EE325 Introductory Econometrics (Section 1 semester 1/2020)

Assignment 1

Instruction: Write your answer in either paper or digital paper. However, if you write on paper, please scan it and save as a PDF file. Submission is via BE-Moodle as a PDF file for both cases. (Please keep the file below 10MB as that is the maximum per file capacity for student.)

Due date: Tuesday, September 1,2020 (Before class starts at 11.00 A.M.)

1. Explain the difference between cross-sectional, time-series, and panel data.

2. Let X and Y be discrete random variables with the joint PDF displayed in the following table. Answer the following questions.

		X		
		1	2	3
Y	2	0.1	0.2	0.1
	4	0.3	0.2	**a**

2.1 Find the value of **a** in the table and explain why.

2.2 Find $E(X)$ and $E(Y)$.

2.3 Find $Var(X)$ and $Var(Y)$.

2.4 Find $E(X|Y=4)$ and $Var(Y|X=3)$.

2.5 Define Z as $X-Y$, Find $Var(Z)$.

2.6 Find the $E(E(Y|X))$ and show that $E\big(E(Y|X)\big)=E(Y)$.

3. Let X be a continuous random variable, the PDF is given by

$$f(x)=\begin{cases} \mathrm{a}+\mathrm{bx}^2 & ; 0\le x\le 1 \\ 0 & ; elsewhere \end{cases}$$

If the expected value $E(X)=\frac{3}{5}$, find the value of a and b.

4. Let X and Y be continuous random variables, the PDF is given by

$$f(x,y)=\begin{cases} x^2+\dfrac{xy}{k} & ; 0\le x\le 1\,, 0\le y\le 2 \\ 0 & ; elsewhere \end{cases}$$

Answer the following questions.

4.1 Find the value of k.

4.2 Find $P\left(\frac{1}{4}<X<\frac{1}{2}, Y>1\right)$.

4.3 Find $P\left(X>\frac{1}{2}|Y=2\right)$.

1. Explain the difference between cross-sectional, time-series, and panel data

- cross-sectional collected at one point or a period of time it collected by many subjects
- Panel data collected overtime not at one point at a time
- Time-series collected in a data point in time order but the difference is time factor

2. Let X and Y be discrete random variables with the joint PDF displayed in the following table. Answer the following questions.

		X		
		1	2	3
Y	2	0.1	0.2	0.1
	4	0.3	0.2	a $= 0.1$

2.1 Find the value of **a** in the table and explain why.

2.2 Find $E(X)$ and $E(Y)$.

2.3 Find $Var(X)$ and $Var(Y)$.

2.4 Find $E(X|Y=4)$ and $Var(Y|X=3)$.

2.5 Define Z as $X-Y$, Find $Var(Z)$.

2.6 Find the $E(E(Y|X))$ and show that $E(E(Y|X)) = E(Y)$.

2.1) $a = 1$ because sum of all possibilities $= 1$

2.2) $(1)(0.1+0.3) + (2)(0.4) + 3(0.2) = 1.8 = E(x)$

$E(Y) = (2)(0.4) + 4(0.6) = 0.8 + 2.4 = 3.2 = E(Y)$

2.3) $var(x) = \Sigma(x_i - \mu)^2 \cdot f(x_i)$

$= (1-1.8)^2 \cdot 0.4 + (2-1.8)^2 \cdot 0.4 + (3-1.8)^2 \cdot 0.2$

$= (0.256) + 0.016 + 0.288 = 0.56$

$var(Y) = (\Sigma x_i - \mu)^2 \cdot f(x_i)$

$= (2-3.2)^2 \cdot 0.4 + (4-3.2)^2 \cdot 0.6$

$= 0.576 + 0.384 = 0.96$

2.4) $E(X \mid Y=4) = \left(\frac{0.3}{0.6}\right) \cdot 1 + \left(\frac{0.2}{0.6}\right) \cdot 2 + \left(\frac{0.1}{0.6}\right) \cdot 3$

$$= \frac{1}{2} + \frac{2}{3} + \frac{3}{6} = \frac{3+4+3}{6} = \frac{10}{6}$$

2.5) $Z = X - Y$

$$\mathrm{Var}(Z) = \mathrm{Var}(X - Y)$$

$$= \mathrm{Var}(X) - \mathrm{Var}(Y)$$

$$= 0.56 - 0.96 = -0.4$$

2.6) $E(Y \mid X) = \sum y f(y \mid x) = \sum y f(y)$

$$= E(Y)$$

$$\therefore E(Y \mid X) = E(E(Y))$$

3)

3. Let X be a continuous random variable, the PDF is given by

$$f(x) = \begin{cases} a + bx^2 & ; 0 \le x \le 1 \\ 0 & ; elsewhere \end{cases}$$

If the expected value $E(X) = \frac{3}{5}$, find the value of a and b.

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 a + bx^2\,dx = a + \frac{bx^3}{3} = \boxed{a + \frac{b}{3} = 1}$$

$$\frac{3}{5} = \int_0^1 x(a + bx^2)\,dx$$

$$\frac{3}{5} = \int_0^1 ax + bx^3\,dx$$

$$\frac{ax^2}{2} + \frac{bx^4}{4}\,dx$$

$$\frac{3}{5} = \frac{a}{2} + \frac{b}{4}$$

$$3a + b = 3 \quad — ①$$

$$12 = 10a + 5b \quad — ②$$

$$b = 3 - 3a \quad — ①'$$

นำ ①' ใส่ ②

$$12 = 10a + 5(3 - 3a)$$

$$12 = 10a + 15 - 15a$$

$$12 = -5a + 15$$

$$-3 = -5a$$

$$\frac{3}{5} = a$$

$$\therefore b = \frac{6}{5}$$

4. Let X and Y be continuous random variables, the PDF is given by

$$f(x,y) = \begin{cases} x^2 + \dfrac{xy}{k} & ; 0 \le x \le 1 , 0 \le y \le 2 \\ 0 & ; elsewhere \end{cases}$$

Answer the following questions.

4.1 Find the value of k.

4.2 Find $P\left(\frac{1}{4} < X < \frac{1}{2}, Y > 1\right)$.

4.3 Find $P\left(X > \frac{1}{2} | Y = 2\right)$.

4.4 Are X and Y independent of each other? Show a proof to your answer.

4.5 Find the correlation coefficient of X and Y (ρ_{XY}).

x : $\int_0^1 f(x)\, dy$

y : $\int_0^2 f(y)\, dx$

$\int_0^1 x^2 + \frac{xy}{k}\, dy \rightarrow \frac{1}{k}($

5. Let $X \sim N(\mu, \sigma^2)$, given that you have two estimators in hand which are

(1) $\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$

(2) $\hat{X} = \frac{n}{(n-2)^2} \sum_{i=1}^{n} X_i$

Answer the following questions.

5.1 Show the bias of each estimator and explain which estimator is unbiased.

5.2 Which estimator is more efficient? Show the answer with a proof.

5.3 Which estimator that you pick to represent the population and why?

5.1) $E(\bar{x}) = E\left(\frac{x_i}{n}\right)$

$= \frac{1}{n} E[x_1 + x_2 + \dots + x_n]$

$= \frac{n\mu}{n} = \mu$ = bias

$E(\hat{x}) = E\left(\frac{n}{(n-2)^2} \cdot \Sigma x_i\right)$

$= \frac{n}{(n-2)^2} \cdot n\mu$

$= \frac{\mu n^2}{(n-2)^2}$

$= \frac{n}{(n-2)} - \frac{\mu n}{(n-2)^2}$ = bias

5.2) $var(\bar{x}) = var\left(\frac{x_i}{n}\right)$

$= \frac{1}{n^2} \cdot \cancel{var(x_i)}$

$= \frac{1}{n^2} \cdot \sigma^2 n = \frac{\sigma^2}{n}$

$Var(\hat{x}) = Var\left(\frac{n}{(n-2)^2}\right)$

$= \frac{n^2}{(n-2)^4} var\cancel{\left(\frac{1}{(n-2)^2}\right)}$ $n\sigma^2$

$Var(\hat{x}) = \frac{n^3 \sigma^2}{(n-2)^4}$

more efficient

5.3) I choose $var(\bar{x}) < var(\hat{x})$ because it is more efficient

4.4 Are X and Y independent of each other? Show a proof to your answer.

4.5 Find the correlation coefficient of X and Y (ρ_{XY}).

5. Let $X \sim N(\mu, \sigma^2)$, given that you have two estimators in hand which are

(1) $\bar{X} = \sum_{i=1}^{n} \frac{X_i}{n}$

(2) $\hat{X} = \frac{n}{(n-2)^2} \sum_{i=1}^{n} X_i$

Answer the following questions.

5.1 Show the bias of each estimator and explain which estimator is unbiased.

5.2 Which estimator is more efficient? Show the answer with a proof.

5.3 Which estimator that you pick to represent the population and why?