22P361 (Pages: 2) Name: ……

Reg. No: ……..

THIRD SEMESTER M.Sc. DEGREE EXAMINATION, NOVEMBER 2023

(CBCSS-PG)

(Regular/Supplementary/Improvement)

CC19P MST3 C11 / CC22P MST3 C11 - APPLIED REGRESSION ANALYSIS

(Statistics)

(2019 Admission onwards)

Time: Three Hours Maximum: 30 Weightage

Part A

Answer any ***four*** questions. Each carries 2 weightage.

1. Define a general Gauss Markov linear model.
2. Define hat matrix and state its properties.
3. Consider the regression model $E[Yi] = \beta_0 + \beta_1 x_i + \beta_2(3x_i^2 - 2)(t = 1,2,3)$, where $x_1 = -1, x_2 = 0$, and $x_3 = +1$. Find the least squares estimates of β_0, β_1, and β_2. Show that the least squares estimates of β_0 and β_1, are unchanged if $\beta_2 = 0$.
4. Why is unequal variance a problem? Explain about different variance stabilizing transformations.
5. Explain Mallows- C_p statistics.
6. Discuss the assumptions of the Poisson regression model.
7. Explain the problem of variable selection.

(4 × 2 = 8 Weightage)

Part B

Answer any ***four*** questions. Each carries 3 weightage.

8. Explain the situations in which ordinary least square estimates are not efficient. State and prove Gauss - Markov theorem in the case of generalized least square method.
9. Suppose that $Y \sim N_n(X\beta, \sigma^2 I_n)$, where X is $n \times p$ of rank p.

   (i) Find $\mathrm{Var}(S^2)$.

   (ii) Evaluate $E(Y'A_1Y - \sigma^2)^2$ for $A_1 = \frac{1}{n-p+2}[I_n - X(X'X)^{-1}X']$.

   (iii) Prove that $Y'A_1Y$ is an estimate of a σ^2 with a smaller mean-squared error than S^2.
10. What are the different methods used in the diagnostics of the leverage and influential observations?
11. Differentiate between parametric and non-parametric regression procedures. What are the usual estimation procedures used in non-parametric regression?

12. Explain the concept of autocorrelation and list the commonly used test procedures for detecting the presence of autocorrelation.
13. What is the need for piecewise polynomial fitting? Discuss the method of splines in this context.
14. Explain the problem of regression for binary response variable and develop the method of maximum likelihood to estimate the parameters in a logistic regression model.

(4 × 3 = 12 Weightage)

Part C

Answer any ***two*** questions. Each carries 5 weightage.

15. If $Y \sim N_n(X\beta, \sigma^2 I_n)$, where X is $n \times p$ of rank p, then show that

    (i) $\hat{\beta} \sim N_p(\beta, \sigma^2 (X'X)^{-1})$

    (ii) $\frac{(\hat{\beta}-\beta)'(X'X)(\hat{\beta}-\beta)}{\sigma^2} \sim \chi^2_{(p)}$

    (iii) $\hat{\beta}$ is independent of $S^2 = \frac{RSS}{n-p}$

    (iv) $\frac{RSS}{\sigma^2} = \frac{(n-p)S^2}{\sigma^2} \sim \chi^2_{(n-p)}$.
16. Explain the problem of ill-conditioning in polynomial regression. Describe how orthogonal polynomials can be used to overcome the ill-conditioning. Explain the fitting of orthogonal polynomials.
17. Discuss various scaled residuals. Explain the model adequacy checking using residuals.
18. Explain the parametric estimation and inferential problems on GLM.

(2 × 5 = 10 Weightage)

****