This question paper contains 3 printed pages.

Your Roll No. ....

Sl. No. of Ques. Paper	: **1389**
Unique Paper Code	: **2371502**
Name of Paper	: **Linear Models**
Name of Course	: **B.Sc. (Hons.) Statistics (Erstwhile FYUP)**
Semester	: **V**
Duration	: **3 hours**
Maximum Marks	: **75**

F-7

(Write your Roll No. on the top immediately on receipt of this question paper.)

*Attempt six questions in all, selecting **three** questions from each Section.*

SECTION I

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

2. 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 $\rho(X)=k<p<n$. Let $\psi=c'\beta$ be an estimable function, then in the class of all linear unbiased estimators of ψ obtain the BLUE of ψ by the method of least squares. Also obtain an unbiased estimator of σ^2. 12½

3. (a) Suppose $Y\sim N_3(0,I)$ and let $A=\frac{1}{3}\begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}$. Are $Y'AY$ and $y_1+y_2+y_3$ independent?

(b) Suppose $Y=(Y_1, Y_2, \ldots\ldots, Y_n)'$ 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. 3½, 9

4. (a) Four objects A, B, C, D are involved in a weighing experiment. Put together they weighed Y_1 grams ; when A and C are put in the left pan of the balance and B and D are put in the right pan, a weight of Y_2 grams was necessary in the right pan for the balance. With A and B in the left pan and C and D in the right pan, Y_3 grams were needed in the right pan. Finally A and D in the left pan and B and C in the right pan, Y_4 grams were needed in the right pan. If the observations Y_1, Y_2, Y_3, Y_4 are all subject to uncorrelated errors with common variance σ^2, obtain the BLUEs of individual weights and the total weight of the four objects, and variance of the estimate of the total weight of four objects.

(b) Show that, for a general linear model $\underset{n\times1}{Y} = \underset{n\times p}{X}\ \underset{p\times1}{\beta} + \underset{n\times1}{\varepsilon}$:

$$\sum_{i=1}^{n} V(\bar{Y}_i)/n = \text{trace}\,\frac{\{x(x'x)^{-1}x'\}\sigma^2}{n} = p\sigma^2/n$$

(c) 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. 5,3½,4

SECTION II

5. (a) For the simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \in_i$, $(i=1, 2, ..., n)$, where $\in_i \sim$ NID$(0, \sigma^2)$. Obtain the least square estimates of β_0 and β_1. Also verify the bias and variance properties of $\hat{\beta}_0$ and $\hat{\beta}_1$.

(b) Consider the simple linear regression model $Y = \beta_0 + \beta_1 X + \in$ with $E(\in) = 0$, $V(\in) = \sigma^2$ and $\in$'s are uncorrelated. Show that:

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

(ii) $E(MSR) = \sigma^2 + \beta_1^2 S_{xx}$. 6½, 6

6. (a) For the simple linear regression model $y_i = \beta_0 + \beta_1 x_i + \in_i$, $(i=1, 2, ..., n)$, where $\in_i \sim$ NID$(0, \sigma^2)$, obtain the maximum likelihood estimates of β_0, β_1 and σ^2. Also show the amount of bias in the estimate of σ^2.

(b) What is a parametric function? Derive a necessary and sufficient condition for which a parametric function is estimable. Consider the model $E(Y_{ij}) = \alpha_i + \beta_j$, $i=1, 2; j=1, 2$. Find the condition under which $l_1\alpha_1 + l_2\alpha_2 + m_1\beta_1 + m_2\beta_2$ is an estimable function. 6½,6

7. (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 that we are fitting a straight line and wish to make the standard error of the slope as small as possible. Suppose that the "region of interest" for x is $-1 \le x \le 1$. Where should the observations $x_1, x_2, ..., x_n$ be taken? Discuss the practical aspects of this data collection plan. 6, 6½

8. (a) Write a short note on bias in regression estimates. Suppose the postulated model is $E(Y)=\beta_0+\beta_1x_1$ but the true model is $E(Y)=\beta_0+\beta_1x_1+\beta_2x_2$. Show that both $\widehat{\beta_0}$ and $\widehat{\beta_1}$ are biased by an amount that depends on the values of x's.

(b) We fit a straight line model to a set of data using the formulas:

$$b=(X'X)^{-1}X'Y,\ \bar{Y}=Xb$$

with the usual definitions. We define $H=X(X'X)^{-1}X'$. Show that:—

$$\text{SS(due to regression)}=Y'HY=\hat{Y}'\hat{Y}=\hat{Y}'H^3\hat{Y}$$ 8½,4