

No. of Printed Pages : 5

MST-017

M. SC. (APPLIED STATISTICS)

(MSCAST)

Term-End Examination

June, 2026

MST-017 : APPLIED REGRESSION ANALYSIS

Time : 3 Hours

Maximum Marks : 50

Note : (i) *Question No. 1 is compulsory.*

(ii) *Attempt any **four** questions from the remaining question nos. 2 to 6.*

(iii) *Use of scientific calculator (non-programmable) is allowed.*

(iv) *Symbols have their usual meanings.*

1. State whether the following statements are True or False. Give reasons in support of your answers : 5×2=10

(a) The model :

$$Y = \beta_0 + \sqrt{\beta_1}X + \beta_2\exp(X) + \varepsilon$$

is a linear model.

- (b) A multiple regression model of Y on X_1 , X_2 and X_3 fitted on 20 observations will have 16 degrees of freedom for testing the significance of fitted model.
- (c) If the PRESS statistic and total sum of squares are 8800 and 12800, respectively, the predictive R^2 will be 0.3125.
- (d) If the coefficient of determination for the j th explanatory variable is 0.85, then its corresponding value of tolerance will be 6.67.
- (e) We apply the complementary log-log model in the case of more than three categories of the response variable.

2. The monthly sales (in ₹ '000), price (in ₹) and types of a product are recorded in the following table :

Sales	Price	Type	Sales	Price	Type
186	46	A	208	20	B
218	52	A	248	33	B
182	39	A	335	78	B
241	65	A	264	51	B
224	54	A	242	43	B
193	49	A	268	44	B
262	76	A	315	63	B
240	71	A			

- (i) Fit the multiple linear regression model for sales.
- (ii) Obtain the linear regression model for each group separately. 10
3. To assess the relationship between the heights (Y) and diameters (X) of trees, the model $Y = 1.74 + 0.54 X$ is fitted on the following data :

Tree	Height	Diameter
1	4.14	4.2

	[4]	MST-017
2	4.87	5.55
3	3.24	3.33
4	5.34	6.91
5	4.34	4.4
6	5.07	6.05
7	3.84	3.83
8	5.44	7.11
9	3.54	3.6
10	5.26	6.65

Draw the residual and normal probability plots. 10

4. (a) For a simple linear regression model fitted on 15 observations, compute the values of AIC and BIC. (Given that $SS_{\text{Res}} = 146.3333$). 5
- (b) Differentiate between Forward selection and Backward elimination methods. 5

5. (a) To study the effect of smoking on the lung cancer, the following data are given : 6

	Lung Cancer	No Lung Cancer
Smokers	152	48
Non-smokers	28	172

Fit a logistic model on the given data.
Also obtain the log-odds and odds ratio.

- (b) Explain the polynomial regression with example. 4
6. Write the short notes on the following : 5+5
- (i) Heteroscedasticity
 - (ii) Shrinkage methods

× × × × ×