Name: **Enrolment No:**	

UPES

End Semester Examination, December 2023

Course: Multivariate Statistics — **Semester : VII**

Program: B.SC(Mathematics by Research) — **Time : 03 hrs.**

Course Code: MATH4015P — **Max. Marks: 100**

Instructions: All questions are compulsory.

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q 1	Find b and A such that the density function $\frac{1}{2\pi} exp\{-\frac{x^2+y^2+4x-6y+13}{2}\}$ can be written of normal density function.. Also find $\mu_x, \mu_y, \sigma_x, \sigma_y, \sigma_{xy}$.	**4**	**CO1**
Q 2	Show that $d^2(x,y) = \Sigma_{j=1}^{p}\left(x_j - y_j\right)^2$ is equal to $d^{(x,y)} = \left(v_x - v_y\right)^2 + p(\bar{x}-\bar{y})^2 + 2v_x v_y\left(1 - r_{xy}\right)$, where $v_x^2 = \Sigma_{j=1}^{p}\left(x_j - \bar{x}\right)^2, \bar{x} = \Sigma_{j=1}^{p} x_j/p$ and r_{xy} is the correlation.	**4**	**CO2**
Q 3	Consider a sample data involving 3 variables with mean and covariance as follows: $\bar{y} = \begin{pmatrix} 28.1 \\ 7.18 \\ 3.09 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 140.54 & 49.68 & 1.94 \\ 49.68 & 72.25 & 3.68 \\ 1.94 & 3.68 & 0.25 \end{pmatrix}$ Test H_0 for $\mu = (15,\ 6,\ 2.85)'$. (Use $T^2_{.05,3,9} = 16.766$)	**4**	**CO2**
Q 4	Show that $E[\hat{y}_i - E(y_i)][\hat{y}_i - E(y_i)]' = E[\hat{y}_i - E(\hat{y}_i)][\hat{y}_i - E(\hat{y}_i)]' + [E(\hat{y}_i) - E(y_i)][E(\hat{y}_i) - E(y_i)]'$	**4**	**CO2**
Q 5	Show that $b_{1,p}$ and $b_{2,p}$ are invariant under the transformation $z_i = Ay_i + b$, where A is non singular.	**4**	**CO3**

SECTION B
(4Qx10M= 40 Marks)

S. No.		Marks	CO
Q 6	Find the covariance matrix for the following data:	**10**	**CO1**

Person	Height (x)	Weight (y)
1	69	153
2	74	175
3	68	155
4	70	135
5	72	172
6	67	150
7	66	115
8	70	137

9	76	200
10	68	130

Q 7

Suppose a data is given for five variables $y_i, (i = 1,2,..,5)$ whose sample mean vector and covariance matrix are given as follows:

$$\bar{y} = \begin{pmatrix} 36.09 \\ 25.55 \\ 34.09 \\ 27.27 \\ 30.73 \end{pmatrix}, S = \begin{pmatrix} 65.09 & 33.65 & 47.59 & 36.77 & 25.43 \\ 33.65 & 46.07 & 28.95 & 40.34 & 28.36 \\ 47.59 & 28.95 & 60.69 & 37.37 & 41.13 \\ 36.77 & 40.34 & 37.37 & 62.82 & 31.68 \\ 25.43 & 28.36 & 41.13 & 31.68 & 58.22 \end{pmatrix}$$

Suppose $z = 3y_1 - 2y_2 + 4y_3 - y_4 + y_5, w = y_1 + 3y_2 - y_3 + y_4 - 2y_5$ be two linear functions. Find the correlation between z and w.

10 | **CO2**

Q 8

Let $f(x, y) = \begin{cases} 2, & 0 \le x \le 1, 0 \le y \le 0 \\ 0, & otherwise \end{cases}$

Find:

a. $F(x, y)$
b. $F(x)$
c. $f(x)$
d. $f(x|y)$
e. $E(X^n Y^m)$

Prove that X and Y are independent.

10 | **CO1**

Q 9

Find the maximum distance by single linkage clustering for the following distance matrix:

City	Distance					
Atlanta	0	536.6	516.4	590.2	693.6	716.2
Boston	536.6	0	447.4	833.1	915	881.1
Chicago	516.4	447.4	0	924	1073.4	971.5
Dallas	590.2	833.1	924	0	527.7	464.5
Denver	693.6	915	1073.4	527.7	0	358.7
Detroit	716.2	881.1	971.5	464.5	358.7	0

OR

10 | **CO3**

	Find the maximum distance Ward's method for the following distance matrix:		

City	Distance					
Atlanta	0	536.6	516.4	590.2	693.6	716.2
Boston	536.6	0	447.4	833.1	915	881.1
Chicago	516.4	447.4	0	924	1073.4	971.5
Dallas	590.2	833.1	924	0	527.7	464.5
Denver	693.6	915	1073.4	527.7	0	358.7
Detroit	716.2	881.1	971.5	464.5	358.7	0

SECTION-C

(2Qx20M=40 Marks)

Q 10	Define the following terms: a. Marginal Distribution b. Statistical Independence c. Maximum Likelihood Estimators d. Generalized Variance e. Canonical correlation	**20**	**CO1**
Q 11	Find five points in two dimension such that the interpoint distances d_{ij} in two dimensions are approximately equal to the values of δ_{ij} in D. $D = (\delta_{ij}) = \begin{pmatrix} 0 & 2\sqrt{2} & 2\sqrt{2} & 2\sqrt{2} & 2\sqrt{2} \\ 2\sqrt{2} & 0 & 4 & 4\sqrt{2} & 4 \\ 2\sqrt{2} & 4 & 0 & 4 & 4\sqrt{2} \\ 2\sqrt{2} & 4\sqrt{2} & 4 & 0 & 4 \\ 2\sqrt{2} & 4 & 4\sqrt{2} & 4 & 0 \end{pmatrix}$ OR Given the measurements on the first and second adult sons in a sample of 10 families.	**20**	**CO3**

First Son		Second Son	
Head Length y_1	Head Breadth y_2	Head Length x_1	Head Breadth x_2
191	155	179	145
195	149	201	152
181	148	185	149
183	153	188	149
176	144	171	142
208	157	192	152
189	150	190	149
197	159	189	152
188	152	197	159
192	150	187	151

Test independence of (y_1, y_2) and (x_1, x_2) for the sons data.