Causal Inference

Chapter 8. Selection Bias

Yuha Park

July 27, 2021

Seoul National University

Contents

What is selection bias

What is selection bias

Definition of selection bias:

- Bias resulting from **conditioning** on the **common effect** of treatment and outcome

The structure of selection bias

Estimate the average causal effect of wasabi intake A on the one-year risk of death Y.

- A : wasabi intake (1: yes, 0: no)
- C : censoring indicator (1: censored, 0: uncensored)
- U : atherosclerosis (unmeasured)
- L : heart disease (1: yes, 0: no)
- Y : death (1: yes, 0: no)

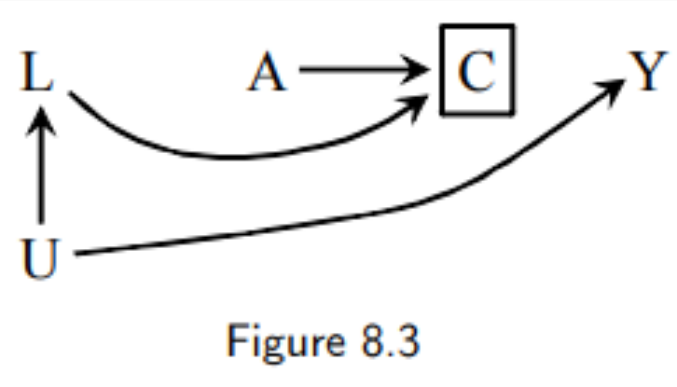

Figure 8.3

◯ The open path : $A \rightarrow C \leftarrow L \leftarrow U \rightarrow Y$

The structure of selection bias

Because of selection bias,

$$(0.89 =) \quad \frac{P(Y = 1|A = 1, C = 0)}{P(Y = 1|A = 0, C = 0)} \neq \frac{P(Y^{a=1} = 1)}{P(Y^{a=0} = 1)} \quad (= 1)$$

If the analysis was not conditioned on the common effect (collider) C,

$$\frac{P(Y = 1|A = 1)}{P(Y = 1|A = 0)} = \frac{P(Y^{a=1} = 1)}{P(Y^{a=0} = 1)}$$

Examples of selection bias

- Differential loss to follow-up
- Missing data bias, nonresponse bias
- Healthy worker bias
- Self-selection bias, volunteer bias
- Selection affected by treatment received before study entry

Selection bias vs. confounding

- **Confounding** : biases due to the presence of common causes
- **selection bias** : biases due to conditioning on common effects

- **Randomization** protects against confounding, but not against selection bias when the selection occurs after the randomization

How to adjust for selection bias

How to adjust for selection bias

- The identifiability conditions of **exchangeability, positivity, and consistency** hold for C as well as for A

- If these conditions hold, we can adjust the selection bias by **IP weighting** (or by standardization)

Application of IP weighting for selection bias adjustment

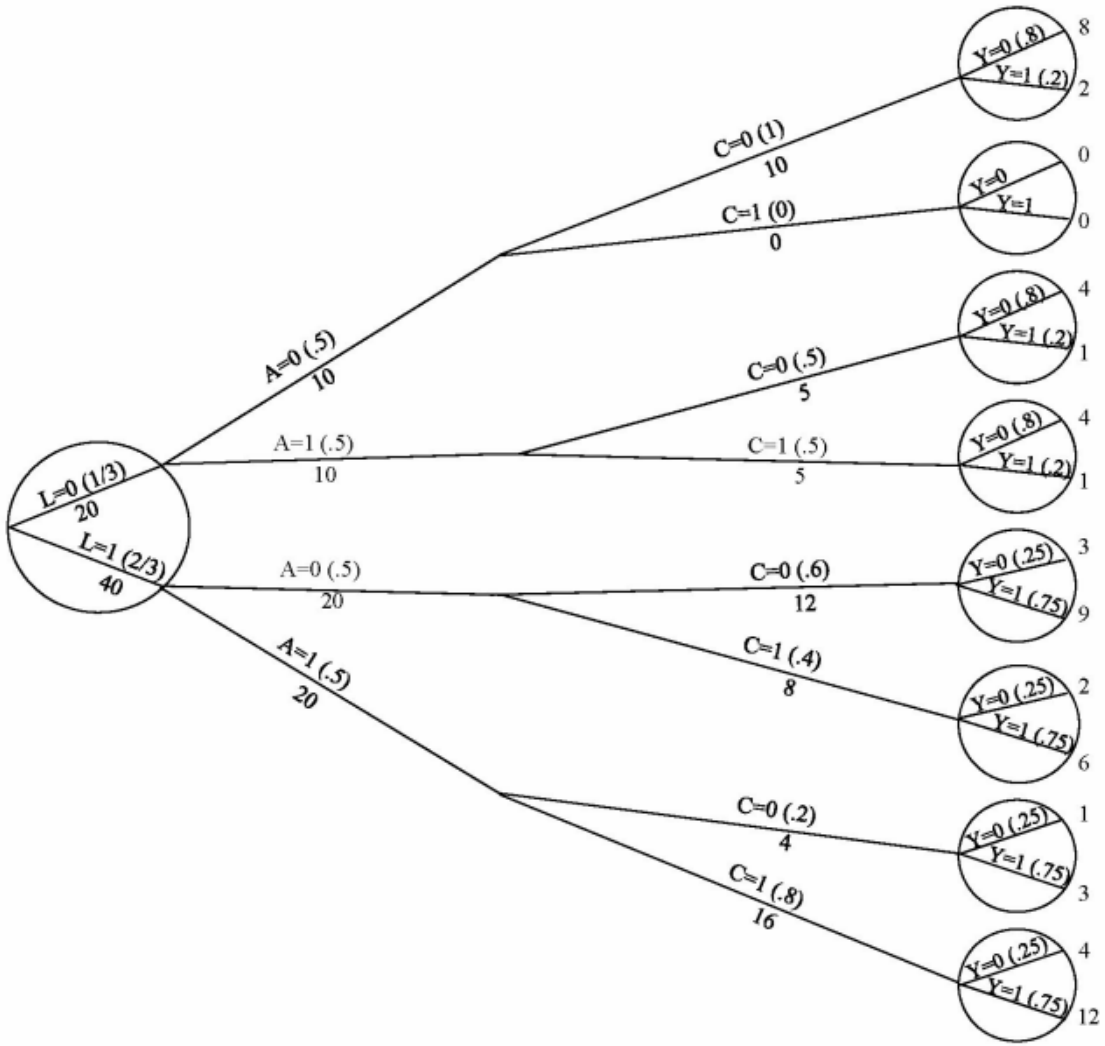

Figure 8.10

Pseudo-population for selection bias adjustment

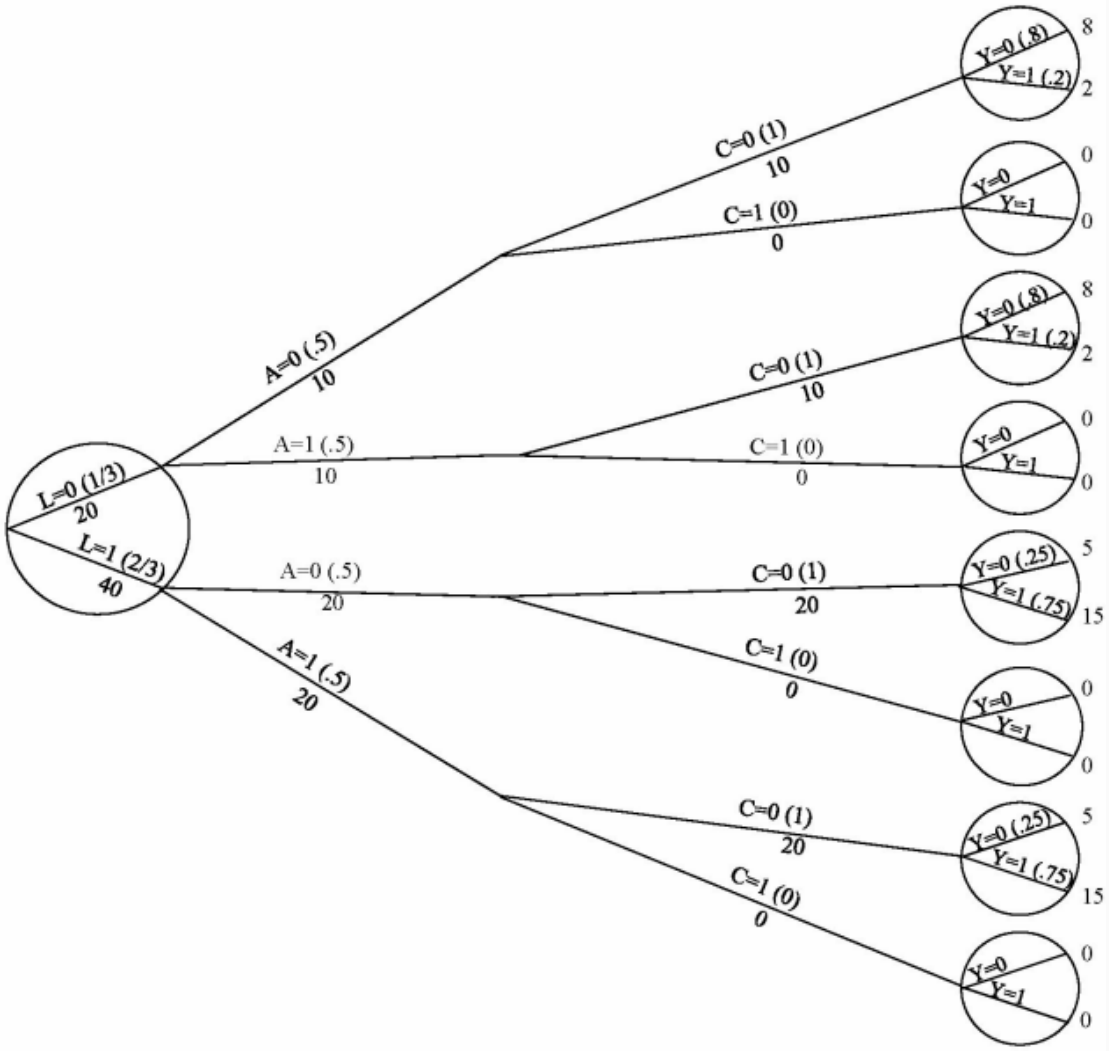

Figure 8.11

Stratification for selection bias adjustment

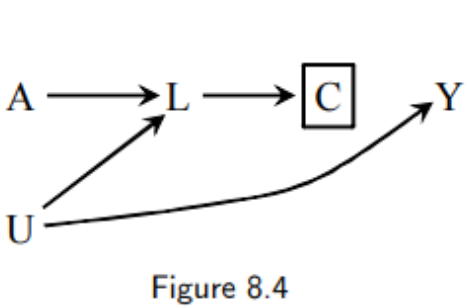

Figure 8.4

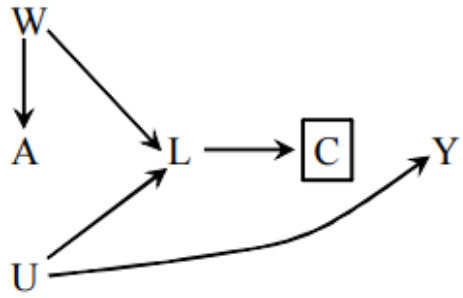

Figure 8.6

- The backdoor path (in the Figure 8.4) : $\mathrm{C} \leftarrow L \leftarrow U \rightarrow Y$
- The open path (in the Figure 8.4) : $\mathrm{A} \rightarrow L \leftarrow U \rightarrow Y$
- The conditional risk ratio :

$$\frac{P[Y=1|A=1,C=0,L=l]}{P[Y=1|A=0,C=0,L=l]} \neq 1$$

(the effect of treatment among the uncensored with $L = l$)