2024

ECONOMICS — HONOURS

Paper : CC-10

(Introductory Econometrics)

Full Marks : 65

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Group - A

2×10

1. Answer ***any ten*** questions :

(a) State whether the following statements are true or false :

(i) The population parameters are variables while the sample estimators are constant.

(ii) Autocorrelation is a problem related primarily to time-series data.

(iii) The t test of significance requires that $\hat{\beta}_1$ and $\hat{\beta}_2$ follow normal distribution.

(iv) If the disturbance term in the CLRM is not normally distributed, the OLS estimators are not unbiased.

(b) Let the population model be given by :

$gpa = \beta_0 + \beta_1 \text{ study} + U_i$; where gpa = the university grade point average and study = hours spent per week studying.

(i) What is the expected sign of β_1?

(ii) What is the interpretation of β_0?

(c) Which of the following can cause the usual OLS t statistics to be invalid?

(i) Heteroskedasticity.

(ii) A simple correlation coefficient of 0.95 between two independent variables that are in the model.

(iii) Omitting an important explanatory variable.

(d) When is an estimator said to be biased?

(e) The regression equation of Y on X estimated on the basis of 10 observations is given by :

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i ,$$

where $\sum Y_i = 1110$; $\sum X_i = 1700$; $\sum X_i Y_i = 205500$; $\sum X_i^2 = 322000$

Determine $\hat{\beta}$.

Please Turn Over

(f) It is given that

Source of variation	Sum of Squares	Degrees of Freedom
RSS	337.27	08
TSS	8890.00	09

Determine the $\bar{R}^2$.

(g) Let *kids* denote the number of children ever born to a woman and *educ* denote years of education for the woman. A simple model relating number of children ever born to a woman to years of education is $kids = \beta_0 + \beta_1\ educ + U_i$. Mention any two factors contained in U_i.

(h) Suppose you want to examine if there is a correlation between amount of food eaten and blood pressure, while controlling for body weight. If correlation coefficient between blood pressure and food eaten (r_{12}) is 0.87, correlation coefficient between blood pressure and body weight (r_{13}) is 0.66 and correlation coefficient between food eaten and body weight (r_{23}) is 0.57, then find the correlation between blood pressure and food eaten eliminating the effect of body weight.

(i) In a two variable PRF, $E(Y|X_i) = \beta_1 + \beta_2 X_i$. If the slope coefficient is zero, then what does that mean?

(j) How can you measure the reliability of the least squares estimators?

(k) What additional assumptions are required in the multiple regression model compared to the two variable regression model?

(l) Suppose the true Cobb-Douglas equation $\ln Y_i = \alpha_0 + \alpha_1 \ln L_{1i} + \alpha_2 \ln L_{2i} + \alpha_3 \ln K_i + U_i$ is wrongly written as : $\ln Y_i = \beta_0 + \beta_1 \ln L_{1i} + \beta_2 \ln K_i + U_i$.

(i) Will the condition $E(\hat{\beta}_1) = \alpha_1$ and $E(\hat{\beta}_2) = \alpha_3$ be satisfied?

(ii) Will the answer in (i) hold if it is known that L_2 is an irrelevant input in the production function?

(m) If $r_{12} = 0.97$; $r_{13} = 0.99$ and $r_{23} = 0.97$, find $R_{1.23}$.

(n) Is the model $\ln Y_i = \beta_1 + \beta_2\left(\frac{1}{X_i}\right) + u_i$ a linear regression model? Why?

(o) If $\hat{\beta}_{YX}$ and $\hat{\beta}_{XY}$ represent the estimated coefficients in the regression of Y on X and X on Y, respectively, show that

$$\hat{\beta}_{YX}\,\hat{\beta}_{XY} = r^2,$$

where r is the correlation coefficient between X and Y.

Group - B

Answer ***any three*** questions.

2. Suppose that you estimate the consumption function $Y_i = \alpha_1 + \alpha_2 X_i + U_{1i}$ and the savings function $Z_i = \beta_0 + \beta_1 X_i + U_{2i}$

(a) What is the relation between α_2 and β_2 (if any)?

(b) Can you compare the R^2 terms of the two models? Why or why not? 3+2

3. Given a sample of 50 observations and 4 explanatory variables, for each of the following cases what can you say about autocorrelation if :

   (a) $d = 1.05$

   (b) $d = 1.40$

   (c) $d = 2.50$

   (d) $d = 3.97$

   (e) $d = 2.00$

   Where d is the Durbin-Watson d statistics

   [Corresponding to 5% level of significance, $n = 50$ and $k = 5$, $d_L = 1.335$ and $d_U = 1.771$].

   1+1+1+1+1

4. What do you mean by multicollinearity? What are its sources? Why can not the regression coefficients be measured accurately when multicollinearity is less than perfect? 1+2+2

5. To estimate the relationship between sales and wages of the employees of a firm three models are estimated as shown below :

$$\text{wage} = \beta_1 + \beta_2 \text{ sales} + u$$
$$\log(\text{wage}) = \beta_3 + \beta_4 \text{ sales} + u$$
$$\log(\text{wage}) = \beta_5 + \beta_6 \log(\text{sales}) + u$$

   (a) Explain how the interpretation of the values of the slope coefficients will vary in these models.

   (b) Does the interpretation change if wage is measured in thousand rupees and sales in lakh rupees in model 1? 4+1

6. From a sample of 3000 employees the following relationship is estimated

$$\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i ,$$

   where Y_i is the wage (in thousand rupees) and X_i is the age (in years). The following informations are shown below :

$$\sum x_i y_i = 3050.24 \quad \sum x_i^2 = 22356.23 \quad \sum y_i^2 = 72320.22$$
$$\sum X_i = 82380.23 \quad \sum Y_i = 29540.25 \quad RSS = 75000.26$$

   where $x_i = X_i - \bar{X}$ and $y_i = Y_i - \bar{Y}$

   (a) Use the above information to compute $\hat{\alpha}$ and $\hat{\beta}$.

   (b) Interpret the slope coefficient.

   (c) Estimate the standard error of the estimated regression coefficient. 2+1+2

Group - C

Answer *any three* questions.

7. (a) What is the difference between population regression function and sample regression function?

(b) Consider the following estimated regression equation :

$\widehat{sleep}$ =3638.25 – 0.148 totwrk – 11.13 educ + 2.20 age

SE = (112.28) (0.017) (5.88) (1.45)

$R^2 = 0.113$; $n = 706$

where 'sleep' = total number of minutes slept per week;

'totwrk' = total number of minutes worked per week;

'educ' = the number of years of education

'age' = age in years

(i) Why is the sign of coefficient of work negative? Interpret the coefficient. Is it significant?

(ii) Discuss the sign and magnitude of the estimated coefficient on education.

(iii) Test the null hypothesis that all the coefficients are simultaneously zero against the alternative hypothesis that they are not simultaneously zero.

(iv) Is the coefficient of age significant?

[Given $F_{0.01,(3,702)} = 3.78$ and $t_{0.01,702} = 2.576$]

2+(3+1+3+1)

8. (a) From the data for 45 developed countries, the following regression results were obtained :

$$\widehat{Log\ C} = 4.30 - 1.34 \text{ Log } P + 0.17 \text{ Log } Y$$

SE = (0.09) (0.32) (0.20)

$\bar{R}^2 = 0.27$

where C = cigarette consumption (packs per year)

P = real price of cigarette per pack

Y = per capita real income

(i) Discuss the sign and magnitude of coefficient on Log Y. What does this coefficient stand for?

(ii) What is the elasticity of demand for cigarette with respect to price? Is it significant?

(iii) How would you determine R^2 from the given $\bar{R}^2$? [Given $t_{0.01,42} = 2.704$]

(b) Define coefficient of determination. What role does it play in econometric analysis?

(3+2+2)+(1+2)

9. (a) Write a short note on importance of dummy variable in econometric analysis.

(b) Consider the following model :

$$\hat{Y} = 46.67 + 26.25\ D_1 - 6.07\ \log(X)$$
$$SE = (43.41)\ (5.59) \qquad (3.88)$$
$$R^2 = 0.237;\ n = 105$$

where Y = number of hours of training per employee at firm level

X = number of employees

D_1 is a dummy variable such that

$D_1 = 1$ if the firm receives a job training grant,

$= 0$ otherwise

(i) Interpret the regression coefficient of the dummy variable in the above model. Is it statistically significant?

(ii) How does the size of employment affect the number of hours of training per worker provided by the firm? [Given $t_{0.01,102} = 2.638$]

5+(3+2)

10. (a) Suppose you have increased the number of explanatory variables in a multiple linear regression model. What will be its impact on R^2 and adjusted R^2?

(b) In a regression model $Y = \alpha + \beta X + U$. Suppose that we multiply each X value by a constant 2. Examine whether it will change the residuals and the fitted values of Y. 4+6

11. (a) A regression of average daily earnings (E) measured in rupees on age (measured in years) using a random sample of workers yields

$$\hat{E} = 696.7 + 9.6\ age$$
$$R^2 = 0.023,\quad RSS = 1542.2$$

(i) What are the units of RSS and R^2?

(ii) What is the predicted earning of a 30 years old worker?

(iii) What is the average increase in earning for one year increase in age?

(b) Explain how R^2 and adjusted R^2 measure the goodness of fit of the regression. Do you think that in the regression analysis our objective will always be to obtain a high R^2 or adjusted R^2?

(2+1+1)+(4+2)