INDIAN STATISTICAL INSTITUTE,
Chennai Centre
M.Stat. : 2016-18
(Year I - Semester I)

Semester Examination: Regression Analysis

Date: 18th November, 2016 Duration: 3 Hours

1. Consider the following cumulative logit models of an ordinal response variable Y with two continuous variables X_1 and X_2.

$$\log it1 = -6.8 - 0.2X_1 + 0.1X_2$$
$$\log it2 = -3.9 - 0.2X_1 + 0.1X_2$$

(a) Compute the probabilities for different levels of Y at $x_1 = 10$ and $x_2 = 60$.
(b) An observation in the dataset is recorded as $y = 2$ for $x_1 = 7$ and $x_2 = 80$. Calculate Pearson residual and Deviance residual for this observation.
(c) Compute Odds ratios and interpret.

[5+10+5 = 20]

2. a) Describe the formal structure of generalized linear model.
b) Describe the estimation method for logistic regression parameters using IRLS procedure and numerical optimization.

[10+7+8 = 25]

3. Describe the procedure of obtaining robust regression model using Huber's t function.

[10]

4. Find the LOESS estimate of Y for x = 20 using span as 6 for the following data:

Observation	1	2	3	4	5	6	7	8	9	10
x	16.0	24.0	22.0	21.5	18.0	23.2	16.9	21.0	14.0	17.1
y	6.0	2.5	3.6	10.9	15.1	9.5	13.0	14.2	5.2	12.9

[15]

5. 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.

$$\text{Sample Size} = 25,\ X'X = \begin{bmatrix} 1 & 0.073 & 0.124 & -0.010 \\ & 1 & 0.971 & -0.308 \\ & & 1 & -0.270 \\ & & & 1 \end{bmatrix} \quad X'Y = \begin{bmatrix} 0.089 \\ 0.980 \\ 0.959 \\ -0.401 \end{bmatrix}$$

i) Find least square regression model and test for its significance.
ii) Estimate ridge parameter and using that, obtain ridge regression model.
iii) Compute eigen values and eigen vectors. Using that, can you suggest how many PCs are to be considered for Principal Component regression?

[10+10+10 = 30]

***