

Instructions

Answer five questions including question number one (1).

No. of questions : Five (06)
No. of pages : Four (04)
Total marks allocated : 40%
Time : Two Hours (02 hrs)

1. Propose the measurements of variables and suitable statistical procedures which can be used in following studies.

- I. To test whether there is a significant difference between dry weights of plants in two nurseries.
- II. To test whether there is a significant reduction of water related conflicts in irrigation scheme after introducing the management of water through user organizations (Farmer Organizations)
- III. To test whether there is a significant reduction of body weight of 20 obese persons after attending a three month aerobic exercise programme.
- IV. To test the impact of three growth regulators for the increasing of body weight of broilers.
- V. To test whether there is a significant difference of preferences of consumers on three different packages (A, B, C) of canned fruits.
- VI. To test whether there is an association between colours and gender of the consumers when selecting plastic toys.
- VII. To test whether there is an association between age of the consumers and daily intake of water.
- VIII. To test whether there is an association between the girth (circumference) of the rubber trees and latex yield of rubber trees.

2. (a) Derive the Karl Pearson Correlation Coefficient (r) from sum of the products of two variables X and Y.

(b) Discuss the weaknesses of (1) Sum of the products, (2) Covariance and (3) Pearson Correlation Coefficient as an indicator of relationship between two variables

(c) Calculate Pearson Correlation Coefficient (r) for the following dataset and interpret results.

Plant number	1	2	3	4	5	6	7	8	9	10
Total root length of been plants (cm)	127	94	87	63	78	113	98	89	64	102
Weight of root nodules of been plants (g)	52	43	37	26	41	47	39	35	31	44

3. (a) What are the steps of the methodology of econometrics

(b) Briefly discuss the use of dummy variables in econometric models

(c) To estimate a regression model to forecast paddy production in yala and maha seasons the time series data of paddy production of Maha (1) and Yala (0) seasons from 1990 to 2010 were used. Different equations estimated were;

$$Y_t = \alpha + \beta_1 T$$

$$Y_t = \alpha + \beta_1 T + \beta_2 D$$

$$Y_t = \alpha + \beta_1 T + \beta_2 DT$$

$$Y_t = \alpha + \beta_1 T + \beta_2 DT + \beta_3 D$$

Develop the equations (a) Common gradient and different intercepts (b) Different gradient and common intercepts and (c) Different gradients and different intercepts, to estimate the production in Yala and Maha seasons.

4. (a) Briefly explain the properties of the error term of the OLS estimate.

(b) Following table gives the quantity produced (Q) and total cost of production in a factory for several years. Estimate the linear cost function of the factory.

Production (Tons)	40	42	48	55	65	79	88	100	120	140
Cost of production (1000 rupees)	150	140	160	170	150	162	185	165	190	185

5. (a) Briefly explain the different components of a time series model.

(b) What are the different methods use to estimate the secular trend of a time series data set.

(c) Following table shows the monthly prices of coconut for last five years. Estimate the monthly index of coconut prices for months of January to December. Comment on the results.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
2006	32	31	28	26	23	20	22	24	27	27	29	31
2007	33	31	28	25	22	22	23	25	28	28	30	31
2008	36	35	30	28	25	25	25	27	30	32	34	35
2009	40	36	35	31	30	28	30	32	33	36	36	38
2010	50	50	45	40	40	36	34	40	42	42	45	48

6. (a) What do you mean my multicollinearity of multiple regression models?

(b) What are the indicators of the presence of multicollinearity of a multiple regression model?

(c) To estimate the potato production of Sri Lanka, production data of last 20 years were used and following models were estimated. Variables: Production in metric tons (Y), Year (T), Season (D: Maha = 1, Yala = 0)

UNWEIGHTED LEAST SQUARES LINEAR REGRESSION OF Y

VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
CONSTANT	132.295	5.02222	26.34	0.0000	
D	44.1818	4.73499	9.33	0.0000	1.0
T	7.59545	0.74867	10.15	0.0000	1.0

R-SQUARED 0.9091 RESID. MEAN SQUARE (MSE) 123.311
 ADJUSTED R-SQUARED 0.8995 STANDARD DEVIATION 11.1045

PREDICTOR	VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
	CONSTANT	117.409	2.63501	44.56	0.0000	
	D	73.9545	3.72647	19.85	0.0000	3.5
	T	10.5727	0.44540	23.74	0.0000	2.0
	DT	-5.95455	0.62989	-9.45	0.0000	4.5

R-SQUARED 0.9848 RESID. MEAN SQUARE (MSE) 21.8217
 ADJUSTED R-SQUARED 0.9822 STANDARD DEVIATION 4.67137

PREDICTOR VARIABLES	COEFFICIENT	STD ERROR	STUDENT'S T	P	VIF
CONSTANT	154.386	8.67485	17.80	0.0000	
T	5.29026	1.66265	3.18	0.0049	1.3
DT	4.61039	1.56756	2.94	0.0084	1.3

R-SQUARED 0.6513 RESID. MEAN SQUARE (MSE) 473.019
 ADJUSTED R-SQUARED 0.6146 STANDARD DEVIATION 21.7490

Select the best model giving reasons to estimate the potato production for future years.
 Estimate the potato yield for the next year (n=21) using the selected equation.