Spring 2020

Quantitative Analysis
Chung-Ming Kuan

Problem Set 1

Due: 3/9

Part One: Hand-Written Exercise

1. Verify the statement in slide 24, Lecture 1. That is, suppose $y_i = \beta_0 + \beta_1 x_{1i} + u_i$, please show that the OLS estimators. (20 points)

   (a) $\hat{\beta}_1 = \frac{\sum_{i=1}^n (y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$

   (b) $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

2. Consider the following regression models:

$$\begin{aligned} &\text{Model A: } y_i = \beta_0 + \beta_1 x_i + u_i \\ &\text{Model B: } y_i = \alpha_0 + \alpha_1 (x_i - \bar{x}) + v_i \end{aligned}$$

   where $\bar{x} = \frac{1}{n} \sum x_i$, and $\mathrm{Var}(y_i) = \sigma^2$ (40 points)

   (a) Find the OLS estimators of β_0 and α_0. Are they identical? Are their variances identical? If not, which variance is larger?

   (b) Find the OLS estimators of β_1 and α_1. Are they identical? Are their variances identical? If not, which variance is larger?

3. Consider the model $y_i = \beta_0 + \beta_1 x_i + u_i$ with $\mathrm{Var}(y_i) = \sigma^2$. Under the Classical Assumptions, The OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased. Let $\tilde{\beta}_1$ be the OLS estimator of β_1 by assuming the intercept is zero. That is, $\tilde{\beta}_1$ is the obtained under the assumption $\beta_0 = 0$. (40 points)

   (a) Calculate $\mathbb{E}(\tilde{\beta}_1)$ in terms of x_i, β_0, and β_1.

   (b) If $\beta_0 \neq 0$, is $\tilde{\beta}_1$ unbiased?

   (c) Calculate the variance of $\tilde{\beta}_1$.

   (d) Compare between $\mathrm{Var}(\tilde{\beta}_1)$ and $\mathrm{Var}(\hat{\beta}_1)$. Is it true that $\mathrm{Var}(\tilde{\beta}_1) \leq \mathrm{Var}(\hat{\beta}_1)$ in general?

   (e) Does the result in (d) violates the Gauss-Markov Theorem, which states that $\hat{\beta}_1$ should have the smallest variance. Explain.

Part Two: Computer Exercise

1. (25 points)

(a) Let $x = c(1 : 150)$

(b) Select the number in x that is greater than 135 or smaller or equal to 5.

(c) Select the number in x that is greater than 70 and smaller than 90.

(d) Select the number in x that is divisible by 4 and 5

2. (25 points)

   (a) Create a series containing 100 number "1" by using **rep()**

   (b) Create a series containing 100 number "1" followed by 50 number "8" by using **rep()** and **c()**

   (c) Create a series **1,4,7,...** with 100 repeat times, which means there are total 300 elements in that series

3. (25 points)

   (a) Draw 150,000 observations from standard normal distribution and name it as "X"

   (b) Evaluate the mean, median, max, min, and variance of X.

   (c) Randomly select 5,000 subsamples from X without replacement, call it Y and calculate its mean and variance.

   (d) Randomly select 5,000 subsamples from X with replacement, call it Z and calculate its mean and variance.

   (e) Find the 45^{th} percentile in X. Also, find the number z such that $Pr(a \leq z) = 0.45$, where $a \sim N(0, 1)$.

   (f) Find the probability of drawing $x \in X$ such that $x \in (-0.55, 1.25]$. Also, find the probability of drawing a, where $a \sim N(0, 1)$ such that $a \in (-0.55, 1.25]$.

4. (25 points)

   Let $\mathbf{X} = \begin{bmatrix} 7 & 2 & 3 \\ 4 & 6 & 7 \\ 9 & 2 & 0 \\ 0 & 9 & 0 \\ 5 & 3 & 5 \end{bmatrix}$ and $\mathbf{Y} = \begin{bmatrix} 6 \\ 2 \\ 4 \\ 2 \\ 1 \end{bmatrix}$.

   (a) Please construct the OLS estimator $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$.

   (b) Given a new observation $x^* = (0, 4, 3)'$, please calculate $\hat{y}$.