

MScEdMAE

MULTIVARIATE ANALYSIS

Time : 3 hours

AUG 2024

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

A1. The following are five measurements on the variables x_1 , x_2 , and x_3 :

x_1	9	2	6	5	8
x_2	12	8	6	4	10
x_3	3	4	0	2	1

- (a) Find the arrays \bar{x} , S and R . [6, 10, 3]
- (b) Represent the data in Box and Whisker diagrams on the same scale and comment the results. [3, 3]

A2. Describe the significance of the following quality control instruments

- (a) Q-Q plots. [5]
- (b) Multivariate control charts. [10]
- (c) Simultaneous confidence intervals. [5]
- (d) Assumptions. [5]

- A3. (a) Evaluate T^2 , for testing $H_0 : \mu = [7, 9]$, using the data $\begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$. [10]
- (b) Specify the distribution of T^2 for the situation in (a). [4]
- (c) Using (a) and (b), test the hypothesis. What Conclusion do you reach? [8, 2]

- A4. (a) Twenty-five portfolio managers were evaluated in terms of their performance. Suppose Y represents the rate of return achieved over a period of time, Z_1 is the manager's attitude toward risk measured on a five-point scale from "very conservative" to "very risky," and Z_2 is years of experience in the investment business. The observed correlation coefficients between pairs of variables are

$$R = \begin{matrix} & Y & Z_1 & Z_2 \\ \begin{matrix} Y \\ Z_1 \\ Z_2 \end{matrix} & \begin{pmatrix} 1.0 & -0.35 & 0.82 \\ -0.35 & 1.0 & -0.60 \\ 0.82 & -0.60 & 1.0 \end{pmatrix} \end{matrix}$$

Interpret the sample correlation coefficients $r_{yz_1} = -0.35$ and $r_{yz_2} = -0.82$. [6]

- (b) 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. [15]

- (c) state the assumptions of multiple regression. [4]

- A5. (a) Define classification as a multivariate technique. [2]

- (b) A biometrics investigating levels of citric acid in lemons from two different plantations, measured the acid levels of 40 lemons from plantation A and 50 lemons from plantation B. The following results were obtained (negative values reflect low acidity and conductivity levels while positive values reflect high values).

$$\bar{\mathbf{X}}_A = \begin{pmatrix} -0.0065 \\ -0.0390 \end{pmatrix}, \bar{\mathbf{X}}_B = \begin{pmatrix} -0.2483 \\ 0.0262 \end{pmatrix}, \mathbf{S}_{pooled}^{-1} = \begin{pmatrix} 131.158 & -90.423 \\ -90.423 & 108.417 \end{pmatrix}.$$

- (i) Obtain the Fisher's linear discriminant function. [10]
(ii) Construct a rule for allocating a new lemon $\mathbf{X}'_0 = \begin{pmatrix} X_{01} & X_{02} \end{pmatrix}$ and use it to classify a lemon with acidity level -0.210 and conductivity -0.044. [13]

END OF QUESTION PAPER