LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

B.Sc. DEGREE EXAMINATION – **STATISTICS**

FIFTH SEMESTER – APRIL 2017

ST 5509- REGRESSION ANALYSIS

Date: 26-04-2017   Dept. No.   Max. : 100 Marks
  01:00-04:00

PART – A

Answer ALL questions: **10 X 2 = 20**

1. Define regression.
2. How will you fit a model using R^2?
3. What do you mean by model diagnostics?
4. What is meant by Mean predicted value?
5. Define Partial regression.
6. Define Multiple regression.
7. Define outliers.
8. What is meant by dummy variable trap?
9. Define Multicollinearity.
10. What is meant by Homoscedasticity?

PART – B

Answer any FIVE questions: **5 X 8 = 40**

11. Derive the least square estimates of β in a simple linear regression model.
12. Explain briefly about model diagnostics.
13. Derive MAE and MAPE for multiple linear regression model.
14. Explain briefly about detection of outlier.
15. Explain briefly about removal of heteroscedasticity.
16. Fit a straight line to the following data using least square method.

X	6	2	10	4	8
Y	9	11	5	8	7

17. An incomplete ANOVA table for a regression model $Y = \beta_0 + \beta_1 X + \varepsilon$ with n = 20 is given below.

Source	DF	SS	MS	F
Regression	-	-	-	165.38
Error	-	166255	-	
Total	-	-		

a) Complete the above table. b) Find R^2.

18. Given the following data, calculate the standard error of the estimators of the regression coefficients.

X	6	2	10	4	8
Y	9	11	5	8	7

1

PART – C

Answer any TWO questions: **2 X 20 = 40**

19. a) How do you check the overall fitness of a regression model? **(10 Marks)**

    b) Explain testing the significance of regression coefficients in a linear regression model .

                                                          (10 Marks)

20. a) Explain briefly about Q-Q plot and P-P plot? **(10 Marks)**

    b) Define Mean percentage Error, Mean absolute percentage error and write their uses **(10 Marks)**

21. The following table gives information on ages and cholesterol levels for a random sample of 10 men.

Age	58	69	43	39	63	52	47	31	74	36
Cholesterol level	189	235	193	177	154	191	213	165	198	181

    i). Taking age as an independent variable and cholesterol level as a dependent variable, compute

    SSxx, SSyy, and SSxy.

    ii). Find the regression of cholesterol level on age.

    iii). Calculate r and r^2. .

    iv). Predict the cholesterol level of a 60-year-old man.

    v). Compute the standard deviation of errors.

22. a) Explain briefly about the residual plots.

    b) What are the sources of multicollinearity? Explain the different methods of diagnosing the

    problem of multicollinearity.

****