Midterm Exam

(April 20, 2021, **1 hour 15 minutes**)

Econometrics (Spring 2022)

Professor: Wonmun Shin

Part I

True or False

(Total 30 points) Read each statement below carefully. Write a **T** if you think the statement is True. Write an **F** if you think the statement is false. Do not forget to state a reason precisely if you answer F. (Grading guide: If your answer (with a reason in case F) is correct, you will get 5 points. If your answer is partially correct (*i.e.* F without a reason, F with an incorrect reason, etc.), you will get partial points.)

1. The central limit theorem (CLT) is used to show that the OLS estimators are unbiased and have the smallest variance among the group of linear and unbiased estimators.

2. In the simple regression model, suppose we multiply each Y_i value by a constant, say 2. Then, it will not change the residuals and fitted values of Y.

3. If the sample correlation coefficient between (X_i, Y_i) for $i = 1, 2, \cdots, n$ is positive, then the OLS estimate of the slope coefficient in the simple regression of Y on X is also positive.

4. The OLS estimator $\hat{\beta}_1$ and $\hat{\beta}_2$ follow the normal distribution asymptotically only if the regression error follows the normal distribution.

5. If you add one more additional independent variable to your linear regression model, the R^2 can increase.

6. At the same significance level, the probability of type I error tends to be smaller when doing a two-sided test than when doing a one-sided test.

Part II

Short Questions

1. **(10 points)** In the simple linear regression $Y_i = \beta_1 + \beta_2 X_i + e_i$, we can obtain the OLS estimators $\hat{\beta}_1$ and $\hat{\beta}_2$. *Student A* considers the newly-defined dependent variable and independent variable by adding constants to Y_i and X_i. Specifically, she defines $Y_i^* = Y_i + c$ and $X_i^* = X_i + d$ (c and d are non-zero constants, and $c \neq d$). As a result, she runs the following regression:

$$Y_i^* = \alpha_1 + \alpha_2 X_i^* + u_i$$

Express the OLS estimators $\hat{\alpha}_1$ and $\hat{\alpha}_2$ in terms of $\hat{\beta}_1$, $\hat{\beta}_2$, c, and d.

2. **(10 points)** From a sample of 100 observations, *Student B* obtained the following results:

$$\sum X_i = 500, \quad \sum Y_i = 400, \quad \sum X_i Y_i = 4,000, \quad \sum X_i^2 = 7,500$$

Suppose that *Student B* wants to obtain the OLS estimates for a simple regression model, $Y_i = \beta_1 + \beta_2 X_i + e_i$. Using the above results, calculate $\hat{\beta}_1$ and $\hat{\beta}_2$.

3. **(10 points)** Consider the following regression output:

$$\widehat{\text{Dust}} = \underset{(0.05)}{0.08} + \underset{(0.003)}{0.006}\ \text{Temp}$$

(standard errors in parentheses)

where *Dust* is density of fine dust (unit: $\mu g/m^3$), *Temp* is temperature (unit: ℃). That is, the above regression analyzes the relationship between the density of fine dust and the temperature. (i) In order to check the significance of the point estimate, *Student C* does the hypothesis testing: $H_0 : \beta_2 = 0$ against $H_1 : \beta_2 \neq 0$ at 5% significance level. Is $\hat{\beta}_2$ significant? (ii) She also tests $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 > 0$ at 5% significance level. Can she reject the null hypothesis? ***(Note: If you need, please refer to the probability table on the last page.)***

Part III

Long Questions

1. **(Total 20 points)** Suppose you have the following estimated models of relationships between consumption (C), income (INC), and time (t):

$$\textbf{Model A}: \widehat{C} = \underset{(2.73)}{0.2} + \underset{(3.44)}{0.5}\, INC \qquad R^2 = 0.5$$

$$\textbf{Model B}: \widehat{\ln C} = \underset{(2.66)}{0.01} + \underset{(4.50)}{0.3}\, \ln INC \qquad R^2 = 0.6$$

$$\textbf{Model C}: \widehat{\ln C} = \underset{(3.02)}{0.01} + \underset{(1.58)}{0.1}\, t \qquad R^2 = 0.6$$

Note that the values in parentheses are t-ratios.

(a) **(4 points)** Is the elasticity of consumption with respect to income in **Model A** constant? If yes, what is it?

(b) **(4 points)** Is the elasticity of consumption with respect to income in **Model B** constant? If yes, what is it?

(c) **(2 points)** How would you interpret 0.1 (coefficient of time variable) in **Model C**?

(d) **(5 points)** Is the slope coefficient in **Model B** significant at 1% significance level? ***(Note: If you need, please refer to the probability table on the last page.)***

(e) **(5 points)** Suppose that $n = 82$ and $\sum \left(C_i - \bar{C}\right)^2 = 2,400$. Calculate the OLS estimator of σ^2 in **Model A**.

2. **(Total 20 points)** Consider the following simple linear regression model given by

$$Y_i = \beta + e_i$$

where the regression error e_i satisfies the classical assumptions.

(a) **(5 points)** Obtain the OLS estimator $\hat{\beta}$, and compare $\hat{\beta}$ and $\bar{Y}$ ($\bar{Y} = \frac{1}{n}\sum Y_i$).

(b) **(5 points)** Is the sum of residuals ($\hat{e}_i$) zero? That is, $\sum \hat{e}_i = 0$? Explain.

(c) **(5 points)** Is the OLS estimator $\hat{\beta}$ unbiased? Explain.

(d) **(5 points)** What is the variance of $\hat{\beta}$? Express it in terms of n and σ^2 (Note: Homeskedasticity assumption means $Var\left(e_i\right) = \sigma^2$ for all i).

- **Probability table for Final Exam**:
  - $P\left(\chi^2(1) \geq 3.841\right) = 0.05$, $P\left(\chi^2(1) \geq 5.023\right) = 0.025$, $P\left(\chi^2(1) \geq 7.879\right) = 0.01$ for χ^2 distribution with degrees of freedom 1.
  - $P\left(\chi^2(2) \geq 5.991\right) = 0.05$, $P\left(\chi^2(2) \geq 7.377\right) = 0.025$, $P\left(\chi^2(2) \geq 9.210\right) = 0.01$ for χ^2 distribution with degrees of freedom 2.
  - $P\left(\chi^2(3) \geq 7.814\right) = 0.05$, $P\left(\chi^2(3) \geq 9.348\right) = 0.025$, $P\left(\chi^2(3) \geq 11.34\right) = 0.01$ for χ^2 distribution with degrees of freedom 3.
  - $P(F(1, +200) \geq 3.84) = 0.05$, $P(F(1, +200) \geq 6.63) = 0.01$ for F distribution with degrees of freedom $(1, +200)$. (+200 indicates for more than 200 degrees of freedom.)
  - $P(F(2, +200) \geq 3.00) = 0.05$, $P(F(2, +200) \geq 4.61) = 0.01$ for F distribution with degrees of freedom $(2, +200)$. (+200 indicates for more than 200 degrees of freedom.)
  - $P(F(3, +200) \geq 2.60) = 0.05$, $P(F(3, +200) \geq 3.78) = 0.01$ for F distribution with degrees of freedom $(3, +200)$. (+200 indicates for more than 200 degrees of freedom.)
  - $P(Z \geq 1.28) = 0.1$, $P(Z \geq 1.645) = 0.05$, $P(Z \geq 1.96) = 0.025$, $P(Z \geq 2.58) = 0.005$ where Z is a standard normal random variable.

(End of Exam, Total 4 Pages)