 LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

**M.Sc.** DEGREE EXAMINATION - **STATISTICS**

FIRST SEMESTER – **NOVEMBER 2012**

# ST 1821 - APPLIED REGRESSION ANALYSIS

Date : 05/11/2012 Dept. No. Max. : 100 Marks

Time : 1:00 - 4:00

Part-A

Answer all the questions: (10x2=20)

1) Define ‘residual’ in a regression model.

2) Explain adjusted R2.

3) What is the variance stabilizing transformation used when σ2 α E(Y) (1-E(Y))?

4) Mention any two sources of multi collinearity.

5) What is the need for standardized regression coefficients?

6) When the regression model is said to be hierarchical?

7) Explain the term auto correlation.

8) Explain AR (1) process.

9) Explain partial correlation coefficient.

10) Explain dummy variable trap.

Part-B

Answer any 5 questions: (5x8=40)

11) How will you verify the assumption of normality and constant variance in a linear regression model? Explain.

12) Consider the model

Y=β0 +β1 x1 +β2 x2 +β3 x3 +ϵ

It is decided to test H0: β1=β3, β2=0

Write the reduced model and the data matrix relevant for the hypothesis, given the data matrix as

X=

13) Explain Studentized Residuals and externally studentized

Residuals.

14) Consider the following ANOVA used for fitting a linear

Regression model with 6 reggressors

ANOVA

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source | df | Sum of squares | Mean square | F |
| Regression | 6 | - | 524.661 | - |
| Residuals | - | 1149 | - |  |
| Total | 29 | | | |

1. Fill in the blanks (4)
2. What is the total number of observations? (1)
3. What conclusion do you draw about the over all fitness of the model? (2)

15) Explain generalized least squares.

16) What are the points to be considered in fitting a polynomial regression model?

17) Explain splines in detail.

18) Explain the random walk model in time series.

Part-C

Answer any 2 questions: (2x20=40)

19)a) An investigator has the following data

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Y | 3.2 | 5.1 | 4.5 | 2.4 |
| X | 5 | 9 | 6 | 4 |

Guide the investigator as to whether the model Y=β0+β1X or Y1/2=β0+β1X is appropriate.

b) Suppose theory suggested that annual income (Y) depended on sex(S), highest degree received (D), and years of experience (E).

The following data is obtained for 10 employees,

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S No. | Y | E | D | S |
| 1 | 13876 | 1 | UG | M |
| 2 | 11608 | 2 | PG | F |
| 3 | 18701 | 1 | PG | M |
| 4 | 11283 | 2 | H.Sc | M |
| 5 | 11767 | 2 | UG | F |
| 6 | 20872 | 2 | PG | M |
| 7 | 11772 | 4 | UG | F |
| 8 | 10535 | 3 | H.Sc | F |
| 9 | 12195 | 3 | PG | M |
| 10 | 12313 | 2 | H.Sc | M |

It is also decided to study the interaction effect of sex with education on Y. Write a suitable linear regression model, with the relevent data matrix.

20) Explain the various methods of diagnosing multicollinearity

and suggest the methods for removing it.

21) Given the following information for fitting a regression

model with 4 regressors. Use forward selection method

to find significant variables that enter at each iteration.

SST=2715.7635 SSRes(x2, x3) =415.4

SSRes(x1) = 1265.6867 SSRes(x2, x4) =868.8

SSRes(x2) = 906.3363 SSRes(x3, x4) =175.7

SSRes(x3) = 1939.4 SSRes(x1, x2, x3) =48.1

SSRes(x4) =883.87 SSRes(x1, x2, x4) =47.9

SSRes(x1, x2) =57.9 SSRes(x1, x3, x4) =50.8

SSRes(x1, x3) =1227 SSRes(x2, x3, x4) =73.8

SSRes(x1, x4) =74.76 SSRes(x1, x2, x3, x4) =47.86

22) a) Explain the methods of studying autocorrelation in a

linear regression model.

b) Explain the Box-Jenkins methodology of ARIMA

modelling.