Standarization and The Parametric G-formula

DongYoon Yang

Seoul National University

August 9, 2021

Table of Contents

1. Introduction
2. Standarization for estimating causal effect
3. Estimating mean outcome
4. IP weighting V.S Standarization

Table of Contents

1. Introduction
2. Standarization for estimating causal effect
3. Estimating mean outcome
4. IP weighting V.S Standarization

Introduction

- Method for estimating the average causal effect
  1. IP weighting : chapter 12
  2. Standarization : This chapter

Table of Contents

1. Introduction
2. Standarization for estimating causal effect
3. Estimating mean outcome
4. IP weighting V.S Standarization

Standarization for estimating causal effect

- $Y = 1$(die), $Y = 0$(survive)
- $A = 1$(treatment), $A = 0$(no treatment)
- Under conditional exchageability, positivity and consistency, marginal counterfactual risk is

$$\mathbb{E}[Y^a] = \sum_l \mathbb{E}[Y|A = a, L = l]P(L = l) \tag{1}$$

- Estimate $\mathbb{E}[Y|A = a, L = l]$ and $P(L = l)$.

Table of Contents

1. Introduction
2. Standarization for estimating causal effect
3. Estimating mean outcome
4. IP weighting V.S Standarization

Estimating mean outcome

- Ideally, we would estimate $\mathbb{E}[Y|A=a, L=l]$ by stratification.
- Consider counfounder L is high-dimensional case.
- In this case, we could use parametric model.(e.g. linear regression)
- The next step is standarizing these means to the distribution of the confounders L for all values l.

$$\mathbb{E}[Y^a] = \sum_l \mathbb{E}[Y|A=a, L=l]P(L=l) = \mathbb{E}_L \mathbb{E}[Y|A=a, L]$$

Table 2.2

	L	A	Y
Rheia	0	0	0
Kronos	0	0	1
Demeter	0	0	0
Hades	0	0	0
Hestia	0	1	0
Poseidon	0	1	0
Hera	0	1	0
Zeus	0	1	1
Artemis	1	0	1
Apollo	1	0	1
Leto	1	0	0
Ares	1	1	1
Athena	1	1	1
Hephaestus	1	1	1
Aphrodite	1	1	1
Cyclope	1	1	1
Persephone	1	1	1
Hermes	1	1	0
Hebe	1	1	0
Dionysus	1	1	0

(a) Original

Second block: All untreated

	L	A	Y
Rheia	0	0	.
Kronos	0	0	.
Demeter	0	0	.
Hades	0	0	.
Hestia	0	0	.
Poseidon	0	0	.
Hera	0	0	.
Zeus	0	0	.
Artemis	1	0	.
Apollo	1	0	.
Leto	1	0	.
Ares	1	0	.
Athena	1	0	.
Hephaestus	1	0	.
Aphrodite	1	0	.
Cyclope	1	0	.
Persephone	1	0	.
Hermes	1	0	.
Hebe	1	0	.
Dionysus	1	0	.

(b) Untreated

Third block: All treated

	L	A	Y
Rheia	0	1	.
Kronos	0	1	.
Demeter	0	1	.
Hades	0	1	.
Hestia	0	1	.
Poseidon	0	1	.
Hera	0	1	.
Zeus	0	1	.
Artemis	1	1	.
Apollo	1	1	.
Leto	1	1	.
Ares	1	1	.
Athena	1	1	.
Hephaestus	1	1	.
Aphrodite	1	1	.
Cyclope	1	1	.
Persephone	1	1	.
Hermes	1	1	.
Hebe	1	1	.
Dionysus	1	1	.

(c) Treated

$$\widehat{\mathbb{E}}[Y^a] = \frac{1}{n}\sum_{i=1}^{n} \widehat{\mathbb{E}}[Y|A = a, L_i] \tag{2}$$

- Impute Y treated(A=1) and Y untreated(A=0) by original dataset.
- Estimate $\widehat{\mathbb{E}}[Y^{a=1}] - \widehat{\mathbb{E}}[Y^{a=0}]$

Table of Contents

1. Introduction
2. Standarization for estimating causal effect
3. Estimating mean outcome
4. IP weighting V.S Standarization

IP weighting V.S Standarization

- Modeling the treatment probability $P(A = a|L)$: IP weighting
- Modeling the outcome $\mathbb{E}[Y|A = a, L = l]$: Standarization
- Model misspecification will introduce some bias.

Definition

A doubly robust estimator in causal inference is a consistent estimator that at least one of the two models(treatment, outcome) is correct.

- To obtain a doubly robust estimate of the average causal effect,
  1. Esitmate IP weight $W^A = P(A|L)$
  2. $\mathbb{E}[Y|A = a, L = l, R]$ where $R = W^A$ if $A = 1$, $R = -W^A$ if $A = 0$.
  3. Standarize mean outcome.

- IP weighting and standarization are estimators of the g-formula, a general method for causal inference.
- We say that standarization is a plug-in g-formula estimator.