# Applied Causal Inferece

## Lecture 3

Alex Stephenson

8-30-2021

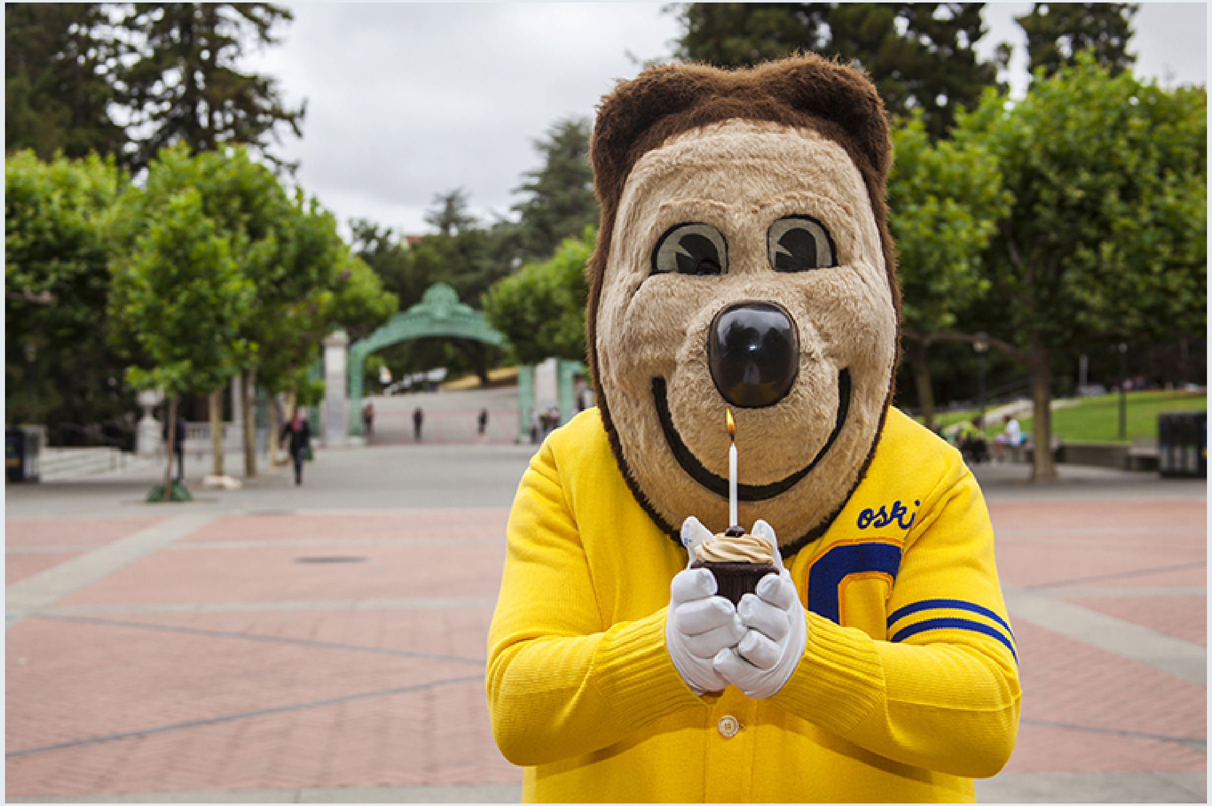
# What did we do last time?

- Covered a framework for research design
- Set up our five questions to describe research
- Discussed some basic ways to think about research ethics

# What are we doing today?

- Defining Causal Inference
- Introducing Potential Outcomes
- Introducing our first estimand

# Our avatar for Causal Inference



# Framework for Causal Inference

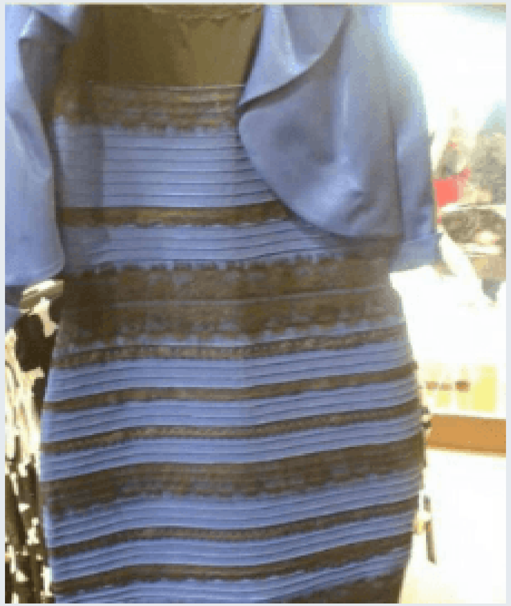
We follow a basic framework (Neyman 1923; Imbens and Rubin 2015) for thinking about causal questions. As a dictum, "No causation without manipulation."

Causality is tied to an action applied to a unit.

- "active" action is the treatment
- "passive" action is the control

# Binary Treatments

- We will be concerned for much of the class with binary treatments
- The framework generalizes to multi-valued treatments and continuous treatments



Is this dress black and blue or white and gold?

# Framework for Causal Inference

The framework has three parts:

1. Potential outcomes
2. Knowledge of the assignment mechanism
3. Stability of multiple observable units

# Potential Outcomes

Each unit has a fixed potential response to an action

Only one outcome will actually be observed, which is the action relative to treatment status.

The unrealized outcome is a counterfactual.

- Think Frost's "The Road not taken."

# Potential Outcomes

Our notation for Potential Outcomes is: $Y_i(0)$, $Y_i(1)$

While there is no formal meaning:

- Conventionally $Y_i(0)$ is called the "control" outcome
- Conventionally $Y_i(1)$ is called the "treatment" outcome

A causal effect is a comparison of potential outcomes.

- Individual Causal Effect (ICE): $\tau_i = Y_i(1) - Y_i(0)$
- We could think of other ICE like ratios

# Potential Outcomes

Summarizing two facts about causal effects:

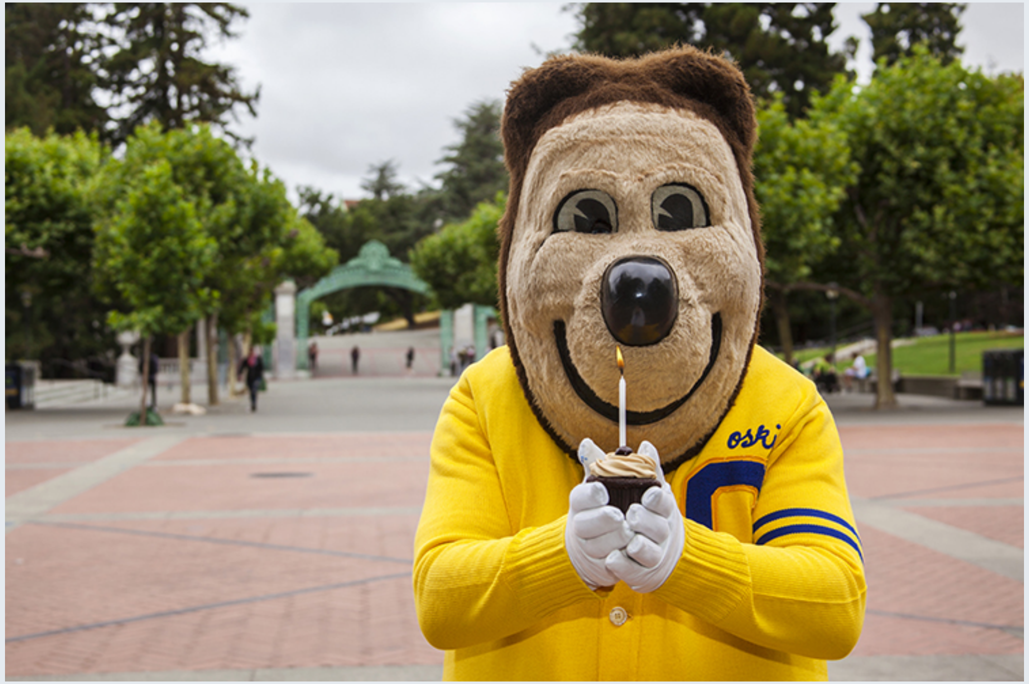
1. Causal effects do not depend on which outcome is actually observed
2. Causal effects is the comparison for the same unit at the same moment in time after treatment

An immediate implication of (2) is that $\tau_i$ cannot be identified, even with infinite data.

# Fundamental Problem of Causal Inference (Holland 1986)

- There is no way to ever observe the ICE.
- We only ever observe one outcome because we can only ever assign one treatment
- In other words, causal inference is fundamentally a missing data problem.

# Only Oski Knows Potential Outcomes




Oski the terrifying demon has access to the full potential outcomes schedule.

As observers of Oski's machinations, we do not. When we conduct research, it is important to keep that in mind.

# Potential Outcomes Example

| County | $Y_0$ | $Y_1$ |
|---|---|---|
| 1 | 10 | 5 |
| 2 | 15 | 10 |
| 3 | 10 | 5 |
| 4 | 20 | 15 |
| 5 | 15 | 10 |

# Assignment Mechanisms

The second part of our framework is the treatment.

Notation for binary case: $D \in \{0, 1\}$

- D is a random variable because it refers to treatment that could be administered hypothetically.
- We'll use $d$ to represent the observed treatment

For example, $Y_i(1)|D_i = 1$ refers to a hypothetical treated unit.

The assignment mechanism is what puts units into either the *treatment group* or *control group*

# The Potential Outcomes Model

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

This little equation is probably the most important equation that hasn't made its way into pop culture.

# Average Treatment Effect

The main *causal estimand* we consider is the Average Treatment Effect (ATE).

- The ATE is defined as $E[Y_i(1) - Y_i(0)]$
- When we go to data, the expected potential outcomes of treatment of treated units are $E[Y_i(1)|D_i = 1]$ and the expected potential outcomes of treatment of control units are $E[Y_i(1)|D_i = 0]$.
- Similarly the expected potential outcomes of control of the treatment units is $E[Y_i(0)|D_i(1)]$ and the expected potential outcomes of control of the control units is $E[Y_i(0)|D_i = 0]$

# Average Treatment Effect (ATE)

Although the ICE is fundamentally unknowable we can use experiments to learn about the Average Treatment Effect (ATE)

We can define the ATE as follows:

$$ATE = E[\pi_i]$$
$$ATE = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0]$$
$$ATE = E[Y_i(1) - Y_i(0)|D_i = 1] + E[Y_i(0)|D_i = 1] - E[Y_i(0)|D_i = 0]$$

The first term in the third line is the ATE among the treated. The second term is **Selection Bias**

# Stability Assumptions

In order to get a situation where we have two potential outcomes, it must be the case that potential outcomes depend **solely** on whether a unit **itself** receives a singular **treatment**.

Solely is an *exclusion* restriction assumption

Itself is a *non-interference* assumption

Treatment here means that there are not different levels or different kind of treatments.

# Stability Assumptions: Excludability

The point of a causal research design is to isolate the causal effect of treatment.

That means that the treatment received $d_i$ must be separated from the treatment group assigned. We call the latter $z_i$ where $z_i \in 0, 1$ is the observed group.

Formally: $Y_i(z_i = 1, d_i) = Y_i(z_i = 0, d_i)$

In words: The value of z is irrelevant to the potential outcomes. Potential outcomes only respond to the treatment.

# No interference

No interference means that a unit's potential outcomes are not affected by the treatment applied to a different unit. It also means that the treatment is the same for all units that receive treatment.

Consider our example about the effect of the coal industry on local government. If the coal industry in one county affects the decisions made by a local government in a different county $\rightarrow$ interference is present.

Formally $Y_i(\mathbf{d}) = Y_i(d)$ where $\mathbf{d}$ describes all treatments administered to every unit

# Where are doing on Wednesday?

- Explain DAGs
- Dive more in depth into estimands
    - Read the Testa piece for Wednesday