Indian Institute of Technology Kharagpur
Department of Humanities and Social Sciences

Class Test - I Full Marks: 20 Time: 1hr. Date: 21/02/2022 (Monday)
Subject No: HS20202/HS41002 Subject Name: Econometric Analysis I

Instructions: **Answer all the questions. Write the answers in sequence of the question paper. Submit the Answer Script within the given timeline in MS Teams only. Submission through email or late submission will not be considered.**

Group A - Select the most appropriate alternative for the following: 1×14 = 14

1. When the explanatory variable of a regression model is "non-stochastic", it is:
   a) Partly random
   b) Fixed in repeated samples
   c) Correlated with the errors
   d) Always confined to a value of unity

2. If a hypothesis test is conducted using the 5 percent significance level:
   i) It is equal to the size of the test
   ii) It is equal to the power of the test
   iii) 2.5 percent of the total distribution will be in the rejection region of each tail for a two-tailed test
   iv) 5 percent of the total distribution will be in the rejection region of each tail for a two-tailed test

   a) (ii) and (iv) only
   b) (i) and (iii) only
   c) (i), (ii) and (iii) only
   d) (i), (ii), (iii) and (iv)

3. The slope of a bivariate linear regression model cannot be estimated if:
   a) The independent variable takes the same value across all the observations
   b) The variance of independent variable is maximized
   c) The independent variable takes values as reciprocal of the dependent variable
   d) None of the above

4. The method of ordinary least squares helps in:
   a) Minimising the distance across the data points
   b) Finding the values of the regression coefficients
   c) Finding the values of the regression coefficients with the smallest values of the residual sum of squares
   d) Finding the values of regression coefficients with the smallest sum of the residuals

5. In regression analysis, the scatter plot around the regression line is measured by the:
   a) Explained sum of squares
   b) Unexplained sum of squares
   c) Mean value of the dependent variable
   d) None of the above

6. If the computed value of the F statistic is greater its critical value, it implies that:
   a) None of the independent variables has a significant effect
   b) One or more independent variables have a significant effect
   c) All the independent variables have a significant effect
   d) None of the above

7. Which of the following is true for R^2?
   a) It is also known as the standard error of the regression
   b) A low value R^2 indicates that the estimated model should be discarded
   c) The value of R^2 usually decreases with inclusion of additional independent variables
   d) It shows the percentage of total variations in the dependent variable that is explained by the explanatory variable(s).

8. If a multiple regression model has a very high R^2, but none of the slope coefficients is significantly different from zero, it indicates that:
   a) The explanatory variables are highly collinear
   b) The dependent variable does not vary by much
   c) The intercept has been omitted from the model
   d) The explanatory variables are highly orthogonal

9. The standard error of an estimated regression model measures the:
   a) Variability of the dependent variable relative to the estimated regression line
   b) Variability of the independent variable(s) relative to its(their) mean
   c) Variability of the dependent variable relative to its mean
   d) The average error that will result if the regression line is used for prediction

10. The adjusted value of the coefficient of determination:
    a) Will always increase with inclusion of additional independent variables to the model
    b) Is equal to the proportion of variations of the dependent variable explained by the regression model
    c) Is always greater than the proportion of variations of the dependent variable explained by the regression model
    d) Is always less than the proportion of variations of the dependent variable explained by the regression model

11. Which of the following would be the consequences of violating one or more assumptions of the classical linear regression model (CLRM)?
    i) The estimated coefficients are not optimal
    ii) The estimated standard errors are not optimal
    iii) The statistical distributions assumed for the test statistics are inappropriate
    iv) Conclusions on the relationships between the dependent and the independent variables may be invalid

    a) ii and iv only
    b) i and iii only
    c) i, ii, and iii
    d) i, ii, iii, and iv

12. Which of the following statements is true about the regression line?
    a) A regression line indicates the average relationship
    b) A regression line is also known as the estimating equation
    c) A regression line can also be termed as the prediction equation
    d) All of the above

13. In a linear regression model, when an additional independent variable is included:
    a) Its coefficient will be significant if the value of the R^2 increases
    b) Its coefficient will not be significant if the value of the R^2 decreases
    c) The importance of a variable cannot be judged only on the basis of R^2 value
    d) None of the above

14. Consider the following two statements:
    i) If the null hypothesis is rejected in one-tailed test, it will be rejected in two-tailed test as well at the same significance level
    ii) For the chosen significance level, the critical value of the t statistic gets closer to zero as the sample size increases

    a) Both statements are true
    b) Neither of the statements is true
    c) Only the first statement is true
    d) Only the second statement is true

Group B - Answer the following questions: $2 \times 3 = 6$

1. Assume that initially the bivariate model $Y_i = \alpha + \beta X_i + u_i$ is estimated. However, when an additional independent variable Z_i is added to this model and the new model $Y_i = \delta + \lambda X_i + \theta Z_i + v_i$ is estimated, the sign of the coefficient of X_i changes. What may be the possible reason(s) for such a finding? Justify your answer with a suitable example.

2. While computing the variance inflation factors (VIFs) of the independent variables in a multiple regression model, should one use adjusted R^2 instead of the ordinary R^2 of the auxiliary regressions? Justify your answer.

3. Suppose while testing the hypothesis $H_0 : \lambda = \theta = 0$ for the model $Y_i = \delta + \lambda X_i + \theta Z_i + v_i$, it is found that the null hypothesis is rejected at 5 percent significance level. However, it is not so when 1 percent significance level is considered. What should be concluded in such a case? Give justification in support of your answer.