

AUG 2023

Time : 3 hours

Total Marks = 75

Candidates may attempt at most **three** questions. Full marks can be obtained for complete solutions to **three** questions. Each question should start on a fresh page.

- A1. (a) Briefly discuss areas of multivariate methods applications in real life. [8]
 (b) Comment on the use of the hotelling T^2 statistic. [2]
 (c) Discuss any four methods of displaying multivariate data. [5]
 (d) Describe the used of multivariate control charts. [10]

- A2. In genetics, it is often important to track the inheritance of characteristics that can be measured several time during an animal lifetime. For sample of $n = 150$, body weight of mice were obtained immediately after birth for the four litters. The sample mean vector and the correlation matrix are given;

$$\bar{x}' = [39.88 \quad 45.08 \quad 48.11 \quad 49.95],$$

$$R = \begin{bmatrix} 1.000 & 0.7501 & 0.6329 & 0.6363 \\ 0.750 & 1.000 & 0.6925 & 0.7386 \\ 0.6329 & 0.6925 & 1.000 & 0.6625 \\ 0.6363 & 0.7386 & 0.6625 & 1.000 \end{bmatrix},$$

$$\lambda_1 = 3.085, \lambda_2 = 0.382, \lambda_3 = 0.342 \text{ and } \lambda_4 = 0.0.217.$$

- (a) Interpret the results fully. [20]
 (b) Sketch the scree plot and comment. [5]
- A3. (a) What are the assumptions of multiple linear regression? [4]
 (b) Comment why the Pearson correlation coefficient is not sufficient in model building. [3]

- (c) The following five measurements were obtained for the dependent variable y and two independent variables x_1 and x_2 .

x_1	9	2	6	5	8
x_2	12	8	6	4	10
y	3	4	0	2	1

Use the matrix method to estimate a regression model for the dataset. [12]

- (d) State the null and alternative hypotheses for testing the significance of all the model parameters involved. [6]

- A4. A selection of seven receipts from the university bookstore was obtained in order to investigate the nature of book sales. Each receipt provided among other things, the number of books sold and the total amount of each sale. the results are presented below.

Consider the following measurements on two variables;

variable 1 (dollar sales); 25 32 45 28 28 60

and

variable 2 (number of books); 5 4 6 3 2 9 Calculate;

- (a) the mean vector, [4]
 (b) variance-covariance matrix, [10]
 (c) the correlation matrix and comment on your result, [6]
 (d) state any hypotheses that can be tested involving the mean vector. [5]

- A5. (a) Distinguish classification and discrimination. [3]
 (b) Consider the two datasets.

$$\mathbf{X}_1 = \begin{bmatrix} 3 & 7 \\ 2 & 4 \\ 4 & 7 \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} 6 & 9 \\ 5 & 7 \\ 4 & 8 \end{bmatrix} \text{ and}$$

$$\mathbf{S}_{pooled} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

- (i) Calculate the linear discriminant function. [15]
 (ii) Classify the object $\begin{bmatrix} 2 \\ 7 \end{bmatrix}$. [7]

END OF QUESTION PAPER