LOYOLA COLLEGE (AUTONOMOUS) CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – **STATISTICS**

THIRD SEMESTER – **APRIL 2025**

PST3MC01 – MULTIVARIATE ANALYSIS

Date: 24-04-2025 Dept. No. Max. : 100 Marks

Time: 09:00 AM - 12:00 PM

	SECTION A – K1 (CO1)	
	Answer ALL the questions	**(5 x 1 = 5)**
1	**Fill in the blanks**	
a)	The multiple correlation coefficient, denoted as RRR, measures the strength of the relationship between a dependent variable and ……. independent variables.	
b)	The formula for Mahalanobis Distance measure is defined as ……..	
c)	In a Repeated Measures design, the same subjects are tested multiple times under different …….	
d)	In Canonical Correlation Analysis, the primary objective is to maximize the …….. between the linear combinations of the two sets of variables.	
e)	……. is a diagrammatic representation of cluster formation in hierarchical clustering	
	SECTION A – K2 (CO1)	
	Answer ALL the questions	**(5 x 1 = 5)**
2	**Definition**	
a)	Cumulative distributive function.	
b)	Partial correlation coefficient.	
c)	Sphericity test.	
d)	Varimax rotation.	
e)	Discriminant analysis.	
	SECTION B – K3 (CO2)	
	Answer any THREE of the following	**(3 x 10 = 30)**
3	If X is a 5x1 vector which is distributed as $N_5(\mu,\sum)$, find the distribution of $\begin{bmatrix} X_2 \\ X_4 \end{bmatrix}$.	
4	The following are the four measurement on the variables X_1 and X_2. X_1: 42, 52, 48, 58 X_2: 4, 5, 4, 3 Calculate the correlation matrix for the above data.	
5	Explain the procedure of MANOVA and repeated measures design.	
6	Discuss varimax rotation and state its uses.	

7	Consider the two dataset $X_1 = \begin{bmatrix} 3 & 7 \\ 2 & 4 \\ 7 & 8 \end{bmatrix}$, $X_2 = \begin{bmatrix} 6 & 9 \\ 5 & 7 \\ 4 & 8 \end{bmatrix}$ and $S_{pool} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. Calculate the linear discriminant function classify the observations $x_0' = [2 \quad 7]$ *and* $[5 \quad 9]$.

SECTION C – K4 (CO3)

	Answer any TWO of the following	**(2 x 12.5 = 25)**
8	Discuss the sample correlation coefficient in a multivariate context, and elaborate its properties.	
9	How does Naïve Bayesian algorithm work for classification problem.	
10	Discuss MANOVA for comparing g population mean vectors in detail.	
11	Explain the procedure of k-means clustering and Pseudo F-statistic for optimizing the number of clusters.	

SECTION D – K5 (CO4)

	Answer any ONE of the following	**(1 x 15 = 15)**
12	Obtain Bivariate Normal distribution from multivariate normal density by substituting p=2.	
13	Calculate partial correlation coefficient $r_{13.2}$ using Pearson correlation approach for the following data:	

X_1	1	2	3	5	3	1	5
X_2	1	1	1	1	2	2	3
X_3	2	3	6	6	8	6	2

SECTION E – K6 (CO5)

	Answer any ONE of the following	**(1 x 20 = 20)**
14	The following are the measurements with the variables x_1, x_2 and x_3:	

x_1	53	35	33	50	33	53	43
x_2	49	44	50	47	52	44	50
x_3	57	35	58	52	36	45	56

Estimate the mean vectors, covariance matrix and correlation matrix. (4+8+8)

15 Find the clusters using average linkage techniques. Use Euclidean distance and draw the dendrogram.

X_1	0.40	0.22	0.35	0.26	0.08	0.45
X_2	0.53	0.38	0.32	0.19	0.41	0.30

$$$$$$$$$$$$