Total No. of Printed Pages: 3

PRN No.	

PAPER CODE	V313- 214-ESE

December 2023 (ENDSEM) EXAM

TY B.TECH (SEMESTER - I)

COURSE NAME: Multivariate Analysis **Branch: AI & DS** **COURSE ODE: ES31204AD**

(PATTERN 2020)

Time: [1Hr. 30 Min] **[Max. Marks: 40]**

(*) 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) **All questions are compulsory. Solve any one sub question from Question 3 and any two sub questions each from Questions 4,5 and 6 respectively.**

Q. No.	Question Description	Max. Marks	CO mapped	BT Level
Q.1	a) Explain the need for Multivariate Statistics in comparison to Univariate Statistics. Provide examples illustrating situations where Multivariate Statistical Modelling is more appropriate. Discuss the limitations of relying solely on Univariate approaches in complex data analysis.	[2]	CO 1	[I] Remember
Q2	a) Consider two variables, A and B, with the following data: $A = \begin{bmatrix} 3 \\ 7 \\ 2 \\ 8 \end{bmatrix}$ $B = \begin{bmatrix} 5 \\ 1 \\ 6 \\ 4 \end{bmatrix}$ Calculate sample mean and covariance matrix	[2]	CO 2	[V] Evaluate
Q3.	a) Explain what is meant by marginal distributions in the context of multivariate distributions. How are marginal distributions useful in understanding the behavior of individual variables within a multivariate setting? Provide an example. b) Define conditional distributions in the context of multivariate distributions. Discuss how conditional distributions can be used to analyze the relationship between variables. Provide a practical example.	[6] [6]	CO 3 CO 3	[I] Remember [II] Understand

Q.4	a) A die is weighted or loaded so that the number of spots X that appear on the up face when the die is rolled has pmf P(x) =x/21, where x=1,2,3,4,5,6. If this loaded die is rolled 21 times, find the probability of rolling 1,2,3,4,5,6.	[5]	CO 4	[IV] Analyze
	b) Suppose we want to know whether or not three different exam preparation programs lead to different mean scores on a certain exam. To test this, we recruit 30 students to participate in a study and split them into three groups. The students in each group are randomly assigned to use one of the three exam preparation programs for the next three weeks to prepare for an exam. At the end of the three weeks, all of the students take the same exam. The exam scores for each group are shown below:	[5]	CO 4	[V] Evaluate
	c) An online medicine shop claims that the mean delivery time for medicines is less than 120 mins with a standard deviation of 30 mins. Is there enough evidence to support this claim at a 0.05 significance level if 49 orders were examined with a mean of 100 mins.	[5]	CO 4	[V] Evaluate
Q.5	a) Imagine your bank conducts a phone survey for customer satisfaction and the results show the following factor loadings:	[5]	CO 5	[II] Understanding
	b) Given the data in the table, reduce the dimensions from 2 to 1 using PCA algorithm	[5]	CO 4	[II] Understanding

Group 1	Group 2	Group 3
85	91	79
86	92	78
88	93	88
75	85	94
78	87	92
94	84	85
98	82	83
79	88	85
71	95	82
80	96	81

Variable	Factor 1	Factor 2	Factor 3
Question 1	0.885	0.121	-0.033
Question 2	0.829	0.078	0.157
Question 3	0.777	0.190	0.540

Describe the relationships between the factors and the observed variables

Feature	Example 1	Example 2	Example 3	Example 4

	X1 | 4 | 8 | 13 | 7 X2 | 11 | 4 | 5 | 14 c) For the given data find the proximity of ordinal attributes Object ID — Attribute 1 — High 2 — Low 3 — Medium 4 — High	[5]	CO 4	[III] Applying
Q.6)	a) Interpret path diagrams in structural equation modeling. b) Using confirmatory factor analysis in SEM measure the latent variable intelligence on the basis of test score spread out in 4 area reading, writing, math and analysis. c) An unfair coin is flipped 100 times, and 61 heads are observed. The coin either has probability 1/3,1/2, or 2/3 of flipping a head each time it is flipped. Which of the three is the MLE?	[5] [5] [5]	CO 6 CO 6 CO 6	[II] Understanding [V] Evaluate [V] Evaluate