Chapter 1. Causality: The Basic Framework

SeongSik Choi

July 7, 2022

Seoul National University

1. Causation $\neq$ Association

Association doesn't imply causation.

Two possibilities exist:

1. **(Reverse causality)** Going to fine dining at least every half year $\rightarrow$ Earning more money
2. **(3rd variable)** Receiving anticancer treatments $\rightarrow$ High mortality

1. Causation $\neq$ Association

1. **(Reverse causality)** Going to fine dining at least every half year $\rightarrow$ Earning more money

Y : Monthly Salary ($)

"Not Going" : $n = 850, \bar{y} = 2500, s = 400$

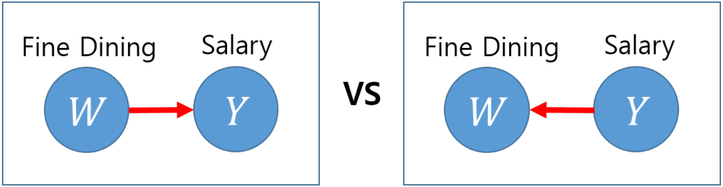
"Going" : $n = 150, \bar{y} = 5700, s = 1100$

Can we say **"If you go to fine dining at least every half year, your salary may be going to increase."**?

1. Causation $\neq$ Association

❶ **(Reverse causality)** Going to fine dining at least every half year $\rightarrow$ Earning more money

Y : Monthly Salary ($)



We can't say **"If you go to fine dining at least every half year, your salary may be going to increase."**

1. Causation $\neq$ Association

❷ **(3rd variable)** Taking intensive care → High mortality

W : Taking intensive care in 2022.03 (0 : No, 1 : Yes)

Y : Survival in 2022.04 (0 : No, 1 : Yes)

		Y	
		0	1
W	0	10	340
	1	20	30

Can we say **"Taking intensive care increases mortality rate by 37%."**?

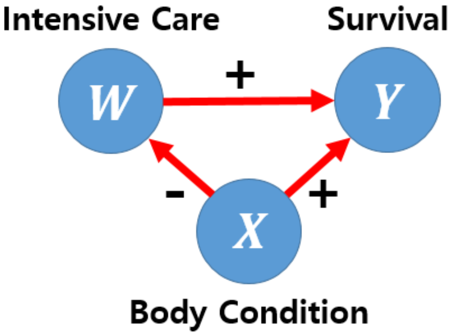
1. Causation $\neq$ Association

❷ **(3rd variable)** Taking intensive care → High mortality

W : Taking intensive care in 2022.03 (0 : No, 1 : Yes)

Y : Survival in 2022.04 (0 : No, 1 : Yes)

X : Body condition in 2022.02 (0 : Bad, 1 : Good)



1. Causation $\neq$ Association

The possibilities of reverse causality and 3rd variable should be considered simultaneously.

Basically, causal studies are difficult to conduct. But journalists want to build a story. So they are tempted to write correlation study as if it is causation study.

Eat Your Breakfast: Study Shows Impact on Grades

Compared to students who frequently ate **breakfast**, students who rarely ate **breakfast** received significantly lower average scores on the GCSE. Socioeconomic ...

2. Notation and Example

Example : One wants to find the causal effects of taking aspirin on headache.

- **Outcome** : $Y \in \{0(\text{No Headache}), 1(\text{Headache})\}$
- **Treatment(Action)** : $T, D, W, (A, Z)$
  $W \in \{0(\text{No Aspirin}), 1(\text{Aspirin})\}$
- **Covariate(Potential confounder, Pre-treatment variable, Attribute, 3rd variable)** : $X, C(\text{Observed}), U(\text{Unobserved})$
  X includes the intensity of the headache before making the decision to take aspirin or not.

3. General aim of causal inference

Example : One wants to find the causal effects of taking aspirin on headache.

Potential (Counterfacual) outcome : $Y_i(w)$.
The outcome for an individual i under a potential treatment w.

Unit-level treatment(causal) effect : $Y_i(1) - Y_i(0)$

ATE(Average Treatment Effect) : $\tau = \mathbb{E}(Y(1) - Y(0))$

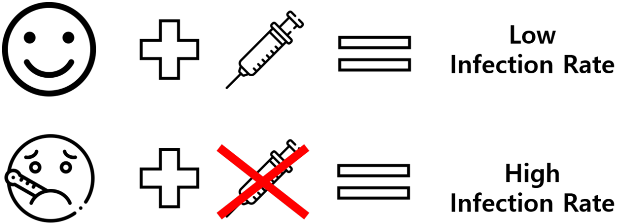
CATE(Conditional Average Treatment Effect) :
$\tau(x) = \mathbb{E}(Y(1) - Y(0)|X = x)$

The general aim of causal inference is estimating parameters such as ATE, CATE, $\cdots$.

4. RCT and observational study

RCT(Randomized Controlled Trial) : Toss a coin to decide subject will receive treatment. When it comes to covariates, Treated Group $\approx$ Untreated Group.

Observational study : Observe the covariates, treatments, outcomes. Treated Group $\not\approx$ Untreated Group.



4. RCT and observational study

To estimate causal effect in observational study, we should adjust the covariate. Some adjusting methods are as below.

1. **Matching** : match treated individual with untreated individual who has similar covariates.
2.
3.
4.

4. RCT and observational study

To estimate causal effect in observational study, we should adjust the covariate. Some adjusting methods are as below.

1. **Matching** : match treated individual with untreated individual who has similar covariates.
2. **Standardization** :
   $\mathbb{E}\left(Y(w)\right) = \sum_x \mathbb{E}\left(Y(w) \mid X = x\right) \mathrm{P}(X = x)$
   $= \sum_x \mathbb{E}(Y|X = x, W = w)\mathrm{P}(X = x)$
   Model $\mathbb{E}(Y|X = x, W = w)$. (Outcome approach)
3. **IP weighting** : $\mathbb{E}\left(Y(w)\right) = \mathbb{E}_{(X,W,Y)}\left[\frac{Y1(W=w)}{\mathrm{P}(W=w|X=x)}\right]$.
   Model $\mathrm{P}(W = w \mid X)$. (Weighting approch).
4.

4. RCT and observational study

To estimate causal effect in observational study, we should adjust the covariate. Some adjusting methods are as below.

1. **Matching** : match treated individual with untreated individual who has similar covariates.
2. **Standardization** :
   $\mathbb{E}\left(Y(w)\right) = \sum_x \mathbb{E}\left(Y(w) \mid X = x\right) \mathrm{P}(X = x)$
   $= \sum_x \mathbb{E}(Y|X = x, W = w)\mathrm{P}(X = x)$
   Model $\mathbb{E}(Y|X = x, W = w)$. (Outcome approach)
3. **IP weighting** : $\mathbb{E}\left(Y(w)\right) = \mathbb{E}_{(X,W,Y)}\left[\frac{Y1(W=w)}{\mathrm{P}(W=w|X=x)}\right]$.
   Model $\mathrm{P}(W = w \mid X)$. (Weighting approch).
4. **Doubly robust** : Model $\mathbb{E}(Y|X = x, W = w)$ and $\mathrm{P}(W = w \mid X = x)$ both. Consistent if either one is correct.

5. Assumptions

Causal assumptions and modeling assumptions are needed.

Causal assumptions : shared in most causal inference researches. By applying these assumptions, we can express our estimand without causal notation in many cases.

Modeling assumptions : related with how to estimate this usual estimand. They are similar with assumptions for traditional ML.

Example(Standardization) :
$\mathbb{E}\left(Y(w)\right) = \sum_x \mathbb{E}\left(Y(w) \mid X = x\right) \mathrm{P}(X = x)$
$= \sum_x \mathbb{E}(Y|X = x, W = w)\mathrm{P}(X = x)$ (Causal assumptions)
Model $\mathbb{E}(Y|X = x, W = w)$. (Modeling assumptions)

5. Assumptions

Causal assumptions :

1. **Unconfounfoundedness, Ignorability, Exogeneity, Conditional independence, Exchangeability** : $[Y(w) \perp\!\!\!\perp W]|X$
   (X : measured) (There isn't any unmeasured confounder.)
2.
3.
   -
   -

5. Assumptions

Causal assumptions :

1. **Unconfoundedness, Ignorability, Exogeneity, Conditional independence, Exchangeability** : $[Y(w) \perp\!\!\!\perp W] | X$
   (X : measured) (There isn't any unmeasured confounder.)
2. **Positivity, Overlap** : $\mathrm{P}(W = w | X = x) > 0, ^{\forall} x, w$
3.
   -
   -

5. Assumptions

Causal assumptions :

1. **Unconfoundedness, Ignorability, Exogeneity, Conditional independence, Exchangeability** : $[Y(w) \perp\!\!\!\perp W]|X$
   (X : measured) (There isn't any unmeasured confounder.)
2. **Positivity, Overlap** : $\mathrm{P}(W = w|X = x) > 0, ^{\forall} x, w$
3. **SUTVA(Stable Unit Treatment Value Assumption)** :
   - No Interference ($W_i \nrightarrow Y_j, ^{\forall} i \neq j$)
   - No hidden variations of treatments

In order to assess the causal assumptions, we should rely on expertise. (not on data!)

6. Samples, Populations, and Super-populations

Population : the finite set of units for which we observe covariates, treatments, and realized outcomes is the set of units we are interested in.

It is useful to view the set of units for which we observe values as drawn randomly from infinite population. In that case,

Finite samples(fs) : the set of units for which we observe values.

Super-populations(sp) : infinite population that finite samples were drawn from.

7. Summary

- Association $\nrightarrow$ causation.(Reverse causality, 3rd variable)
- We want to estimate ATE : $\tau = \mathbb{E}(Y(1) - Y(0))$
- In observational study, we should adjust the covariate. Four adjusting methods(Matching, Standardization, IP weighting, Doubly robust)
- Causal assumptions : Unconfoundedness, Positivity, SUTVA

8. Structure of the Book

PART I. Introduction

PART II. Classical randomized experiments

PART III. Regular assignment mechanisms: Design

PART IV. Regular assignment mechanisms: Analysis

PART V. Regular assignment mechanisms: Supplementary analyses

PART VI. Regular assignment mechanisms with noncompliance: Analysis