Spring 2022

Quantitative Analysis
Chung-Ming Kuan
Jui-Chung Yang

Problem Set 4

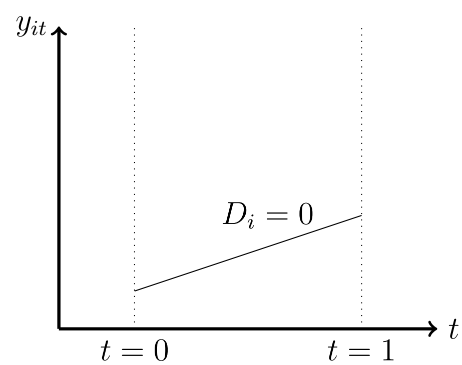
Due: 03/21

Part One: Hand-Written Exercise

1. The government is going to bulid a new railway. Tom wants to understand if a city's GDP will be influenced by the passing of the new railway. He uses a DID model below:

$$y_{it} = \alpha + \beta t + \gamma D_i + \varphi_i + \delta(D_i t) + u_{it},$$

where y_{it} denotes the GDP of city i at time t, and all the other notations are defined the same as in our lecture slides. Suppose both γ and δ are greater than 0, and the regression line for cities which are not passed by the railway ($D_i = 0$) is:



Draw the line for the cities passed by the railway ($D_i = 1$) on the plot above, and indicate β, γ and δ.

2. Consider the model

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

that satisfies the Modern Assumptions. Moreover, let $\mathrm{Var}(x_1) = \sigma^2_{x_1}$ and $\mathrm{Cov}(x_1, x_2) = \sigma_{x_1 x_2}$. Suppose we exclude an important variable x_2 and obtain the corresponding OLS estimator $\tilde{\beta}_1$. That is, we obtain $\tilde{\beta}_1$ from the model $y_i = \beta_0 + \beta_1 x_{1i} + u_i$.

(a) Is $\tilde{\beta}_1$ consistent for β_1?

(b) As the sample size $n \to \infty$ and $\sigma_{x_1 x_2} > 0$, does $\tilde{\beta}_1$ over- or under-estimate β_1? By how much?

(c) As the sample size $n \to \infty$ and $\sigma_{x_1 x_2} > 0$, does $\sqrt{n}(\tilde{\beta}_1 - \beta_1)$ converge to a normal distribution, or any other distributions?

3. Answer the following questions with "True" or "False" and briefly explain them. All notations are defined as in our lecture slides.

   (a) A biased estimator must be inconsistent.

   (b) An unbiased estimator must be consistent.

   (c) $\hat{\boldsymbol{\beta}}$ is inconsistent when $\text{Var}(\mathbf{x_i}\varepsilon_i) = \mathbf{V}$ doesn't exist.

Part Two: Monte Carlo Simulation

- Simulation design:
  - Sample sizes N:
    (i) 10 (ii) 500
  - Number of replications: 1000
  - Data generating process (DGP):
    (i) $y_i \sim N(0,1)$ (ii) $y_i \sim t(4)$ (iii) $y_i \sim t(1)$
  - The statistics:

$$M_N = \frac{1}{\hat{\sigma}_N \sqrt{N}} \sum_{i=1}^{N} \phi(y_i), \quad \text{where} \quad \hat{\sigma}_N^2 = \frac{1}{N} \sum_{i=1}^{N} \left(\phi(y_i) - \frac{1}{N} \sum_{i=1}^{N} \phi(y_i) \right)^2,$$

    with the moment functions:
    (i) $\phi(y_i) = y_i$ (ii) $\phi(y_i) = y_i^3$ (iii) $\phi(y_i) = sin(y_i)$ (iv) $\phi(y_i) = cos(y_i)$

1. For the total of 24 different ways to construct M_N, please plot their corresponding histograms using 1000 replications. Open a new window and combine the 24 graphs on a single plot and place them as 6×4.

2. Please compute the empirical frequencies of the events: $M_N^2 > 3.8414588$ and $M_N^2 > 6.6348966$ for each simulations. Record them under their corresponding graphs. Check if the frequencies are, respectively, sufficiently close to the 5% and 1% nominal levels.

3. Please add the Gaussian kernel density estimate (KDE) of M_N as well as the probability density function (PDF) of $N(0,1)$ for each simulation graph.