STRATHMORE INSTITUTE OF MATHEMATICAL SCIENCES
END OF SEMESTER EXAMINATION FOR BACHELOR OF BUSINESS SCIENCE: ACTUARIAL SCIENCE, FINANCIAL ECONOMICS AND FINANCIAL ENGINEERING
BSE 2205: INTERMEDIATE ECONOMETRICS

2nd April, 2024 | Time: 2.5 hours

Instructions

1. This examination consists of **Five** questions.
2. Answer **Question One**(Compulsory) and **any other two** questions.

Question 1

(a) Consider the vector $X = [x_1\, x_2\, x_3 \ldots x_n]'$ and $Z = [1\ 1\ 1 \ldots 1]'$. What matrix expressions and operations would yield the following?

(i) $\sum_{i=1}^{n} x_i$ {1 mark}

(ii) $\sum_{i=1}^{n} x_i^2$ {1 mark}

(iii) $\begin{bmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{bmatrix}$ {2 marks}

(b) To estimate $\hat{\beta}_{OLS}$ we make several assumptions, one of these assumptions is that $E[UU'] = \sigma^2 I_n$. State the assumption(s) implied by this expression and with the help of appropriate algebra prove that indeed $E[UU'] = \sigma^2 I_n$ {5 marks}

(c) Eugene has collected some data on Stata. The resulting matrix is $X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$.
He seeks to estimate the model, $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i$. Therefore, column j corresponds to variable x_{j+1} Required:

(i) Construct the matrix X that Stata uses to estimate this model {3 marks}

(ii) Given X, what problems will Eugene experience in estimating the model {3 marks}

(iii) State the three specifications of the model that Eugene can estimate to overcome the challenges in 1c(ii) above {3 marks}

(d) Consider the model $y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \epsilon_i$. Required:

(i) State the OLS assumption that would be violated if x_{2i} was endogenous {2 marks}

(ii) What would be the consequences of the endogeneity of x_{2i}? {4 marks}

(iii) One of the methods of dealing with endogeneity requires that we understand how to derive $\hat{\beta}_{OLS}$ but with the generalized method of moments (GMM). Derive $\hat{\beta}_{GMM}$ for the equation in 1(c) above {4 marks}

(iv) Suggest the solution for endogeneity implied in 1c(iii) above {2 marks}

[30 marks]

Question 2

(a) Consider the model $Y = X\beta + U$ where $Y = y_1 y_2 y_3 \ldots \quad y_n{}'$,

1 x x … x_{K2}

1 x_{21} x_{31} … x_{KK1} 3 [1 2 [3 K] [ ]1 2 3 n]

$X =$ 1 x_{22} $_{22}$ $_{3n}$ x_{33} $_{23}$ $_{3n}$ … x , $\beta = \beta\ \beta \quad \beta \quad \ldots\beta\ {}'$ and $U \quad U \quad U \ldots U$

′

… … … … ⋱

1 x x … x_{Kn}

Required:

(i) If U is the residual, derive $\hat{\beta}_{OLS}$ using matrix algebra {3 marks}

(ii) Show that $\hat{\beta}_{OLS}$ is unbiased {3 marks}

(iii) Derive the expression for $var(\hat{\beta}_{OLS})$ {2 marks}

(iv) If $Y = [0 \quad 1 \quad 0 \quad 1 \quad 0]'$ and $x_2 = [0 \quad 1 \quad 1 \quad 0 \quad 0]'$ find $\hat{\beta}_{OLS}$ using the expression derived in 2a(i) above {3 marks}

(v) If $\sigma^2 = \frac{6}{25}$ find $var(\hat{\beta}_{OLS})$ using the expression in 2a(iii) above {3 marks}

(vi) What t-statistic is associated with the slope and intercept parameters given the estimates in 2a(iv) and 2a(v) above? {3 marks}

(b) Using a well labelled diagram distinguish between residuals, errors and bias {3 marks }

[20 marks]

Question 3

You are given the following data sampling process $y_i = \beta_1 + \beta_2 x_{2i} + \epsilon_i$ where:

$$\epsilon_i = \sqrt{x_{2i}} * u_i$$

$$u_i \overset{iid}{\sim} N(0,1)$$

and x_{2i} is a non-stochastic positive variable.

(a) Show that this model is heteroskedastic {2 marks}

(b) If the empirical information is $Y = [4 \quad 2 \quad 5 \quad 7]'$ and $x_{2i} = [1 \quad 1 \quad 4 \quad 4]'$. Estimate

$\hat{\beta}_{OLS}$ {2 marks}

(c) What are the characteristics of $\hat{\beta}_{OLS}$? {2 marks}

(d) Discuss how you would transform the data so that you could remove the heteroskedasticity {2 marks}

(e) Now estimate the model with the empirical information given in section (b), but by GLS. {4 marks}

(f) Show that in this case $var(\hat{\beta}_{GLS}) = \begin{bmatrix} \frac{10}{9} & -\frac{4}{9} \\ -\frac{4}{9} & \frac{5}{18} \end{bmatrix}$ {2 marks}

(g) Supply the robust standard errors that can be used to test the significance of β_1 and β_2 in 2(c) above {2 marks}

(h) The figure below shows edited stata output detailing results for a heteroscedasticity test. Use it to test whether the reference estimation was heteroscedastic. {2 marks}

```
. hettest

Breusch–Pagan/Cook–Weisberg test for heteroskedasticity
Assumption: Normal error terms
Variable: Fitted values of price

H0: Constant variance

    chi2(1) = 105.75
Prob > chi2 = 0.0000
```

(i) Benji ran the following regression $Price_i = \beta_1 + \beta_2 lotsize + u_i$. If this regression was heteroscedastic and price and lot-size are positively related, sketch the distribution of price around the line of best fit. {2 marks}

[20 marks]

Question 4

Consider the following data sampling process

$$Y_t = \beta x_t + \epsilon_t \text{ where } \epsilon_t = 0.6\epsilon_{t-1} + U_t$$

$$U_t \overset{iid}{\sim} N(0,1)$$

You are told that x is exogenous and are also given the following matrices:

$$X'X = \begin{bmatrix} 20 & 10 \\ 10 & 10 \end{bmatrix}', X'y = \begin{bmatrix} 86.6 \\ 68.4 \end{bmatrix}', X'\Psi X = \begin{bmatrix} 72.5 & 36.25 \\ 36.25 & 32.55 \end{bmatrix}', X'\Psi^{-1}X = \begin{bmatrix} 5.75 & 2.875 \\ 2.875 & 3.8125 \end{bmatrix}'$$

and $X'\Psi^{-1}y = \begin{bmatrix} 25.475 \\ 25.29375 \end{bmatrix}'$ where $\sigma^2\Psi$ is $E(\epsilon\epsilon')$

Required:

(a) Assumme that $\epsilon_t \sim N(\mu_\epsilon, \sigma_\epsilon^2)$ for every t. Show that $\mu_\epsilon = 0$ and var $(\epsilon_t) = \frac{25}{16}$ {3marks}

(b) What is the shape and dimension of Ψ? (You don't have to write it out in full) {4 marks}

(c) Estimate β_1 and β_2 using OLS {3 marks}

(d) Discuss the characteristics of $\hat{\beta}_{OLS}$ {5 marks}

(e) Estimate the true value of variance-covariance matrix of $\hat{\beta}_{OLS}$ {3 mark}

(f) Test the null hypothesis that $\beta_2 = 0$ using your OLS estimator of β_2 {2 marks}

[20 marks]

Question 5

You are given the following description of a data sampling process $y_i = \beta_1 + \beta_2 x_i + \epsilon_i$. $iid \sim$ ^{2}).

$\sum$ You have a sample size of 1000 and the following information: Each ϵ_i is ϵ_i $N(0,\sigma$ x_i = 2,000, $\sum x^2_i = 9{,}000$, $\sum x_i y_i = 8{,}000$ and $\sum y_i = 5{,}000$. Required:

(a) Find $E[y_i|x_i]$ {2 marks}

(b) Find var$[y_i|x_i]$ {3 marks}

(c) How does $X'X$ and $X'Y$ look like in this case {4 marks}

(d) Find $\hat{\beta}_{ols}$ given your findings in 5(c) above {4 marks}

(e) What is the true value of $var(\hat{\beta}_{OLS})$ given the above findings? {4 marks}

(f) If $\epsilon_i'\epsilon_i = 300,000$, test the statistical significance of each regression coefficient given that $t_{critical} = 1.96$ {3 marks}

[20 marks]

Be More be Strathmore! ©