

INDIAN STATISTICAL INSTITUTE,
Chennai Centre
M.Stat. : 2015-17
(Year I - Semester I)

Semester Examination: Regression Analysis

Date: 20th November, 2015

Duration: 3 Hours

ANSWER ALL QUESTIONS. TOTAL MARKS IS 100.

1. Consider the binary logistic regression model $\text{logit}(p_i) = -15.0 - 0.4 X_1 + 1.5 X_2$, where p_i denotes the probability of success of an event ($Y = 1$).
 - a) An observation in the dataset is recorded as $y = 0$ for $x_1 = 4$ and $x_2 = 10$. Calculate Pearson residual and deviance residual for this observation.
 - b) Compute Odds ratios and interpret them.

[15]

2. Model Diastolic Blood Pressure as a function of Smoking Status and Cholesterol Level for the following data using logistic regression:

Smoking Status	Cholesterol Level	Diastolic Blood Pressure		
		Low	Normal	High
Non-Smoker	Normal	50	200	150
	High	100	200	200
Smoker	Normal	75	150	300
	High	30	70	200

Compute expected frequency for each cell using the fitted model.

[10 + 10 = 20]

3. Consider the Poisson regression model with log link for the number of mistakes $\ln(\mu_i) = -0.8 + 0.02 X_i$, where X_i denotes the number of pages in a document. In a particular document of 100 pages, 5 mistakes are observed. Compute Pearson residual and deviance residual for this observation. Also, compute the rate ratio and interpret it.

[15]

4. Consider the following summarized data from a study where the response variable (Y) is assumed to be a function of four predictor variables. Actual values of all the variables are transformed using unit length scaling.

$$\text{Sample size} = 15, \quad X^t X = \begin{bmatrix} 1 & 0.23 & -0.8 & -0.25 \\ 0.23 & 1 & -0.15 & -0.97 \\ -0.8 & -0.15 & 1 & 0.03 \\ -0.25 & -0.97 & 0.03 & 1 \end{bmatrix} \quad X^t Y = \begin{bmatrix} 0.73 \\ 0.815 \\ -0.53 \\ -0.821 \end{bmatrix}$$

- Write your observations from the summarized data.
- Find the least square regression model.
- Using the parameters of least square regression model, estimate ridge parameter. Using this ridge parameter value, obtain the ridge regression model.

[10 + 5 + 10 = 25]

5. Describe the formal structure of generalized linear model. Obtain the models of general linear regression and Poisson regression from this structure.

[15]

6. (a) Describe the procedure of obtaining robust regression model using Huber's t – function.

(or)

- (b) Describe the LOESS procedure of non-parametric regression.

[10]
