Causal Inference: Notation and Basic Concepts

Hyunseung Kang

2025-01-21

Concepts Covered Today

- Association between smoking and lung function
- Defining causal quantities with counterfactual/potential outcomes
- A setting where causal effects are identified: unit homogeneity
- References:
  - Pages 1-18 of Shadish, Cook, and Campbell (2002) (for concepts)
  - Chapter 1 of Hernán and Robins (2020)

Motivating Example: Smoking and Lung Function

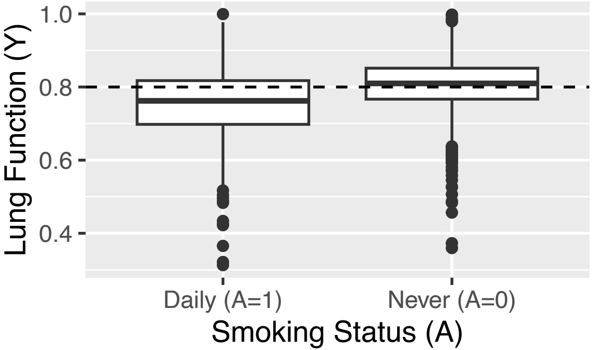
Table 1: A Subset of the Observed Data

Lung function (Y)	Smoking status (A)
0.94	0
0.92	0
0.81	1
0.84	0

Data: 2009-2010 National Health and Nutrition Examination Survey (NHANES).

- Treatment (A): Daily smoker ($A = 1$) vs. never smoker ($A = 0$)
- Outcome (Y): ratio of forced expiratory volume in one second over forced vital capacity. $Y \geq 0.8$ is good lung function!
- Sample size is $n = 2360$.

Association of Smoking and Lung Function



- $\overline{Y}_{\text{daily(A=1)}} = 0.75$ and $\overline{Y}_{\text{never(A=0)}} = 0.81$.
- t-stat = -11.8, two-sided p value: $\ll 10^{-16}$

Daily smoking is **strongly associated** with reduction in lung function.

But, is the strong association evidence for **causality**? After all, association does not imply causation...

Building Intuition for Causality: The Parallel Universe Analogy I

John's Parallel Universe 1 (A=1)

John's Parallel Universe 2 (A=0)

Inspired by recent Marvel movies, I find the parallel universe analogy helpful to conceptualize causal effects.

Consider a particular snapshot in time (e.g., June 1, 2024, John's 25th birthday) in two parallel universes.

- In universe 1, John is a daily smoker.
- In universe 2, John never smoked.
- Beyond smoking status, everything is identical between universe 1 and 2 (John's age, friends, parents, diet, etc.)

Building Intuition for Causality: The Parallel Universe Analogy II

Now suppose John's lung functions are different between the two universes.

- The difference in lung functions can **only be** attributed to the difference in smoking status.
- Why? All variables (except smoking status) are the same between the two parallel universes.
- Between the two parallel universes, **any difference** in the outcomes must be due to a **difference in the treatment status.**

A helpful Youtube clip from movie Sliding Doors.

Counterfactual/Potential Outcomes

Let's define the outcomes from the two parallel universes, often referred to as counterfactual or potential outcomes[1].

- $Y(1)$: lung function that would have been observed if you smoked daily (i.e., parallel world where you smoked)
- $Y(0)$: lung function that would have been observed if you did not smoke (i.e,. parallel world where you didn't smoke)

Similar to the observed data table, we can create counterfactual/potential outcomes data table.

	$Y(1)$	$Y(0)$
John	0.54	0.94
Sally	0.91	0.91
Kate	0.81	0.60
Jason	0.60	0.84

[1] The framework was developed by Neyman (1923), republished in Statistical Science in 1990, and Rubin (1974). See Holland (1986) for more background.

Our First Causal Effect: Individual Causal Effects

Let's take a look at $Y_{\text{John}}(1) - Y_{\text{John}}(0) = -0.4$ and $Y_{\text{Sally}}(1) - Y_{\text{Sally}}(0) = 0$.

- For John, changing smoking status *causes* a change in his lung function since the difference between $Y_{\text{John}}(1)$ and $Y_{\text{John}}(0)$ can only be attributed to the difference in smoking status in the parallel universes.
- Unlike John, changing Sally's smoking status will *not cause* a change her lung function.

Both numbers -0.4 and 0 are **individual causal effects** as they reflect each person's change in the outcome when their smoking status changes.

When causal effects differ from individual to individual, the causal effect is generally said to be **heterogeneous.** If the effects are the same for every individual, the causal effect is generally said to be **homogeneous** or **constant.**

Other Measures of Causal Effects

Suppose we add additional information about the individuals

	$Y(1)$	$Y(0)$	Age (X_1)	Graduated HS? (X_2)
John	0.54	0.94	23	Yes
Sally	0.91	0.91	27	No
Kate	0.81	0.60	32	No
Jason	0.60	0.84	30	Yes

- The **average**[2] **treatment effect (ATE)**: $\mathbb{E}[Y(1) - Y(0)]$
- The **conditional average treatment effect (CATE)**: $\mathbb{E}[Y(1) - Y(0) | X_2 = \text{Yes}]$

These are examples of **causal estimands/parameters** because they are functions of the counterfactual outcomes.

[2]Expectations are defined with respect to a joint cumulative distribution function $F_{Y(1),Y(0)}$ (i.e., super-population framework).

Average Treatment Effect (i.e., *the* Causal Effect) I

Let's consider the ATE $\mathbb{E}[Y(1) - Y(0)]$, by far the most popular causal estimand/measure of a causal effect.

- This is the average of John's, Sally's, etc. causal effects of smoking on lung function.

If this average is zero, then on average, the causal effect of smoking on lung function is zero.

- This doesn't mean that everyone's individual causal effect is zero.
- Some people may have a positive individual causal effect, others may have a negative individual causal effect, and some may have zero individual causal effect.

If this average is negative, then being a daily smoker is, on average, cause decrease in lung function compared to being a never-smoker.

By linearity of expectations, $\mathbb{E}[Y(1) - Y(0)] = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)]$.

Average Treatment Effect (i.e., *the* Causal Effect) II

- In words, the average of everyone's causal effects is also the difference in the average of everyone's lung functions when they are daily smokers (i.e., $\mathbb{E}[Y(1)]$) versus when they are never smokers (i.e., $\mathbb{E}[Y(0)]$).
- While the equality is trivial, it allows us to study the ATE by studying the *marginal* distributions of $Y(1)$ and $Y(0)$ rather than studying the *joint distribution* of $Y(1), Y(0)$.

Some Subtle Points about Causal Effects I

1. Causal effects of a treatment is often defined as a **comparison** to another (possibly inactive) treatment.
   - In addition to linear differences,causal effects can also be defined as contrasts in log scale, $\log(Y(1)) - \log(Y(0))$ so long as $Y(0), Y(1)$ are positive.
   - See Chapter 1.3 of Hernán and Robins (2020) for details.

2. We focus on the **effects of causes** (e.g., effect of daily smoking on lung function) rather than the **causes of effects** (e.g., why does John have poor lung function?). The causes of effects are hard to define because of the problem of infinite regress.

   Example (from Don Rubin): He got lung cancer because he smoked cigarettes. The real reason he smoked is because his parents smoked, and they smoked because they hated each other and they hated each other because...

Some Subtle Points about Causal Effects II

3. Cause-effect relationships have a natural **temporal ordering** where the treatment variable (i.e., smoking status) always precedes the outcome variable (i.e., lung function)
   - You can't have an effect (i.e. outcome) before a cause (i.e. treatment variable).
   - You also can't have causal simultaneity where the outcome and treatment variable simultaneously change each others values at the exact same time. This makes it impossible to determine whether the outcome is causing the treatment variable or vice versa.
4. (Discussed more later) The notation currently does not make a distinction between different kinds of daily smoking on lung function (e.g., John smokes 10 packs of cigars per day versus 1 cigar per day). The notation assumes *no multiple versions of treatment*.

Counterfactual Data Versus Observed Data I

Table 4: Comparison of tables.

(a) Counterfactual table

	$Y(1)$	$Y(0)$
John	0.54	0.94
Sally	0.91	0.91
Kate	0.81	0.60
Jason	0.60	0.84

(b) Observed table

	Y	A
John	0.94	0
Sally	0.91	0
Kate	0.81	1
Jason	0.84	0

In the counterfactual table, we see what everyone's lung function would be if they are never-smokers and daily smokers.

- Comparing John's $Y(1)$ and $Y(0)$ gives us John's causal effect of being a daily smoker versus a never-smoker on his lung function.

Counterfactual Data Versus Observed Data II

- Similary, comparing Sally's $Y(1)$ and $Y(0)$ gives us Sally's causal effect of being a daily smoker versus a never-smoker on her lung function.

In the observed table, we only see everyone's lung function under one particular status of smoking status.

- We only see John's lung function when he is a non-smoker (i.e., $Y_{\text{John}} = 0.94$ when $A_{\text{John}} = 0$). We don't get to see his lung function in the parallel universe when, contrary to fact, he is a daily smoker.
- Similarly, we only see Sally's lung function when she is a non-smoker (i.e., $Y_{\text{Sally}} = 0.91$ when $A_{\text{Sally}} = 0$). We don't get to see her lung function in the parallel universe, when contrary to fact, she is a daily smoker.

Fundamental Problem of Causal Inference (Holland 1986)

Without additional information, it's impossible to study causal effects from the observed data table because we don't get to observe all counterfactual outcomes. This is the **fundamental problem of casual inference** (Holland 1986).

A key goal in causal inference is to learn about both counterfactual outcomes $Y(1), Y(0)$ when you only observe one of them.

- This often involves making (usually untestable) **assumptions** about the counterfactual data and/or the observed data.
- These assumptions are often referred to as assumptions for **causal identification.**

The fundamental problem is closely related to a missing data problem; we'll explore this later.

When Can You Observe Both Counterfactuals? Unit Homogeneity Assumption I

There are situations in the real world where you can observe all counterfactual outcomes. Most of them take place in lab experiments or in manufacturing and all of them fundamentally rely on some domain knowledge to claim that all counterfactual outcomes are observable.

Suppose we want to determine the causal effect of putting a chocolate bar over a candle.

- $Y(1)$: the counterfactual outcome of the chocolate bar if it's over a candle.
- $Y(0)$: the counterfactual outcome of the chocolate bar if it's not over a candle.
- Let's say these outcomes measure whether the chocolate melted (1) or not (0).

When Can You Observe Both Counterfactuals? Unit Homogeneity Assumption II

We put one chocolate bar over a candle and another bar away from the candle, resulting in the following table.

	$Y(1)$	$Y(0)$
1st chocolate bar	1	NA
2nd chocolate bar	NA	0

When Can You Observe Both Counterfactuals? Unit Homogeneity Assumption III

Despite the missing values in the potential outcomes, we can *impute* them from our daily experiences.

- We know that chocolates bars are identical *with respect to their behavior under heat.*
- Therefore, we can obtain the second chocolate bar's missing $Y(1)$ from the first chocolate bar's $Y(1)$.
- Similarly, we know that chocolates don't melt without heat and thus, we can impute the missing first chocolate bar's $Y(0)$ with the second chocolate bar's $Y(0)$.

This phenomena is known as the **unit homogeneity assumption** and is formalized as follows

$$Y_i(1) = Y_j(1) \text{ and } Y_i(0) = Y_j(0) \quad \forall i \neq j$$

When Can You Observe Both Counterfactuals? Unit Homogeneity Assumption IV

Note that we don't even have to randomize which chocolate bar is exposed to heat or not to identify the causal effect of heat on the chocolate bar. We also don't have to sample 10 or 100 chocolate bars to understand the causal effect of exposing chocolate to heat on melting as all chocolate bar.

References I

Hernán, Miguel, and James Robins. 2020. *Causal Inference: What If.* Boca Raton: Chapman & Hall/CRC.

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

Neyman, Jerzy. 1923. "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9." *Statist. Sci.* 5 (4): 465–72.

Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66 (5): 688.

Shadish, W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton, Mifflin, & Company.