

**Course Name: BBE**

**Paper Name: BASIC ECONOMETRICS**

**Semester: IV**

ASSIGNMENT QUESTIONS

Q1 Comment on any five of the following. Give reasons in support of your comment.

5 x 3

- a) The value of adjusted  $R^2$  is always less than  $R^2$ .
- b)  $R^2$  and adjusted  $R^2$  are always positive.
- c) Testing the significance of the slope coefficient in a 2-variable linear regression model is the same as testing the overall significance of the model.
- d)  $Y_i = \beta_1 + \beta_2^3 x_i + \mu_i$  and  $Y_i = \beta_1 + \beta_2 (1/x_i) + \mu_i$  are linear regression models and therefore meet all the assumptions of a CLRM (Investigate separately).
- e) The stochastic error term is irrelevant in the regression analysis as its mean value is always zero.
- f) If you have monthly data for a number of years, you need to use 12 dummies to study seasonal effects if all months are known to exhibit seasonal variations.

Q2. a) To reduce crime, the minister has budgeted more money to put more police force in the city. A regression to study the effect of police deployment ( $Y$  in '000) on the number of reported crimes ( $X$ ) was studied. From the data pertaining to 8 weeks, following results were obtained.

$$\sum X_i = 76 \quad \sum Y_i = 130$$

$$\sum X_i^2 = 116 \quad \sum Y_i^2 = 241.5$$

$$\sum X_i Y_i = 155 \quad \text{RSS} = 34.38$$

Where  $x_i, y_i$  are deviations from their respective mean

- i) Estimate both the regression coefficients and interpret the regression equation. 4
- ii) Test the significance of the slope coefficient at 5% level of significance 3
- iii) Calculate the value of coefficient of determination and interpret it. 3

b) You are given the following regression result

$$\text{Sales } (Y_t) = 4.3863 + 1.08132 \text{ ADV}_t$$

$$t = 4.42 \quad 13.99$$

$$r^2 = 0.938$$

To reduce crime, the minister has budgeted more money to put more police force in the city. A regression to study the effect of police deployment (Y in '000) on the number of reported crimes (X) was studied. From the data pertaining to 8 weeks, following results were obtained.

$$\sum X_i = 76 \quad \sum Y_i = 130$$

$$\sum X_i^2 = 116 \quad \sum Y_i^2 = 241.5$$

$$\sum X_i Y_i = 155 \quad \text{RSS} = 34.38$$

Where  $x_i, y_i$  are deviations from their respective mean

- iv) Estimate both the regression coefficients and interpret the regression equation.
- v) Test the significance of the slope coefficient at 5% level of significance
- vi) Calculate the value of coefficient of determination and interpret it.

b) You are given the following regression result

$$\text{Sales } (Y_t) = 4.3863 + 1.08132 \text{ ADV}_t$$

$$t = 4.42 \quad 13.99$$

$$r^2 = 0.938$$

Find the sample size underlying the result.

- Q43 a) The Price (P) of housing (in '000 of rupees) is regressed on the area (A) of the house (in square feet), number of bedrooms (B), number of baths (BA), presence of pool (DP), family room (DFAM) and fire place (DFR). Pool, family room and fire place are dummy variables that take the following values:

DP = 1 if pool is present, 0 otherwise

DFAM = 1 if family room is present, 0 otherwise

DFR = 1 if fire place is present, 0 otherwise

The following regression results were obtained.

$$P = 39.057 + 0.147 A - 7.046 B - 0.264 BA + 53.196 DP - 21.345 DFAM + 26.188 DFR$$

$$(0.436) \quad (4.869) \quad (-0.245) \quad (-0.006) \quad (2.411) \quad (-0.498) \quad (0.486)$$

$$n = 14 \quad \text{adjusted } R^2 = 0.836$$

Figures in brackets are t values.

- i) Interpret the model.
- ii) Write the equation for price as a function of all quantitative variables and dummy variables when pool, family room and fire place are present.
- iii) How would the equation change if only pool was present, family room and fire place were absent.
- iv) Do you think that presence of pool significantly contributes to the price of house.
- v) Is the model suggested above overall significant (test using the F statistic).

Q4 b) What are the advantages of the dummy variable technique over the Chow test while trying to measure structural stability?

Q4 c) Following is the demand schedule for commodity x.

$$D_x = f(P_x, P_y, Y)$$

Where the  $D_x$  is the demand for commodity x,  $P_x$  is its price,  $P_y$  is the price of related commodity y and Y is the income of the consumer.

How do you measure the elasticity of demand with respect to own price and price of related commodity y if you use a i) double log model ii) linear model.

Q5 a) Following is a model relating to number of people who travel by bus (BUS) and various factors that affect it.

$$BUS = f(FA, IN, POP, DEN, LANDA)$$

Where BUS is a measure of urban transportation by bus in thousands of passenger hours.

FA = Bus fare in rupees.

IN = Average income per capita in rupees.

POP = Population of city in '000.

DEN = Density of population (persons/sq.km)

	Model A	Model B	Model C
Constant	3215.86	3111.18	2815.7
Fare (FA)	-225.66 (440.49)	-295.73 (424.83)	-----
Income (IN)	-0.195 (0.638)	-0.2022 (0.062)	-0.2013 (0.0621)
Population (POP)	1.7168 (0.2265)	1.588 (0.1227)	1.576 (0.1206)
Density (DEN)	0.1182 (0.0580)	0.1490 (0.0357)	0.1534 (0.0349)
Land Area (LANDA)	-1.1953 (1.765)	-----	-----
R <sup>2</sup>	0.909	0.911	0.912

Figures given are coefficients of independent variables and figures in brackets are standard errors.

- i) Interpret coefficients of model A.
- ii) Find out whether the variable Fare (FA) should be added to the model or not.
- iii) Also find out whether land area (LANDA) should be added to the model.
- iv) Test the significance of the population variable in model A and Density in model C.

Q5 b) Given the following model

$$CM_i = 81.79436 + 27273.17 ( 1/PCGNP)$$

$$S.E. = (10.8321) \quad (3759.99)$$

$$r^2 = 0.4590$$

CM = child mortality rate i.e., number of deaths per '000 live births.

PCGNP = per capita GNP in rupees.

- i) What kind of a model is it?
- ii) Interpret the model.
- iii) At what level of PCGNP will a country achieve a child mortality value of zero?

Q6. Write short notes on any three of the following.

- i) JB test.
- ii) SRF versus PRF
- iii) Influence of level of significance on critical Z values.
- iv) Log-Lin versus Lin-Log regression models.

#### CLASS TESTS

Q1 A car manufacturer is interested in trying to estimate how sales are influenced. To do this, the company randomly chooses 10 small cities and offers cars at different prices. Using car sales (SALES) as dependent variable and price (PRICE) as explanatory variable, the company conducts a simple linear regression.

The following least squares results are:

$$\text{SALES} = 161.39 - 48.19\text{PRICE} + e$$

$$Se = (26.16) \quad (12.65)$$

Error sum of squares = 1062.58

Total sum of squares = 4918

Both sales and prices are measured in Rs'000.

- a) Interpret the model.
- b) Use 5% significance level to test the null hypothesis that price has no effect on sales. State the alternative hypothesis.
- c) Calculate and interpret coefficient of determination.
- d) Comment on the overall significance of the model.

#### OR

A regression was run between nominal exchange rate and relative prices from the year 1980 to 1994 (in rupees) and the following results were obtained. Here  $Y_i$  is the exchange rate of US to India and  $X_i$  is the ratio of India CPI to US CPI.

$$Y_i = 6.682 - 4.318 X_i$$

$$Se = (1.22) \quad (1.333) \quad r^2 = 0.528$$

- a) What is the economic interpretation of regression coefficients.

- b) Does the negative value of  $X_i$  make economic sense? What is the underlying economic theory?
- c) What do you think about goodness of fit. Test the slope coefficient at 1% significance level.
- d) Construct 95% confidence interval.

**Q3 a)** What is Sample regression function? Differentiate with population regression function.

b) Explain the significance of the stochastic error term in the regression model.

c) How do we estimate the value of slope coefficient in the linear

regression model.

**OR**

a) State the assumptions of CLRM

b) Prove that the least square estimators are BLUE.

**Q5 a)** Write short note on any **one** of the following:

i) Gauss Markov Theorem

ii) Central limit theorem.

b) Given the data on Consumer Price Index(CPI),  $Y$  and Wholesale price index (WPI),  $X$  for the following

$$\sum Y = 952.1, \sum X = 988.7, \sum Y^2 = 48710.17, \sum X^2 = 53818.43, \sum XY = 51145.55$$

i) Estimate  $Y_i = \beta_1 + \beta_2 X_i$

ii) Construct the ANOVA table

iii) Outline the procedure test for normality of the disturbance term for:

$$\text{Skewness} = 0.38, \text{Kurtosis} = 3.5$$

What would you conclude about the behavior of the disturbance term at 5% level of significance( State null and alternative hypothesis )

## MCQ'S

1. The strength (degree) of the correlation between a set of independent variables  $X$  and a dependent variable  $Y$  is measured by
  - a. Coefficient of Correlation
  - b. Coefficient of Determination
  - c. Standard error of estimate
  - d. All of the above
2. The percent of total variation of the dependent variable  $Y$  explained by the set of independent variables  $X$  is measured by
  - a. Coefficient of correlation
  - b. Coefficient of Skewness
  - c. Coefficient of Determination
  - d. Standard error of estimate
3. A coefficient of correlation is computed to be  $-0.95$  means that
  - a. The relationship between the two variables is weak
  - b. The relationship between the two variables is strong and positive
  - c. The relationship between two variables is strong but negative
4. The locus of the conditional means of  $Y$  for fixed values of  $X$  is
  - a. Conditional expectation function
  - b. Intercept line
  - c. Population regression line
  - d. Linear regression line
5. When we reject the null hypothesis, then our finding is said to be
  - a) 95% probability finding
  - b) 5% confidence finding
  - c) Not statistically significant
6. The sample parameter estimator  $\beta_2$  follows
  - a. T- distribution

- b. Normal distribution
- c. F distribution
- d. Chi square distribution

7. Dummy variable can take the value only 0 or 1

- a. True
- b. False

8. If there exists high multicollinearity, then the regression coefficients are

- a. Determinate
- b. Indeterminate
- c. Infinite value

9. Heteroscedasticity may result due to the presence of

- a. Outliers in the sample
- b. Omission of explanatory variable
- c. Both a and b

10. When error terms across cross section data are correlated, it is known as

- a. Cross correlation
- b. Cross autocorrelation
- c. Serial autocorrelation

11. ANOVA models include

- a. Only quantitative variables
- b. Only qualitative variables
- c. Both a and b

12. Accepting a false hypothesis results in

- a. Type I error
- b. Type II error
- c. Structural error

13. For regression through origin, the intercept is equal to



- a. 1
- b. 2
- c. 0
- d. -1

14. In a log regression model, the regression slope gives

- a. The relative change in Y for an absolute change in X
- b. The percentage change in Y for a given percentage change in X
- c. The absolute change in Y for a percent change in X.

15. The value of adjusted R square is always less than R square

- a. True
- b. False

16. To test for structural breaks in time series data, we use

- a. T test
- b. F test
- c. Chow test

17. If a quantitative variable has m categories we can introduce

- a. Only m-1 dummy variables
- b. Only m dummy variables
- c. Only m+1 dummy variables

18. Multicollinearity is same as multicollinearity

- a. True
- b. False

19. Durbin Watson is used to detect

- a. Autocorrelation
- b. Multicollinearity
- c. Heteroscedasticity

20. Elasticity of a regression model can be measured through

- a. Log log model
- b. Lin log model
- c. Log lin model

**Answers : b,a,a,c,d,b,b,a,c,a,b,c,c,a,c,a,b,a,a**