Spring 2020

Quantitative Analysis
Chung-Ming Kuan

Problem Set 4

Due: 3/30

Part One: Hand-Written Exercise

1. Answer the following questions with "True" or "False". Briefly explain if your answer is True, provide a counter example if your answer is False.

   (a) A biased estimator must be inconsistent.

   (b) An unbiased estimator must be consistent.

   (c) Central Limit Theorem(CLT) always works no matter what the original distribution of random sample is as long as the sample size is sufficiently large.

   (d) If an estimator is weakly consistent, then with proper normalization, it is also asymptotically normal distributed.

   (e) If $\sqrt{n}(\hat{\beta} - \beta_o)$ is asymptotically normal distributed, then $\hat{\beta}$ is weakly consistent for β_o.

2. Verify the statement on slide 18, Lecture 4. That is, for the simple linear regression: $y_i = b_0 + b_1 x_i + u_i$, we have $\hat{\beta}_0 \xrightarrow{p} b_0$ under Modern Assumption 0.

3. Consider the model without intercept

$$y_i = \beta_1 x_i + u_i$$

that satisfies the Modern Assumptions and $\mathbb{E}(x_i) = 0$, $\text{Var}(x_i) = \sigma_x^2$. Now suppose x_i can't be observed directly, instead, we observe w_i, which is x_i plus a random error ϵ_i:

$$w_i = x_i + \epsilon_i,$$

with $\mathbb{E}(\epsilon_i) = \mathbb{E}(\epsilon_i x_i) = \mathbb{E}(\epsilon_i u_i) = 0$, and $\text{Var}(\epsilon_i) = \sigma_\epsilon^2$. So instead of estimating the model $y_i = \beta_1 x_i + u_i$, which is impossible due to the fact that x_i is unobservable, we now estimate

$$y_i = \gamma_1 w_i + u_i,$$

which is known as the model with *measurement errors.*

   (a) Please find the probability limit of the OLS estimator $\hat{\gamma}_1$.

(b) Is $\hat{\gamma}_1$ consistent for β_1? If not, then as the sample size $n \to \infty$, does $\hat{\gamma}_1$ over- or under-estimates β_1? By how much?

(c) As $n \to \infty$, what would happen to $\hat{\gamma}_1$ if $\sigma_x^2 \to \infty$? Please briefly explain the intuition behind your results.

(d) As $n \to \infty$, what would happen to $\hat{\gamma}_1$ if $\sigma_\epsilon^2 \to \infty$? Please briefly explain the intuition behind your results.

4. Consider the model

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

that satisfies the Modern Assumptions. Moreover, let $\text{Var}(x_1) = \sigma_{x_1}^2$ and $\text{Cov}(x_1, x_2) = \sigma_{x_1 x_2}$. Suppose we exclude an important variable x_2 and obtain the corresponding OLS estimator $\tilde{\alpha}_1$. That is, we obtain $\tilde{\alpha}_1$ from the model $y_i = \alpha_0 + \alpha_1 x_{1i} + u_i$.

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

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

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

5. Verify the statement on slide 41, Lecture 4. That is, write down the 3×3 matrix $\mathbf{R}\tilde{\mathbf{D}}\mathbf{R}'$ using the notation d_{ij}, where

$$\mathbf{R} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}$$

and

$$\tilde{\mathbf{D}} = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1(k+1)} \\ d_{21} & d_{22} & \cdots & d_{2(k+1)} \\ \vdots & \vdots & \ddots & \vdots \\ d_{(k+1)1} & d_{(k+1)2} & \cdots & d_{(k+1)(k+1)} \end{bmatrix}.$$

Part Two: Monte Carlo Simulation (Due: 3/30)

- 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) $\phi(y_i) = cos(y_i)$

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

- Please compute the empirical frequencies of the events: $M_N > Z_{.99} = 2.3$ and $M_N > Z_{.95} = 1.6$ for each simulations(Z_q is the q^{th} quantile of standard normal distribution). Record them under their corresponding graphs, see if the frequencies are, respectively, sufficiently close to the 5% and 1% nominal levels.

- 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 graphs.