Causality 1: Basic Concepts

The Causality Problem in Social Science

- Question: Whether a treatment, T, has an effect on Y.
- $\delta_i = Y_i^t - Y_i^c$

  where Y_i^t refers to the result of Y when treatment t is presented to individual i and Y_i^c refers to the result of Y when treatment t is absent (i.e., control) to the same individual i. We can define this individual-level causal effect.
- We cannot have both Y_i^t and Y_i^c. An easiest way to solve this problem is random assignment of T.
- However for social science, we cannot do this. (E.g., we cannot provide college education and no-college education to the same individual. Nor can we randomly assign college education to kids.)
- What we observe, in reality, is either Y_i^t when $T_i = 1$ or Y_i^c when $T_i = 0$. (E.g., we don't have the outcome of no-college education for kids who are college educated. We don't have the outcome of college education for kids who are not college educated.) In this sense, causal inference is a problem of missing data.
- To solve this problem, methodologists have developed "counter-factual analyses" to infer causality from observational data.
- Counter-factual framework: Given that the δ_i cannot be calculated for any individual and therefore that Y_i^t and Y_i^c can be observed only on mutually exclusive subsets of the population, what can be inferred about the distribution of the δ_i from an analysis of Y_i (= observed outcome) and T_i (= observed treatment).
- Causal inference, in other words, is equivalent to answering "what if" question.

Average Treatment Effects (ATE) and the Standard (or Näive) Estimator

- ATE of the standard estimator:

$$\begin{aligned}\bar{\delta} &= \bar{Y}^t - \bar{Y}^c \\ &= E[Y^t|T=1] - E[Y^c|T=0]\end{aligned} \tag{1}$$

  E.g., the effect of college on earnings (= $\bar{\delta}$) is the difference between the average earnings of the college educated (=$\bar{Y}^t$) and the average earnings of the no college educated (=$\bar{Y}^c$) net of other control variables. This is what we get with regression models.
- What we really estimate with the standard estimator is

$$\bar{\delta} = \bar{Y}^t_{i\in T} - \bar{Y}^c_{i\in C} \tag{2}$$

  where $\bar{Y}^t_{i\in T}$ refers to the mean of Y for the treated when individual i is assigned to the treatment and $\bar{Y}^c_{i\in C}$ refers to the mean of Y for the untreated (i.e., control group) when individual i is assigned to the control group.

- In this setting, we do not have information on what will happen if the treated individuals are untreated ($= \bar{Y}^c_{i \in T}$) and what will happen if the untreated individuals are treated ($= \bar{Y}^t_{i \in C}$).
- The real causality (ATE) should be:

$$\bar{\delta} = \pi(\bar{Y}^t_{i \in T} - \bar{Y}^c_{i \in T}) + (1 - \pi)(\bar{Y}^t_{i \in C} - \bar{Y}^c_{i \in C}) \tag{3}$$

  where π is the probability of assigning into the treatment. Thus, the ATE is the weighted average between the effect of treatment for the treated and the effect of treatment for the untreated.

- If we assume $\pi = .5$,

$$\begin{aligned}\bar{\delta} &= .5(\bar{Y}^t_{i \in T} - \bar{Y}^c_{i \in T}) + .5(\bar{Y}^t_{i \in C} - \bar{Y}^c_{i \in C}) \\ &= .5(\bar{Y}^t_{i \in T} - \bar{Y}^c_{i \in C}) + .5(\bar{Y}^t_{i \in C} - \bar{Y}^c_{i \in T})\end{aligned} \tag{4}$$

  Thus, $\bar{Y}^t_{i \in T} - \bar{Y}^c_{i \in C}$ is an unbiased estimate only if $\bar{Y}^t_{i \in T} = \bar{Y}^t_{i \in C}$ and $\bar{Y}^c_{i \in T} = \bar{Y}^c_{i \in C}$. That is, the effect of the treatment should be equal to the treated and the untreated, and the effect of untreatment should be identical between the treated and the untreated. E.g., the effect of college education should be identical for those who went to college and who didn't.

- In order to satisfy these equality conditions, a sufficient condition is that treatment assignment T_i be uncorrelated with the potential outcome distributions of Y^t_i and Y^c_i. The principal way to achieve this uncorrelatedness is through random assignment to the treatment.
- From equation (4), we know 2 sources of bias:

$$\begin{aligned}\bar{Y}^t_{i \in T} - \bar{Y}^c_{i \in C} &= \frac{\bar{\delta}}{.5} + (\bar{Y}^c_{i \in T} - \bar{Y}^t_{i \in C}) \\ &= \bar{\delta} + (\bar{Y}^c_{i \in T} - \bar{Y}^t_{i \in C}) - \frac{\bar{\delta}}{.5}\end{aligned}$$

  More generally, when we do not assume $\pi = .5$,

$$\bar{Y}^t_{i \in T} - \bar{Y}^c_{i \in C} = \bar{\delta} + (\bar{Y}^c_{i \in T} - \bar{Y}^t_{i \in C}) + (1 - \pi)(\bar{\delta}_{i \in T} - \bar{\delta}_{i \in C}) \tag{5}$$

  Thus, the standard estimator is the true effect of the treatment, $\bar{\delta}$ plus two biases: first, $(\bar{Y}^c_{i \in T} - \bar{Y}^t_{i \in C})$ and second, $(1 - \pi)(\bar{\delta}_{i \in T} - \bar{\delta}_{i \in C})$.

  1. The first source of bias, $(\bar{Y}^c_{i \in T} - \bar{Y}^t_{i \in C})$ is equal to the difference between the treatment and control groups in the absence of treatment. E.g., difference in the labor market outcome for BA+ and non-BA if none are college educated. BA+ would earn more even if they don't have college education because they are smarter at the first place and more ambitious.
     $\longrightarrow$ Selection effect
  2. The 2nd source of bias, $(1 - \pi)(\bar{\delta}_{i \in T} - \bar{\delta}_{i \in C})$, is the difference in the treatment effect for those in the treatment and control groups. E.g., the effect of college education will differ between those who are actually college educated and those who are not college educated. Recall negative and positive selection debate (Brand and Xie 2010 ASR).
     $\longrightarrow$ Treatment heterogeneity

- An example: effect of college education

    1. Real effect: College education makes people more productive and smarter.
    2. Selection: Smarter and more productive people tend to go to college more likely.
    3. Treatment heterogeneity: Productivity may have increased more for those attended a college than those those did not attend a college.

Average Treatment Effects for the treated (ATT)

- ATE is not necessarily of theoretical interest. Sometimes the average treatment effects for the treated (ATT $= \bar{Y}^t_{i\in T} - \bar{Y}^c_{i\in T}$) only can be of researchers' interest.
- For example, the effect of a job training for those who didn't go to college would be negative if we include college goers in the control group. But the effect of job-training for college goers is not our interest.
- ATU: Average treatment effects for the untreated.

Key Assumption of the Causality

- SUTVA assumption: a change in treatment status of any individual does not affect the potential outcomes of other individuals. This is known as "the stable unit treatment value assumption (SUTVA)."
    - E.g., the supply of more college educated workers will not change the earnings for the no-college educated workers.
- Conditional Independence (CIA), Strict Ignorability, Exogeneity, or Selection on observables: observed covariates X comprises a sufficient set of joint causes of the treatment on the outcome so that, conditional on covariates, variation in the treatment is as good as randomly assigned. That is, the treatment is unrelated to either unobserved heterogeneity (omitted variable bias) or agents' purposive choice of treatment (self selection or endogeneity).
    - Note that CIA is not a primarily statistical, but rather a fundamental theoretical statement.

Several Causality Models

- Natural Experiments
- Fixed Effects Model
- Difference-in-difference
- Inverse Probability Treatment Weight (IPTW)

- Instrumental Variable (IV)
- Propensity Score Matching