Microeconometrics Module

Lecture 2: Miscellaneous Concepts

Swapnil Singh

Lietuvos Bankas and KTU

Course Link

Introduction

- Clarification of some concepts
    1. Definition of *estimand*, *estimate*, and estimator
    2. Different causal estimands
    3. Understanding identification

Estimand, Estimate, and Estimator

- **Estimand**: A parameter or a function of the population that we aim to measure.
  - *True value* that we are interested in knowing
  - Example: The average income of all households in a country.
- **Estimator**: A rule or a formula that tells us how to calculate estimand based on sample data.
- **Estimate**: The actual value obtained by applying the estimator to a given set of sample data.
  - Example: The calculated value of $50,000 as the average income from the sample data.

Causal Estimands

Different Causal Estimands

- Three types of estimands are interesting for us:
  1. Average treatment effect i.e. τ_{ATE}
  2. Average treatment effect on the treated i.e. τ_{ATET}
  3. Conditional average treatment effect i.e. τ_{CATE}
- **Average treatment effect** i.e. τ_{ATE}

$$\begin{aligned} \tau_{ATE} &= \mathbb{E}(\tau_i) \\ &= \mathbb{E}[Y_i(1) - Y_i(0)] \end{aligned}$$

  where $\mathbb{E}(\cdot)$ is the population operator
- Note that until now we are not talking about the type of data we have

Different Causal Estimands

- **Average treatment effect on the treated** i.e. τ_{ATET}

$$
\begin{aligned}
\tau_{ATET} &= \mathbb{E}(\tau_i | D_i = 1) \\
&= \mathbb{E}[Y_i(1) - Y_i(0) | D_i = 1] \\
&= \underbrace{\mathbb{E}[Y_i(1) | D_i = 1]}_{observed} - \mathbb{E}[Y_i(0) | D_i = 1]
\end{aligned}
$$

Different Causal Estimands

- **Conditional average treatment effect** i.e. τ_{CATE}

$$\begin{aligned} \tau_{CATE}(x) &= \mathbb{E}(\tau_i | X_i = x) \\ &= \mathbb{E}[Y_i(1) - Y_i(0) | X_i = x] \end{aligned}$$

where X_i is some observed characteristic

Understanding Identification

Types of identification

- Two types of identification[1]:
  1. Point identification, or as Prof. Beatrice Cherrier calls in her X/Twitter post Ashenfelter-Angrist-Lalonde-Card-Krueger approach
  2. Partial identification, see Tamer (2010)[2]
- This module: point identification
- But, let's first dig into understanding the meaning of identification

[1] Lewbel, A. (2019). The Identification Zoo: Meanings of Identification in Econometrics. Journal of Economic Literature, 57 (4): 835-903.
[2] Tamer, E. (2010). Partial identification in econometrics. Annu. Rev. Econ., 2(1), 167-195.

Understanding identification

- Let θ denote an unknown parameter or a *set* of unknown parameters
- Our objective is to learn/estimate θ

Understanding identification

- Let θ denote an unknown parameter or a *set* of unknown parameters
- Our objective is to learn/estimate θ

Ex: In linear regression models, i.e. $y = \alpha + \beta x + \epsilon$, θ can be α, β

Understanding identification

- Let θ denote an unknown parameter or a *set* of unknown parameters
- Our objective is to learn/estimate θ

Ex: In linear regression models, i.e. $y = \alpha + \beta x + \epsilon$, θ can be α, β

Ex: τ_{ATE}

- Identification means, if we knew the population the data is drawn from, can we infer information about θ
- Another way to think:
  - variation in $\theta \implies$ variation in the distribution of data

Point identification

- We want to identify and estimate θ
- Assume, there exists ϕ
  - ϕ can be learned from the data

Point identification

- We want to identify and estimate θ
- Assume, there exists ϕ
  - ϕ can be learned from the data

Ex. of ϕ:
distribution function,
conditional means,
quantiles, etc.

Point identification

- We want to identify and estimate θ
- Assume, there exists ϕ
  - ϕ can be learned from the data

Ex. of ϕ: distribution function, conditional means, quantiles, etc.

- Essentially, we will leverage ϕ to get to know θ
- To bridge this knowledge, we need a model
- Model puts restriction on what values ϕ can take

Point Identification: Example 1

- Suppose there are two variables y and x
- Our model is that there is a **linear** relationship between y and x, i.e.

$$y = x\theta + \varepsilon, \quad \mathbb{E}(x^2) \neq 0, \quad \mathbb{E}[ex] = 0$$

- Let's say we can learn second moments of (y, x) i.e. we know what ϕ is
- Then, we know

$$\theta = \frac{\mathbb{E}(xy)}{\mathbb{E}(x^2)}$$

- θ is point identified
- However, if we were not able to calculate second moments, then θ is not point identified

Identification of treatment effect

Strong Ignorability

Definition
D_i is strongly ignorable conditional on $\boldsymbol{X}_i$ if

1. Potential outcomes are independent of treatment, conditional on X_i i.e.

$$Y_i(1), Y_i(0) \perp\!\!\!\perp D_i | X_i$$

2. The treatment is not too rare or common, i.e.

$$\exists \epsilon > 0, \text{ s.t. } \epsilon < Pr(D_i = 1 | X_i) < 1 - \epsilon$$

ATE Point Identification

Theorem
If D_i is strongly ignorable conditional on X_i, then

$$\begin{aligned} ATE &= \mathbb{E}(\tau_i) \\ &= \int_{x \in \Omega(X_i)} \left[\mathbb{E}(Y_i | D_i = 1, X_i = x) - \mathbb{E}(Y_i | D_i = 0, X_i = x) \right] f(x) dx \end{aligned}$$

Next

- Endogeneity

Questions?