

Notes of Causal Inference

Two views: Regression Analysis and Potential Outcomes

Yue Wu

1 Introduction

Mathematically equivalent estimators and different views from economists and statisticians. We talk about Binary treatment (and possibly extend to continuous setting).

Similarity

Focusing on finite population of N units under study, providing inferences solely about this finite population.

Differences

For regression, the randomness comes from the relation of error term and D_i . For each unit i , D_i is non-stochastic, ε_i is a random variable and so is Y_i . Parameter of interest is $\mathbb{E}[Y_i | D_i = 1] - \mathbb{E}[Y_i | D_i = 0]$.

For potential outcomes, the randomness comes from the relation of assignment mechanism and potential outcomes. For each unit i , $Y_i(0), Y_i(1)$ are non-stochastic, D_i is a random variable and so is Y_i . Parameter of interest is $\tau = N^{-1} \sum_{i=1}^N (Y_i(1) - Y_i(0))$.

2 The OLS/DIM estimator

A general setting of data we have is $\{D_i, Y_i\}_{i=0}^N$, where D_i is binary, and the general goal is to make inference about the effect of D on Y . And OLS/DIM estimator is defined as

$$\hat{\tau} = \bar{Y}_1 - \bar{Y}_0$$

2.1 Regression View

Assumptions

Here we just introduce the basic assumptions of a regression model.

$$Y_i = \beta_0 + \tau D_i + \varepsilon_i$$

where the error terms ε_i are uncorrelated with mean 0 and variance σ^2 .

What does the assumptions mean?

1. Constant treatment effect of D on Y . By saying 'constant', we mean that for each unit i under study, the effect of D on Y is the same, shared constant τ .

The effect of D on Y can be interpreted as the change in the conditional expectation of Y associated with a one-unit change in D .

If D is binary, the effect of D on Y can be expressed as

$$\tau = \mathbb{E}[Y_i \mid D_i = 1] - \mathbb{E}[Y_i \mid D_i = 0]$$

2. (?)Independent error terms. This assumption allows us to write:

$$\varepsilon_i \mid \mathbf{D} \quad \leftrightarrow \quad \varepsilon_i \mid D_i$$

$$\mathbb{E}[\varepsilon_i \mid \mathbf{D}] = 0 \quad \leftrightarrow \quad \mathbb{E}[\varepsilon_i \mid D_i] = 0$$

$$Y_i \mid \mathbf{D} \quad \leftrightarrow \quad Y_i \mid D_i$$

$$\mathbb{E}[Y_i \mid \mathbf{D}] \quad \leftrightarrow \quad \mathbb{E}[Y_i \mid D_i]$$

Estimand

The parameter of interest is the coefficient τ .

Property of OLS estimator

Randomness comes from the random error term, and $\hat{\tau}$ is an unbiased estimator of τ because we assume zero expectation of error.

2.2 Potential outcomes View

Assumptions

For each individual, the pair of potential outcomes are $\{Y_i(1), Y_i(0)\}$.

Assumption 1. Stable Unit Treatment Value Assumption(SUTVA) [!!!cite]

If $D_i = D'_i$, then $Y_i(\mathbf{D}) = Y_i(\mathbf{D}')$.

Assumption 2. Random Assignment. The treatment assignment D_i is random:

$$Pr(\mathbf{D} = c) = Pr(\mathbf{D} = c')$$

for all c and c' such that $\iota^T c = \iota^T c'$, where ι is the N-dimensional column vector with all elements equal to one.

What does the assumptions mean?

SUTVA allows us to write:

$$\varepsilon_i \mid \mathbf{D} \quad \leftrightarrow \quad \varepsilon_i \mid D_i$$

$$\mathbb{E}[\varepsilon_i \mid \mathbf{D}] = 0 \quad \leftrightarrow \quad \mathbb{E}[\varepsilon_i \mid D_i] = 0$$

$$Y_i \mid \mathbf{D} \quad \leftrightarrow \quad Y_i \mid D_i$$

$$\mathbb{E}[Y_i \mid \mathbf{D}] \quad \leftrightarrow \quad \mathbb{E}[Y_i \mid D_i]$$

Estimand The parameter of interest is the Average Treatment Effect (ATE)

$$\tau = N^{-1} \sum_{i=1}^N (Y_i(1) - Y_i(0))$$

Property of OLS/DIM estimator Randomness comes from the assignment mechanism. And $\hat{\tau}$ is an unbiased estimator of τ because we assume the treatment assignment D_i is random with regard to potential outcomes.

3 The IV estimator

A general setting of data we have is $D_i, Z_i, Y_{i=1}^N$, where D_i, Z_i are binary variables, and the general goal is to make inference about the effect of D on Y . Z_i is called the instrument variable.

The IV estimator is defined as the ratio of sample covariances. [!!!cite]

3.1 Regression View

Assumptions

3.2 Potential outcomes View

Assumptions

4 Summary

A proofs