Segment 1: Fundamentals of Causal Inference

Section 05: Solving Fundamental Problem

Causal Effect

The causal effect of Z on Y for the i^{th} patient is defined by

$$Y_i^t - Y_i^c$$

Z causally affects Y if

$$Y_i^c \neq Y_i^t$$

The "**fundamental problem of causal inference**:" we only observe either Y_i^c *or* Y_i^t

Solving the Fundamental Problem

Two different ways to think about obtaining the “other” potential outcome:

1. The Scientific Solution
   - Exploit what we *know* about the situation to dictate the unobserved potential outcome
2. The Statistical Solution
   - Look in the data to find a good “close substitute” for the unobserved potential outcome

The “Scientific Solution”

Science (and everyday life) routinely infers causal relationships

- Does flipping a light switch cause the light to go on?
- Does turning the car steering wheel make the car turn?
- Pulling the rip cord causes the parachute to deploy

Why is it easy to declare “causality” here?

The "Scientific Solution"

Science (and everyday life) routinely infers causal relationships

- Does flipping a light switch cause the light to go on?
- Does turning the car steering wheel make the car turn?
- Pulling the rip cord causes the parachute to deploy

Why is it easy to declare "causality" here?

The scientific solution is to exploit what we "know" about the setting to infer the "other" potential outcome

Scientific Solution

Exploit certain assumptions to *imply* the value of the unobserved potential outcome

- *Temporal Stability*: The value of Y_i^c does not depend on *when* c is applied.
- *Causal Transience*: The value of Y_i^t is not affected by prior exposure to c.
  - Stability + transience $\Rightarrow$ Sufficient to sequentially expose Y_i to c then t.
- *Unit Homogeneity*: Assume $Y_i^t = Y_{i'}^t$ and $Y_i^c = Y_{i'}^c$
  - $\Rightarrow$ Compare Y_i^c to $Y_{i'}^t$
  - Think of a laboratory experiment

Example: Driving A Car

Does turning a car's steering wheel *cause* the car to turn? Let's design a trial....

1. Drive a car straight without turning the wheel.
   - The car goes straight
   -
2. Turn the wheel
   - The car turns
   -

Example: Driving A Car

Does turning a car's steering wheel *cause* the car to turn? Let's design a trial....

1. Drive a car straight without turning the wheel.
   - The car goes straight
   -
2. Turn the wheel
   - The car turns
   -

(At least) 2 reasons why we are comfortable assuming we have observed both potential outcomes:

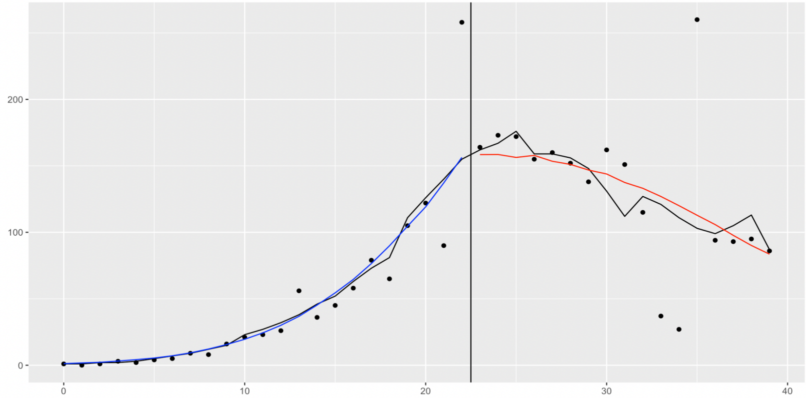
- We know how cars work!
- We are comfortable assuming *temporal stability*, that is, that outcome under"control condition" (no wheel turn) would be the same at any other point in time
- $\Rightarrow$ Had we never turned the wheel, the care would have never turned

Pre-Post Designs and Temporal Stability

A hypothetical "pre-post" design in the dietary example (Gelman, Hill, Vehtari Ch. 18.1):

- We see that Caitlin has a blood pressure of 155 after taking the supplements
  -
- What would be a good "close substitute" for the unobserved Y_i^c?
  - What if we knew her pre-study blood pressure was 160?
  - Temporally stable?

Example: COVID Deaths



The “Statistical Solution”

When the “other” potential outcome is not self-evident...

- Make use if a *population* of multiple units
  - E.g., multiple distinct individuals, same person observed at different point in times, etc.
- Estimate the “other” potential outcome using other units
  - E.g., use units that actually received the “other” treatment as a “close substitute”
- Estimate *average causal effects* that average over the units in a sample or population

Statistical Solution

Use the population U to make use of multiple units exposed to different treatments to make causal inferences

- Observe the same person under different treatments at different points in time
- Observe different people under different treatments at the same time

Does not "automatically" solve the fundamental problem, but can provide a basis for estimating average causal effects:

$$E[Y^t - Y^c] = E[Y^t] - E[Y^c]$$

Use average values of units exposed to t and average values of units exposed to c.

Statistical Solution

Use the population U to make use of multiple units exposed to different treatments to make causal inferences

- Observe the same person under different treatments at different points in time
- Observe different people under different treatments at the same time

Does not "automatically" solve the fundamental problem, but can provide a basis for estimating average causal effects:

$$E[Y^t - Y^c] = E[Y^t] - E[Y^c]$$

Use average values of units exposed to t and average values of units exposed to c. **Success will depend critically on what leads units to receive t or c.**

Example: One-Person Clinical Trial

HIV vaccine

- Give 1 HIV negative person a vaccine $(Z_i = 1)$
- Follow up for 10 years
- Observe that the person does not become infected with HIV
    -
- Did the vaccine *cause* the person remain uninfected?

Cause and Effect?

Key Question: Would this person have become infected absent the vaccine?

- Could compare what *did* happen to what *would have* happened
  - Uninfected with vaccine + infected with placebo $\Rightarrow$ vaccine *caused* the person to remain uninfected
- Of course, we only get to observe one of these...

Scientific solution?

Randomized Experiment

Randomize 1000 patients to receive vaccine ($Z = 1$) or placebo ($Z = 0$)

- No infections in the vaccine arm
- 30% of the patients in the control arm become infected
- Use *average outcome* in placebo recipients as a "close substitute" for *average outcome* in vaccine recipients
  - Observed placebo outcomes represent what *would have happened* to the vaccine recipients
- Conclude that the vaccine *causes* a person to remain HIV free

Randomization is key!