Chennai Mathematical Institute
Multivariate Statistical Analysis
End Semester Examination

Time: 3 hours **Maximum Marks: 30**

PART-A: Answer ALL questions (3 x 5 = 15)

1. Let X_1, X_2, X_3, X_4, and X_{15} be independent and identically distributed random vectors with mean vector μ and covariance matrix ∑. Find the mean vector and covariance matrices for each of the two linear combinations of random vectors

$0.2X_1 + 0.2X_2 + 0.2\,X_3 + 0.2X_4 + .2X_5$ and $X_1 - X_2 + X_3 - X_4 + X_5$

in terms of the mean vector μ and the covariance matrix ∑. Also, find the covariance between the two linear combinations of random vectors.

2. Given the data

$$X = \begin{bmatrix} 2 & 12 \\ 8 & 9 \\ 6 & 9 \\ 8 & 10 \end{bmatrix}$$

(a) Evaluate T^2 for testing *Ho:* $\mu' = [7, 11]$, using the data

(b) Specify the distribution of T^2 for the situation in (a).

(c) Using (a) and (b), test *Ho* at the $\alpha = .05$ level. What conclusion do you reach?

3. Consider the two data sets

$$X_1 = \begin{bmatrix} 3 & 7 \\ 2 & 4 \\ 2 & 7 \end{bmatrix} \quad \text{and} \quad X_2 = \begin{bmatrix} 6 & 9 \\ 5 & 7 \\ 4 & 8 \end{bmatrix}$$

for which $\bar{x}_1 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$ $\bar{x}_2 = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$ and $S_{pooled} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

(a) Calculate the linear discriminant function.

(b) Classify the observation $x_0' = [2\ 7]$ as population $\prod_1$ or population $\prod_2$ with equal priors and equal costs.

PART-B
Answer the following (1 x 15 = 15)

The following data pertains to house price data in Boston. Examine the data and compare using the Maximum Likelihood Method(MLM)(with and without varimax), Principal Component method(PCM), and Principal Factor Method (PFM). Interpret the results. (Exclude CHAS from the data analysis.)