Total No. of Questions – [3]

Total No. of Printed Pages: 02

PRN No.	22110544

PAPER CODE	U313-214 (IE)

October 2023 (INSEM) EXAM

TY B.TECH (Artificial Intelligence & Data Science) (SEMESTER - I)

COURSE NAME: MULTIVARIATE ANALYSIS **COURSE CODE: ES31204AD**

(PATTERN 2020)

Time: [40 Min] **[Max. Marks: 20]**

(*) Instructions to candidates:

1) **Figures to the right indicate full marks.**
2) **Use of scientific calculator is allowed**
3) **Use suitable data wherever required**
4) **Solve any two sub questions from Question 1 and 2 and any one sub question from Question 3**

<table>
<tr><th>Q. No.</th><th>Question Description</th><th>Max. Marks</th><th>CO mapped</th><th>BT Level</th></tr>
<tr><td>Q.1</td><td>a) Describe the different data types commonly encountered in multivariate analysis. Provide examples of categorical data types.

b) Consider a dataset containing information about sales (in thousands of dollars), advertising expenditure (in thousands of dollars and market shared (in percentage) for a set of products in particular industry. Calculate the mean and standard deviation of sales (X1), advertising expenditure (X2) and market share (Y)

c) You have collected data on the heights (in inches), weights (in pounds), and ages (in years) of 30 individuals. Calculate the covariance matrix for these three variables based on the following data:
Heights: [65, 68, 72, 61, 70, 74, 63, 67, 71, 69, 72, 65, 68, 73, 62, 70, 75, 64, 68, 71, 66, 69, 72, 67, 70, 74, 66, 71, 75, 68]
Weights: [150, 155, 175, 140, 165, 180, 145, 155, 170, 160, 175, 150, 155, 185, 140, 170, 190, 145, 155, 170, 150, 160, 175, 155, 165, 185, 150, 170, 190, 155]
Ages: [25, 30, 35, 20, 28, 38, 22, 31, 36, 29, 37, 24, 29, 39, 21, 28, 40, 23, 31, 35, 26, 32, 38, 30, 33, 40, 27, 36, 40, 32]</td><td>[4]

[4]

[4]</td><td>[CO1]

[CO1]

[CO1]</td><td>[1] Remember

[3] Apply

[3] Apply</td></tr>
</table>

Product	Sales (X1)	Advertising Expenditure (X2)	Market Share (Y)
A	50	10	12%
B	45	12	9%
C	60	8	15%
D	55	11	11%
E	48	9	10%

Q2	a) Given a dataset with three variables: X, Y, and Z, compute the generalized variance (GV) for this dataset. Assume you have the following data: X: [10, 15, 20] Y: [5, 8, 12] Z: [7, 10, 14]	[4]	[CO2]	[3] Apply
	b) Given two random variables, X and Y, with the following sample values: X: [2, 4, 6, 8] Y: [1, 3, 5, 7] Calculate the sample covariance matrix S and the sample correlation matrix R using matrix operations.	[4]	[CO2]	[3] Apply
	c) Suppose you have a dataset with two variables, A and B, and you want to calculate the sample mean of a new variable C, where C = 2A - 3B. If the sample values are as follows: A: [10, 15, 20] B: [5, 8, 12] Calculate the sample mean of C.	[4]	[CO2]	[3] Apply
Q3.	a) For the given function $f(x.y) = \begin{cases} 4xy & ;0 < x < 1,\ 0 < y < 1 \\ 0 & ;elsewhere \end{cases}$ Find the marginal density of x and y	[4]	[CO3]	[3] Apply
	b) For X be $N_3(\mu,\Sigma)$ with μ'=[-3 1 4] & $\sum \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ Find whether (X1, X2) and X3 are dependent variables or not.	[4]	[CO3]	[3] Apply