

**M. SC. (APPLIED STATISTICS)
(MSCAST)**

Term-End Examination

June, 2025

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 :

$$Y = \beta_1 \exp(X_1) + \exp(\beta_2) X_2 + \varepsilon$$

is a linear model.

- (b) If total and residual sum of squares for a dataset are 124 and 15.5, respectively, then the coefficient of determination is 0.875.

- (c) We apply Box-Tidwell approach to obtain suitable variance stabilising transformation on the response variable.
 - (d) The ridge regression shrinks some of the coefficients to exactly zero.
 - (e) We use baseline category logit models when we have ordered categories of a response variable.
2. The systolic blood pressure (mmHg) and age (years) of 12 patients are given as follows :

Patient No.	Systolic Blood Pressure	Age
1	122	28
2	132	36
3	133	37
4	119	24
5	120	27
6	117	25
7	126	30
8	116	23
9	118	24
10	121	29
11	127	35
12	115	23

- (i) Fit an appropriate simple linear regression model, and
- (ii) Test the significance of fitted model at 5% level of significance. (Given that $t_{\text{tab}} = 2.201$) 5+5
3. Explain the Box-Cox approach to transform the response variable. Also explain how to choose the optimum value of λ in this approach. 10
4. The marks obtained in an examination and the time spent on study of 10 students are shown in the following table :

Student	Marks	Time spent in studying
1	80	11
2	90	14
3	76	10
4	82	12
5	54	6
6	59	8
7	88	13
8	43	4
9	48	5
10	55	7

Use Goldfeld-Quandt test to check the heteroscedasticity present in errors of the model. Also, interpret the results. 10

5. The data regarding amount of dose (x_i), total number of patients received medicine (n_i) and number of cured patients (y_i) are recorded in the following table :

S. No.	x	n	y
1	5	24	60
2	10	18	48
3	15	12	40
4	20	20	80
5	25	26	104

Fit a logistic model considering initial values of $\hat{\beta}_0 = -0.2$ and $\hat{\beta}_1 = -0.04$ upto only one iteration. 10

6. (a) Differentiate between leverage and influence points. 5
- (b) Describe Poisson regression model with a suitable example. 5

× × × × ×