9. Write short notes on any three of the following:
   (a) Orthogonal polynomials
   (b) Estimable functions
   (c) Extra sum of square principle
   (d) Linear hypothesis

(4, 4, 4½)

2/6/22 (M)

[This question paper contains 4 printed pages.]

Your Roll No..............

Sr. No. of Question Paper : 1381 **A**

Unique Paper Code : 32371402

Name of the Paper : Linear Models

Name of the Course : **B.Sc. (H) Statistics under CBCS**

Semester : IV

Duration : 3.5 Hours Maximum Marks : 75

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Attempt **SIX** questions in all.

1. Derive the analysis of variance of two-way classified data with one observation per cell under fixed effects model. Also obtain the expectation due to all different sources of variation. (12½)

2. (a) Develop a prediction interval for the future observation y_0 corresponding to a specified level x_0 of the regressor variable x in the simple linear regression model.

(b) Suppose $X_i, Y_i, Z_i, i = 1,2,\dots,n$ are $3n$ independent observations with common variance σ^2 and expectations $E(X_i) = \theta_1, E(Y_i) = \theta_2, E(Z_i) = \theta_1 - \theta_2, i = 1,2,\dots,n$. Find the BLUEs of θ_1, θ_2. Compute $cov(\widehat{\theta_1}, \widehat{\theta_2})$ and the residual sum of squares. Also find BLUE of $\theta_1 + \theta_2$.

(5,7½)

3. (a) Write a short note on bias in regression estimates? Suppose the postulated model is $E(Y) = \beta_0 + \beta_1 x_1$ but the true model is $(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$. Show that both $\widehat{\beta_0}$ and $\widehat{\beta_1}$ are biased by an amount that depends the values of x's.

(b) We fit a straight line model to a set of data using the formulas $b = (X'X)^{-1}X'Y$ and compute $\hat{Y} = Xb$ with the usual definitions. We define $H = X(X'X)^{-1}X'$. Show that

$$SS(\text{due to regression}) = Y'HY = \hat{Y}'\hat{Y} = \hat{Y}'H^3\hat{Y}.$$

(6,6½)

4. State and prove Cochran's theorem.

(12½)

5. (a) Suppose $Y = (Y_1, Y_2, \dots, Y_n)'$to be a vector of n independent standard normal variates then a necessary and sufficient condition for $Y'AY$ to be distributed as chi-square variate with k d.f. is that A is an idempotent matrix of rank k.

(b) Suppose $y_i, (i = 1,2,\dots,n)$ is a random sample from a standard normal distribution. Show that $\sum_{i=1}^{n} y_i$ and $\sum_{i=1}^{n}(y_i - \bar{y})^2$ are independently distributed.

(8, 4½)

6. Suppose we estimate the following relationship between wages and education controlling for years of experience (a quadratic form) in the workforce:

$Wage = \beta_1 + \beta_2 Educ + \beta_3 Exper + \beta_4 Exper^2 + \varepsilon$,

Using a random sample of 54 observations. The least square estimates and the corresponding estimated covariance matrix are given by

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} -13 \\ 2.5 \\ 0.5 \\ -0.01 \end{bmatrix}, \widehat{Var(b)} = \begin{bmatrix} 4 & -0.2 & -0.125 & 0.01 \\ -0.2 & 0.16 & 0.01 & 0 \\ -0.125 & 0.01 & 0.04 & 0.12 \\ 0.01 & 0 & 0.12 & 0.09 \end{bmatrix}$$

(a) What is the covariance between $\widehat{\beta_1}$ and $\widehat{\beta_4}$?

(b) What is a 95% confidence interval for β_4? (Given $t_{(0.05,50)} = 2.01$)

(c) Is there a statistically significant effect of education on wages in this regression at the 5% level? Give interpretation also.

(d) What is the estimated standard error for $(\widehat{\beta_2} + 3\,\widehat{\beta_3})$?

(e) What is the *t-statistic* from a test that the marginal impact of years of experience on wages equals zero for someone with two years of experience?

(1, 3½, 3, 3½, 1½)

7. (a) Suppose $\underset{\sim}{Y} \sim N_3(\underset{\sim}{0}, I)$ and let $A = \frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$.

(i) What is the distribution of $Y'AY$ and $Y'(I-A)Y$?

(ii) Are the two quadratic forms defined in part (a) independent?

(iii) Are $\underset{\sim}{Y}'A\underset{\sim}{Y}$ and $y_1+y_2+y_3$ independent?

(b) Write the simple linear regression model in matrix notation. Hence obtain the least squares estimators of the unknown parameters and their variances.

(6, 6½)

8. (a) If you fit the models (i) $Y = \beta_0 + \beta_1 x + \varepsilon$ and (ii) $Y = \beta_1 x + \varepsilon$ for a given data, with usual assumptions, which of the above models is appropriate, justify your choice? Give three reasons for your justification.

(b) What do you understand by analysis of covariance? State the mathematical model used in analysis of covariance for one way classification with one covariate. Explain the hypothesis to be used.

(c) Explain the significance of partial F-test and sequential F-test.

(5, 4, 3½)