\overline{Y_1}|T=1, D(T=1)) - (\overline{Y_0}|T=0, D(T=0))
\end{aligned}
$$

# Common Estimates under Noncompliance

Let $T_i$ be treatment *assigned*, $D_i$ be treatment *received*.

- Intent-to-treat effect (as-assigned)

$$\begin{aligned} ITT &= (\overline{Y_1}|T=1) - (\overline{Y_0}|T=0) \\ &= (\overline{Y_1}|T=1, D(T=1)) - (\overline{Y_0}|T=0, D(T=0)) \end{aligned}$$

- As-treated effect

$$ASTRE = (\overline{Y_1}|D=1) - (\overline{Y_0}|D=0)$$

# Unbiasedness

$$E(\widehat{\theta}) = \theta$$

# Unbiasedness

$$E(\hat{\theta}) = \theta$$

- $E\left[\widehat{\overline{Y_1 - Y_0}}\right] = \overline{Y_1 - Y_0}$

# Unbiasedness

$$E(\hat{\theta}) = \theta$$

- $E\left[\widehat{\overline{Y_1 - Y_0}}\right] = \overline{Y_1 - Y_0}$
- $E\left[\widehat{\beta_1}\right] = \beta_1$

# Unbiasedness

What estimator used to estimate $\overline{Y_1 - Y_0}$?

# Unbiasedness

What estimator used to estimate $\overline{Y_1 - Y_0}$?

The difference-in-means estimator.

# Unbiasedness

What estimator used to estimate $\overline{Y_1 - Y_0}$?

The difference-in-means estimator.

$$\widehat{\overline{Y_1 - Y_0}} = (Y_1|T_i = 1) - (Y_0|T_i = 0)$$

## Unbiasedness

What estimator used to estimate $\overline{Y_1 - Y_0}$?

The difference-in-means estimator.

$$\widehat{\overline{Y_1 - Y_0}} = (Y_1 | T_i = 1) - (Y_0 | T_i = 0)$$

Is it a "good" estimator?

## Unbiasedness

The difference-in-means estimator is *unbiased* for the true average treatment effect.

# Unbiasedness

The difference-in-means estimator is *unbiased* for the true average treatment effect.

(See `01-unbiased.R`)

# Potential Outcomes Model: Estimands

- Individual TE

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

- Average treatment effect

$$ATE = E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

- Average treatment effect for the treated

$$ATT = E(Y_1|T = 1) - E(Y_0|T = 1)$$

# Potential Outcomes Model: Estimands

The average treatment effect (ATE):

$$
\begin{aligned}
E(Y_1 - Y_0) &= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1} - Y_{i0}) \\
&= \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i1}) - \frac{1}{2n}\sum_{i=1}^{2n}(Y_{i0}) \\
&= E(Y_1) - E(Y_0)
\end{aligned}
$$

# Potential Outcomes Model: Estimands

- If we know $(Y_1, Y_0)$ indep of $T$
- Then,

$$
\begin{aligned}
E(Y_1) &= E(Y_1|T=1) \\
E(Y_0) &= E(Y_0|T=0)
\end{aligned}
$$

- Then, can substitute

$$
\begin{aligned}
ATE &= E(Y_1) - E(Y_0) \\
&= E(Y_1|T=1) - E(Y_0|T=0)
\end{aligned}
$$

# Potential Outcomes Model: Estimands

- If we know $(Y_1, Y_0)$ indep of $T$
- Then,

$$
\begin{aligned}
E\left(Y_1\right) &= E(Y_1|T=1) \\
E\left(Y_0\right) &= E(Y_0|T=0)
\end{aligned}
$$

- Then, can substitute

$$
\begin{aligned}
ATE &= E(Y_1) - E(Y_0) \\
&= E(Y_1|T=1) - E(Y_0|T=0)
\end{aligned}
$$

Observed diff in Tr and Co group means gives ATE!

# Potential Outcomes Model: Estimands

Holland (1986): "*prima facie effect*":

$$E(Y_t|S = t) - E(Y_c|S = c)$$

# Potential Outcomes Model: Estimands

Holland (1986): "*prima facie effect*":

$$E(Y_t|S=t) - E(Y_c|S=c)$$

"It is important to recognize that $E(Y_t)$ and $E(Y_t|S=t)$ are *not* the same thing ..."

# Potential Outcomes Model: Estimands, Interpretation

Gerber and Green (2012):

When $Y(1)$ and $Y(0)$ indep of $T$,

$$ATE = E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 0)$$

# Potential Outcomes Model: Estimands, Interpretation

Gerber and Green (2012):

When $Y(1)$ and $Y(0)$ indep of $T$,

$$
\begin{aligned}
ATE &= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1) + \\
&\qquad E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)
\end{aligned}
$$

# Potential Outcomes Model: Estimands, Interpretation

Gerber and Green (2012):

When $Y(1)$ and $Y(0)$ indep of $T$,

$$
\begin{aligned}
ATE &= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1) + \\
&\quad\; E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= [E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1)] + \\
&\quad\; [E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)]
\end{aligned}
$$

# Potential Outcomes Model: Estimands, Interpretation

Gerber and Green (2012):

When $Y(1)$ and $Y(0)$ indep of $T$,

$$
\begin{aligned}
ATE &= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1) + \\
&\quad E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= [E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1)] + \\
&\quad [E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)] \\
&= E(Y_i(1) - Y_i(0)|T_i = 1) + \\
&\quad E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)
\end{aligned}
$$

# Potential Outcomes Model: Estimands, Interpretation

Gerber and Green (2012):

When $Y(1)$ and $Y(0)$ indep of $T$,

$$\begin{aligned}
ATE &= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1) + \\
&\quad E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0) \\
&= [E(Y_i(1)|T_i = 1) - E(Y_i(0)|T_i = 1)] + \\
&\quad [E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)] \\
&= E(Y_i(1) - Y_i(0)|T_i = 1) + \\
&\quad E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)
\end{aligned}$$

$$\underbrace{E(Y_i(1) - Y_i(0)|T_i = 1)}_{\text{ATT}} + \underbrace{E(Y_i(0)|T_i = 1) - E(Y_i(0)|T_i = 0)}_{\text{Selection Bias}}$$

# Assignment of Treatment

# What can be a Treatment?

- Timing clearly defined

# What can be a Treatment?

- Timing clearly defined
- Covariates: causally prior to treatment

# What can be a Treatment?

- Timing clearly defined
- Covariates: causally prior to treatment
- "Attributes", "immutable characteristics": difficult to isolate

# What can be a Treatment?

- Timing clearly defined
- Covariates: causally prior to treatment
- "Attributes", "immutable characteristics": difficult to isolate

## What can be a Treatment?

- Timing clearly defined
- Covariates: causally prior to treatment
- "Attributes", "immutable characteristics": difficult to isolate

Sen and Wasow (2016) : "Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics" (*elements* of attributes varyingly manipulable)

# Attributes

"Causal effect of race"?

```r
data(resume, package = "qss")
dim(resume)
```

```
[1] 4870    4
```

# Attributes

"Causal effect of race"?

```
data(resume, package = "qss")
dim(resume)
```

```
[1] 4870    4
```

| | 0 | 1 |
|---|---|---|
| black | 2278 | 157 |
| white | 2200 | 235 |

# Attributes

"Causal effect of race"?

```r
data(resume, package = "qss")
dim(resume)
```

```
[1] 4870    4
```

| | 0 | 1 |
|---|---|---|
| black | 2278 | 157 |
| white | 2200 | 235 |

Bertrand and Mullainathan (2004): "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination"

# Random Assignment Mechanisms

- Simple/complete randomization
  (Bernoulli trial, prob $\pi$)

# Random Assignment Mechanisms

- Simple/complete randomization
  (Bernoulli trial, prob $\pi$)
- Complete randomization / random allocation
  (fixed proportion to tr)

# Random Assignment Mechanisms

- Simple/complete randomization
  (Bernoulli trial, prob $\pi$)
- Complete randomization / random allocation
  (fixed proportion to tr)
- Blocked randomizations
  (fixed proportion to tr, w/in group)

# Random Assignment Mechanisms

- Simple/complete randomization
  (Bernoulli trial, prob $\pi$)
- Complete randomization / random allocation
  (fixed proportion to tr)
- Blocked randomizations
  (fixed proportion to tr, w/in group)
- Cluster randomizations
  (assignment at higher level)

# The Potential Outcomes Model: Assignment

Observed outcome: $Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$

The assignment mechanism *selects* which potential outcome we observe.

# The Potential Outcomes Model: Assignment

Observed outcome: $Y_i = Y_i(1) \cdot T_i + Y_i(0) \cdot (1 - T_i)$

The assignment mechanism *selects* which potential outcome we observe.

(Gerber and Green (2012) use $Y_i = Y_i(1) \cdot d_i + Y_i(0) \cdot (1 - d_i)$ to highlight that we observe pot outcome from treatment actually taken, not hypothetical or assigned treatment.)

# The Potential Outcomes Model: Assignment

"Assignment mechanisms" are really missing-data-generating procedures.

# Ignorability

Assignment mechanism is *ignorable* if $Y_{mis}$ conditnly indep of $T$

$$P(T|X, Y_{obs}, Y_{mis}) = P(T|X, Y_{obs})$$

# Ignorability

Assignment mechanism is *ignorable* if $Y_{mis}$ conditnly indep of $T$

$$P(T|X, Y_{obs}, Y_{mis}) = P(T|X, Y_{obs})$$

Nothing in unobserved $Y_{mis}$ informs relationship between $Y_{obs}$, $T$.

# Unconfoundedness

Some ignorable mechanisms are *unconfounded*, too.

$$P(T|X, Y_{obs}, Y_{mis}) = P(T|X)$$

# Unconfoundedness

Some ignorable mechanisms are *unconfounded*, too.

$$P(T|X, Y_{obs}, Y_{mis}) = P(T|X)$$

Nothing in $Y$ informs $T$.

# Unconfoundedness

Some ignorable mechanisms are *unconfounded*, too.

$$P(T|X, Y_{obs}, Y_{mis}) = P(T|X)$$

Nothing in $Y$ informs $T$.

These are special cases of conditional independence.

# An Assignment Mechanism

Little & Rubin (2000):

| Patient | Y | T |
|---|---|---|
| 1 | 6 | 1 |
| 2 | 12 | 1 |
| 3 | 9 | 0 |
| 4 | 11 | 0 |

# An Assignment Mechanism

Little & Rubin (2000):

| Patient | Y | T |
|---|---|---|
| 1 | 6 | 1 |
| 2 | 12 | 1 |
| 3 | 9 | 0 |
| 4 | 11 | 0 |

Clearly, treatment is harmful. $\overline{Y(1)} - \overline{Y(0)} = 9 - 10 = -1$.

# An Assignment Mechanism

Little & Rubin (2000):

| Patient | Y(0) | Y(1) | $\tau$ | T |
|---|---|---|---|---|
| 1 | | 6 | | 1 |
| 2 | | 12 | | 1 |
| 3 | 9 | | | 0 |
| 4 | 11 | | | 0 |
| Mean | 10 | 9 | | |

Clearly, treatment is harmful.

$$\overline{Y(1)|T=1} - \overline{Y(0)|T=0} = 9 - 10 = -1$$

# An Assignment Mechanism: Perfect Doctor

Little & Rubin (2000):

| Patient | Y(0) | Y(1) | $\tau$ | T |
|---|---|---|---|---|
| 1 | (1) | 6 | | 1 |
| 2 | (3) | 12 | | 1 |
| 3 | 9 | (8) | | 0 |
| 4 | 11 | (10) | | 0 |
| | | | | |
| Mean | 10 | 9 | | |

# An Assignment Mechanism: Perfect Doctor

Little & Rubin (2000):

| Patient | Y(0) | Y(1) | $\tau$ | T |
|---|---|---|---|---|
| 1 | (1) | 6 | (5) | 1 |
| 2 | (3) | 12 | (9) | 1 |
| 3 | 9 | (8) | (-1) | 0 |
| 4 | 11 | (10) | (-1) | 0 |
| | | | | |
| Mean | 10 | 9 | (3) | |

# An Assignment Mechanism: Perfect Doctor

Little & Rubin (2000):

| Patient | Y(0) | Y(1) | $\tau$ | T |
|---|---|---|---|---|
| 1 | (1) | 6 | (5) | 1 |
| 2 | (3) | 12 | (9) | 1 |
| 3 | 9 | (8) | (-1) | 0 |
| 4 | 11 | (10) | (-1) | 0 |
| Mean | 10 | 9 | (3) | |

Clearly, treatment is beneficial:

$$\overline{Y(1)} - \overline{Y(0)} = 9 - 6 = 3$$

# An Assignment Mechanism: Perfect Doctor

Little & Rubin (2000):

| Patient | Y(0) | Y(1) | $\tau$ | T |
|---|---|---|---|---|
| 1 | (1) | 6 | (5) | 1 |
| 2 | (3) | 12 | (9) | 1 |
| 3 | 9 | (8) | (-1) | 0 |
| 4 | 11 | (10) | (-1) | 0 |
| | | | | |
| Mean | 10 | 9 | (3) | |

Clearly, treatment is beneficial:

$$\overline{Y(1)} - \overline{Y(0)} = 9 - 6 = 3$$

This assg mechanism is non-ignorable, confounded.

So, the *assignment mechanism* defines

So, the *assignment mechanism* defines how *causal* our empirical estimate may be.

# Next:
# Inference for Experiments

Next:
Inference for Experiments
(or, “There is only one test”)

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4): 991–1013.

Gerber, Alan S., and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation.* New York, NY: WW Norton.

Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60.

Sen, Maya, and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics." *Annual Reviews of Political Science* 19: 499–522.