**FINM3123 Introduction to Econometrics**

**Quiz 2**

**Name: ­\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Student ID: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Date: 17 December 2024**

**Time allowed: 60 minutes Full mark: 60**

1. **Multiple choice questions (12 points):**

*Identify the letter of the choice that* ***best*** *completes the statement or answers the question. 2 points for each question.*

1. If a categorical independent variable contains 3 categories, then \_\_\_\_\_\_\_\_\_ dummy variable(s) will be needed to uniquely represent these categories in a regression model containing an intercept.
2. 1
3. 2
4. 3
5. 4
6. Consider the following regression equation: y = β0+β1x1+…βkxk+ u. In which of the following cases, is ‘y’ a discrete variable?
7. y indicates the gross domestic product of a country
8. y indicates the total volume of rainfall during a year
9. y indicates household consumption expenditure
10. y indicates the number of children in a family
11. What will you conclude about a regression model if the Breusch-Pagan test results in a small p-value?

a. The model contains homoskedasticty.

b. The model contains heteroskedasticty.

c. The model contains dummy variables.

d. The model omits some important explanatory factors.

1. Which of the following is true?

a. In ordinary least squares estimation, each observation is given a different weight.

b. In weighted least squares estimation, each observation is given an identical weight.

c. In weighted least squares estimation, less weight is given to observations with a higher error variance.

d. In ordinary least squares estimation, less weight is given to observations with a lower error variance.

1. Which of the following is true of Regression Specification Error Test (RESET)?

a. It tests if the functional form of a regression model is misspecified.

b. It detects the presence of dummy variables in a regression model.

c. It helps in the detection of heteroskedasticity when the functional form of the model is correctly specified.

d. It helps in the detection of multicollinearity among the independent variables in a regression model.

1. The classical errors-in-variables (CEV) assumption is that \_\_\_\_\_.

a. the error term in a regression model is correlated with all observed explanatory variables

b. the error term in a regression model is uncorrelated with all observed explanatory variables

c. the measurement error is correlated with the unobserved explanatory variable

d. the measurement error is uncorrelated with the unobserved explanatory variable

**II. Problems (48 points)**

1. (12 points) We estimate a model that allows for wage differences among four groups: married men, married women, single men, and single women, denoted by *marrmale*, *marrfem*, *singmale*, and *singfem* respectively. Column (1) of the following table shows the result of regressing log(*wage*) on *educ*, *exper*, and group dummy variables, with single men as the base group. Column (2) shows the result of the same regression, using married women as the omitted category. Numbers in parentheses are standard errors. As far as this is possible, determine the numbers represented by all the letters (A to L) in Column (2). If there are values that cannot be determined, please point them out.

|  |  |  |
| --- | --- | --- |
|  | (1) | (2) |
| constant | 0.321  (0.100) | A  (B) |
| *marrmale* | 0.213  (0.055) | C  (D) |
| *marrfem* | –0.198  (0.058) | —— |
| *singfem* | –0.110  (0.056) | E  (F) |
| *singmale* | —— | G  (H) |
| *educ* | 0.079  (0.007) | I  (J) |
| *exper* | 0.027  (0.005) | K  (L) |
| *n* | 500 | 500 |

1. (22 points) Consider the linear probability model , where .
2. (6 points) What is the value of ? Show your work.
3. (6 points) Show that .
4. (4 points) Is heteroskedastic? Explain.
5. (6 points) Illustrate the procedure for estimating the model parameters and .
6. (14 points) Consider a linear model (Regression 1) for log(wage) (lwage) as a function of education (educ), experience (exper), and tenure (tenure) based on a dataset of employees. The output is shown below:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Let be the log(wage) prediction from this regression, and let be the residual from this regression. The following regression (Regression 2) is performed:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

Using the outcomes of Regression 2, perform an alternative White test of Regression 1 at the 5% significance level through the *F* statistic. What can we conclude about Regression 1?



