INDIAN STATISTICAL INSTITUTE
CHENNAI CENTRE
M.Stat. : 2016-18
(Year I – Semester I)

Mid-Semester Examination – Regression Analysis

Date: 8th September 2016 Duration : 3 Hours

ANSWER ALL QUESTIONS. TOTAL MARKS IS 100.

1. Consider the simple linear regression model $Y = \beta_0 + \beta_1 X + \varepsilon$. From a random sample of size 50, the correlation coefficient between X and Y is observed as 0.7 and the lag 1 autocorrelation coefficient of residuals is observed as 0.25.

   a) Test for the significance of the regression model (using ANOVA).
   b) Test for autocorrelation of residuals using DW statistic.

   [10]

2. Consider the multiple linear regression model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$, where Y, X_1 and X_2 all are transformed using unit length scaling. From a random sample of size 30, the sample correlation coefficients are computed as : $r_{x1x2} = 0.1$, $r_{x1y} = 0.8$ and $r_{x2y} = 0.6$.

   a) Find the regression model and test for its significance.
   b) An observation in the above dataset is recorded as $y = 0.65$ for $x_1 = 0.5$ and $x_2 = 0.4$ (all values are scaled). Compute studentized residual, leverage, Cook's D for this observation and interpret them.

   [15 + 15 = 30]

3. Find an appropriate model for the following data:

X	1	2	3	4	5	6	7	8	9	10
Y	150	2000	10000	30000	40000	20000	5000	500	50	5

   Predict the value of Y at $X = 5.5$. [15]

4. Develop an appropriate model using a cubic spline for the following data set:

X	1	2	3	4	5	6	7	8	9	10	11
Y	50	80	90	60	25	5	30	60	150	400	750

[15]

5. Consider the following data on time (X) and oxygen demand (Y). Fit an asymptotic concave non-linear regression model. Mention the starting values of parameters of the chosen model. Compute R^2 of the fitted model.

X	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y	11.8	16	20	24	27	29.5	33	35.1	36.2	38.8	39.3	41	43	45	46.1

[20]

6. Describe on any one of the following:

   a) White's test for checking homoscedasticity of residuals.
   b) Cochran-Orcutt procedure for eliminating autocorrelation.
   c) Methods for detecting multicollinearity.

[10]

*****