Causual Inference Note

Kirby CHEN

Academic Year 2024-2025

Contents

1 PSM Model Setup

For an individual i, the outcome depends on whether they receive a certain treatment:

$$y_i = \begin{cases} y_{1i}, & \text{if } D_i = 1 \\ y_{0i}, & \text{if } D_i = 0 \end{cases} \tag{1}$$

- D_i indicates whether individual i receives the treatment, where 1 represents treated, and 0 represents untreated.

- y_{1i} represents the outcome for individual i if treated.

- y_{0i} represents the outcome for individual i if untreated.

Given the observable covariates x_i, the probability of an individual i receiving the treatment is defined as:

$$p(x_i) = \Pr(D_i = 1 \mid x = x_i) = E(D_i \mid x_i) \tag{2}$$

Based on Equations (1) and (2), the **Average Treatment Effect on the Treated (ATT)** is given by:

$$ATT = E[y_{1i} - y_{0i} \mid D_i = 1] \tag{3}$$

$$= E[E[y_{1i} - y_{0i} \mid D_i = 1, p(x_i)]] \tag{4}$$

$$= E[E[y_{1i} \mid D_i = 1, p(x_i)] - E[y_{0i} \mid D_i = 0, p(x_i)] \mid D_i = 1] \tag{5}$$

1.1 Definition of ATE

The treatment effect is defined as a random variable:

$$y_{1i} - y_{0i} \tag{6}$$

The expected value of this effect is called the **Average Treatment Effect (ATE)**:

$$ATE = E(y_{1i} - y_{0i}) \tag{7}$$

ATE represents the expected treatment effect for a randomly selected individual from the population, regardless of whether they received the treatment.

1.2 Definition of ATT

If we consider only those who actually participated in the treatment, we define the **Average Treatment Effect on the Treated (ATT)**:

$$ATT = E(y_{1i} - y_{0i} \mid D_i = 1) \tag{8}$$

ATT measures the **average effect of the treatment for those who received it**.

1.3 Importance of ATT vs. ATE

For policymakers, **ATT is often more important** because it directly measures the effect on those who received the intervention. However, ATE and ATT are generally **not equal**.

1.4 Estimation Challenges and Selection Bias

Since we cannot observe both y_{0i} and y_{1i} for the same individual, estimating ATE or ATT is challenging. A naive comparison between the treated and untreated groups results in selection bias:

$$E(y_{1i} \mid D_i = 1) - E(y_{0i} \mid D_i = 0) \tag{9}$$

This difference consists of two terms:

$$E(y_{1i} \mid D_i = 1) - E(y_{0i} \mid D_i = 1) + E(y_{0i} \mid D_i = 1) - E(y_{0i} \mid D_i = 0) \tag{10}$$

- The first term represents the **true ATT**:

$$ATT = E(y_{1i} \mid D_i = 1) - E(y_{0i} \mid D_i = 1) \tag{11}$$

- The second term represents **selection bias**:

$$E(y_{0i} \mid D_i = 1) - E(y_{0i} \mid D_i = 0) \tag{12}$$

If selection bias is present, a direct comparison between treated and untreated individuals **does not accurately estimate ATT or ATE**.

1.5 PSM Assumptions

1.6 Common Support Assumption

For any possible value of x_i, the propensity score must satisfy:

$$0 < p(x_i) < 1 \tag{13}$$

This assumption ensures that there is **overlap between the treated and control groups**, making it possible to find comparable units.

1.7 Balancing Assumption

$$D_i \perp (y_{1i}, y_{0i}) \mid p(x_i) \tag{14}$$

This assumption states that **conditional on the propensity score $p(x_i)$, treatment assignment is as good as random**. That is, for a given $p(x_i)$, there are no systematic differences between the treatment and control groups, meaning the treatment effect is entirely due to the treatment itself.

1. Propensity Score Theorem The theorem states that if treatment assignment satisfies:

$$(y_0, y_1) \perp D \mid x \tag{15}$$

then it also holds that:

$$(y_0, y_1) \perp D \mid p(x) \tag{16}$$

Proof: Since D is a binary variable, we need to show that:

$$P(D = 1 \mid y_0, y_1, p(x)) \tag{17}$$

2

is independent of y_0, y_1. Using the expectation rule:

$$\begin{aligned}
P(D = 1 \mid y_0, y_1, p(x)) &= E[D \mid y_0, y_1, p(x)] \\
&= E_{y_0, y_1, x}[E(D \mid y_0, y_1, x) \mid y_0, y_1, p(x)] \quad \text{(by iterated expectation)} \\
&= E_{y_0, y_1, x}[E(D \mid x) \mid y_0, y_1, p(x)] \quad \text{(by ignorability assumption)} \\
&= E_{y_0, y_1, x}[p(x) \mid y_0, y_1, p(x)] \\
&= p(x)
\end{aligned}$$

Since the probability only depends on $p(x)$, this confirms that $(y_0, y_1) \perp D \mid p(x)$.

1.8 2. Overlap Assumption (Common Support)

To ensure valid matching, every possible value of x should exist in both treatment and control groups. This is known as the **overlap assumption** or **matching assumption**.

Definition (Overlap Assumption):
$$0 < p(x) < 1, \quad \forall x \tag{18}$$

This ensures that there is a **common support** where treated and control groups have comparable observations.

Implications:

- Ensures that treated and control groups share a common range of propensity scores.

- Helps avoid extrapolation in estimating treatment effects.

- Also called the **matching assumption**.

Graphical Representation (Figure 28.1): A graphical illustration shows the common support region, where both treated and control groups have overlapping propensity scores. Observations outside this range are often dropped to improve match quality.

1.9 3. Steps for Estimating Treatment Effects Using Propensity Score Matching

The general procedure for propensity score matching follows three steps:

1. **Select Covariates x_i:** Include all relevant variables that influence both treatment assignment D and potential outcomes (y_0, y_1). This ensures the **ignorability assumption** holds.

2. **Estimate the Propensity Score $p(x)$: Rosenbaum and Rubin (1985)** suggest using a logit model:

$$p(x) = Pr(D = 1 \mid x) \tag{19}$$

   Including higher-order terms and interactions improves matching quality.

3. **Perform Matching Based on $p(x)$:** If the model is correctly specified, the **covariate distributions should be similar** between treated and control groups post-matching.

   Data Balancing Check:

$$\bar{x}_{\text{treat}} \approx \bar{x}_{\text{control}} \tag{20}$$

   Ensuring that the mean covariate values are similar across matched groups confirms that matching was successful.