[This question paper contains 2 printed pages.]

Sr. No. of Question Paper : **6708** **D** **Your Roll No...............**

Unique Paper Code : 237503

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

Name of the Paper : STHT – 503 : Linear Models

Semester : V

Time : 3 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, selecting **three** questions from each Section.

SECTION I

1. Consider the general linear model

$$\underset{\sim}{Y}_{n\times 1} = X_{n\times p}\underset{\sim}{\beta}_{p\times 1} + \underset{\sim}{\epsilon}_{n\times 1}$$

with $E\left(\underset{\sim}{\epsilon}\right) = 0$, $V\left(\underset{\sim}{\epsilon}\right) = \sigma^2 I$ and $\rho(X) = k < p \leq n$. Let $\Psi = \underset{\sim}{c}'\underset{\sim}{\beta}$ be an estimable function, then in the class of all linear unbiased estimators of $\underset{\sim}{\Psi}$ obtain the BLUE $\underset{\sim}{\Psi}$ by the method of least squares. Also obtain an unbiased estimator of σ^2. (12½)

2. State and prove Cochran's theorem. (12½)

3. Derive the analysis of covariance for two way classified data with one covariate. (12½)

4. (a) Consider the model

$E(Y_1) = \beta_1 + \beta_2,$
$E(Y_2) = 2\beta_1,$
$E(Y_3) = \beta_1 - \beta_2$

with usual assumptions. Find sum of squares due to error.

(b) Suppose $y_i (i = 1, 2, ..., n)$ 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.

(c) Obtain the variances of the estimated parameters in ANOVA for one way classification under fixed effects model. (4,4,4½)

SECTION II

5. (a) Stating clearly the underlying assumptions of the simple linear regression model with no intercept term, obtain the least square estimate of the regression parameter along with its variance.

(b) 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. (6,6½)

6. (a) Write short notes on the following :

(i) Residual analysis

(ii) Use of orthogonal polynomials in fitting polynomial models

(b) Suppose that we have fitted the straight-line model $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1$ but the response is affected by a second variable x_2 whereas the true regression function is $E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.

(i) Is the least-squares estimator of the slope in the original simple linear regression model unbiased ?

(ii) Obtain the bias in $\hat{\beta}_1$. (7,5½)

7. (a) Discuss the problem of testing for lack of fit in simple linear regression model.

(b) Consider the simple linear regression model

$$Y = \beta_0 + \beta_1 x + \epsilon$$

with $E(\epsilon) = 0$, $V(\epsilon) = \sigma^2$ and ϵ's are uncorrelated. Show that

(i) $E(MSE) = \sigma^2$, and

(ii) $E(MSR) = \sigma^2 + \beta_1^2 S_{xx}$ (5½,7)

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

(b) What is a parametric function ? Derive a necessary and sufficient condition for which a parametric function is estimable. (7,5½)

[This question paper contains 2 printed pages.]

Sr. No. of Question Paper : 1193 E **Your Roll No.**...............

Unique Paper Code : 237503

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

Name of the Paper : STHT-503 : Linear Models

Semester : V

Duration : 3 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, selecting **three** from each Section.

SECTION I

1. (a) Let $Y = (Y_1, Y_2, ..., Y_n)'$ be a vector of n independent standard normal variates. Let $Q_1 = \underset{\sim}{Y}' A_1 \underset{\sim}{Y}$ and $Q_2 = \underset{\sim}{Y}' A_2 \underset{\sim}{Y}$ be distributed as χ^2 with n_1 and n_2 degrees of freedom respectively. Show that the necessary and sufficient condition for Q_1 and Q_2 to be independently distributed is $A_1A_2 = 0$.

   (b) Write note on Coefficient of Determination. (9,3½)

2. Derive the analysis of variance for two way classified data with one observation per cell under fixed effects model. (12½)

3. For a given model $\underset{\sim}{Y}_{n\times 1} = X_{n\times p}\underset{\sim}{\beta}_{p\times 1} + \underset{\sim}{\epsilon}_{n\times 1}$ with $E(\underset{\sim}{\epsilon}) = \underset{\sim}{0}$, $V(\underset{\sim}{\epsilon}) = \sigma^2 I$ and $\rho(X) = r < p$, prove that, there are r independent estimable functions of the parameters for the model. Obtain the BLUE of these estimable functions and also their variances. (12½)

4. Eight experiments are to done at the coded level $(\pm 1, \pm 1)$ of two predictor variables x_1 and x_2. Two experimenters A and B suggest the following designs :

A: Take one observation at each of $(x_1, x_2) = (-1, -1)$ and $(1,1)$ and three observations at each of $(-1,1)$ and $(1,-1)$.

B: Take two observations at each of four sites.

If a model $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \epsilon$ is to be fitted by least squares but it is feared there may be some additional quadratic curvature expressed by extra terms $\beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2$, evaluate the anticipated biases in the estimated coefficients b_0, b_1 and b_2 for each design. (12½)

SECTION II

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

   (b) Obtain the confidence interval on mean response at a particular point under simple linear model. (6½,6)

6. (a) For the general linear model $\underset{\sim}{Y}_{n\times 1} = X_{n\times p} \underset{\sim}{\beta}_{p\times 1} + \underset{\sim}{\epsilon}_{n\times 1}$, set the appropriate hypothesis for testing the significance of regression and develop the test for significance of individual regression coefficients. Explain why this test is known as marginal test.

   (b) Write note on stepwise regression method. (8,4½)

7. (a) Discuss the problem of testing for lack of fit in a simple linear regression model.

   (b) Write note on the Orthogonal columns in X matrix. (7,5½)

8. (a) Suppose X_i, Y_i, Z_i, $i = 1, 2, ..., n$ are 3n independent observations with common variance σ^2 and expectations $E(X_i) = \theta_1$, $E(Y_i) = \theta_2$, $E(Z_i) = \theta_1 - \theta_2$, $i = 1, 2,..., n$. Find the BLUEs of θ_1, θ_2 and $\theta_1 + \theta_2$. Also find $\text{cov}(\hat{\theta}_1, \hat{\theta}_2)$ and compute the residual sum of squares.

   (b) Explain partial F-test and sequential F-test. (9,3½)

[This question paper contains 2 printed pages.]

Sr. No. of Question Paper : 6236 F-5 Your Roll No...............

Unique Paper Code : 2371502

Name of the Paper : Linear Models

Name of the Course : **Erstwhile FYUP B.Sc (H) Statistics**

Semester : V

Duration : 3 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, selecting **three** questions from each section.

SECTION I

1. Derive the analysis of covariance for two way classified data with one covariate. (12½)

2. State and prove Cochran's theorem. (12½)

3. Consider the general linear model

$$Y_{n\times 1} = X_{n\times p}\, \beta_{p\times 1} + \varepsilon_{n\times 1}$$

with $E(\varepsilon)=0$, $V(\varepsilon)=\sigma 2\ I_n$ and $\rho(X)=p<n$, where Y is observed, X is known and β and $\sigma 2$ are unknown. Show that the BLUE of β is given by the method of least squares. Also obtain an unbiased estimator of $\sigma 2$. (12½)

4. (a) Suppose the hypothesis of homogeneity of k-treatment means is rejected in ANOVA testing for one way classification under fixed effect model, how would you proceed to test the hypothesis of equality of two specific treatment means?

(b) Suppose $Y \sim N_3(0,I)$ and let

$$A=\frac{1}{3}\begin{bmatrix} 1 & 0 & -\sqrt{2} \\ 0 & 0 & 0 \\ -\sqrt{2} & 0 & 2 \end{bmatrix}$$

P.T.O.

Find the distribution of $Y'AY$, stating the appropriate theorem to be used and also find the distribution of $Y'(I-A)Y$. Are they independent? (6½,6)

SECTION II

5. (a) Write a short note on bias in regression estimates? Suppose the postulated model is $E(Y) = \beta_1 x$ but the true model is $E(Y) = \beta_0+\beta_1 x$. Show that $\hat{\beta}_1$ is biased by an amount that depends on β_0 and the values of x's.

   (b) Explain the role of orthogonal polynomials in fitting polynomial models in one variable. (8½,4)

6. (a) For the simple linear regression model $Y_i = \beta_0+\beta_1 x_i + \epsilon_i$, $(i = 1,2, \ldots , n)$, where $\epsilon_i \sim NID(0, \sigma^2)$. Explain the hypotheses testing on β_0 and β_1. Also obtain the confidence interval on β_0 and β_1.

   (b) For the simple linear regression model $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, $(i = 1, 2, \ldots, n)$, where $\epsilon_i \sim NID(0, \sigma^2)$,

   (i) Show that $\operatorname{cov}(\bar{y},\hat{\beta}_1) = 0$.

   (ii) Show that $\operatorname{cov}\left(\hat{\beta}_0,\hat{\beta}_1\right) = -\dfrac{\sigma 2\bar{x}}{S_{xx}}$. (6½,6)

7. (a) Stating clearly the underlying assumptions of the simple linear regression model through the origin, obtain the least squares estimate of the regression parameter along with its variance.

   (b) If you fit the models (i) $Y = \beta_0 + \beta_1 x + \epsilon$ and (ii) $Y = \beta_1 x + \epsilon$ for a given data. Which of the above models is appropriate, justify your choice? Give three reasons for your justification. (6½,6)

8. (a) Consider a linear model $Y_{n\times 1} = \underset{n\times p-r}{X_1}\ \underset{p-r\times 1}{\beta_1} + \underset{n\times r}{X_2}\ \underset{r\times 1}{\beta_2} + \varepsilon_{n\times 1}$ with usual assumptions.

   How will you test the significance of the contribution of regressors in X_2 when other regressors in X_1 are already in the model?

   (b) Write notes on the following

   (i) Coefficient of determination

   (ii) Residual analysis (6½,6)

(150)