



VIT

Vellore Institute of Technology
(Declared to be University under section 3 of UGC Act, 1956)

SCHOOL OF ADVANCED SCIENCES
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2025-2026

REG.NO.:

SLOT: A1

Programme Name & Branch : B. Tech
Course Code and Course Name : CBS1009 & Computational Statistics
Faculty Name(s) : Dr. R. Padma & Dr. V. Kanimozhi
Class Number(s) : VL2025260104405, 4403
Date of Examination : 17/08/2025
Exam Duration : 90 minutes

Maximum Marks: 50

General instruction(s):

- Answer All Questions
- M - Max mark; CO - Course Outcome; BL - Blooms Taxonomy Level (1 - Remember, 2 - Understand, 3 - Apply, 4 - Analyse, 5 - Evaluate, 6 - Create)
- Course Outcomes (Type the CO statements covered in this question paper. Use the CO number as per the syllabus copy)

CO1: Analyse and interpret statistical data using multivariate normal distributions
CO2: Learn the approaches to point estimation of parameters.

Q. No	Question	M	CO	BL
1.	A climate researcher records annual measurement of average summer temperature ($^{\circ}\text{C}$) in a coastal region X and the average annual rainfall (mm) as Y . Data shows (X, Y) follows a bivariate normal population with $\mu_1 = 28, \mu_2 = 1100, \sigma_{11} = 4, \sigma_{22} = 10000$ and $\rho_{12} = -0.6$ (a) Write out the squared generalized distance expression (b) Determine and sketch the contour that contains 90% of the probability.	10	1	2
2.	Lenovo company produces three laptop models: Basic (X_1), Pro (X_2), and Ultra (X_3). The battery life in hours (X_1, X_2, X_3) follows a multivariate normal distribution with parameters: $\mu = \begin{pmatrix} 8 \\ 10 \\ 14 \end{pmatrix}$ and $\Sigma = \begin{pmatrix} 1.44 & 0.8 & 0.5 \\ 0.8 & 2.25 & 1.2 \\ 0.5 & 1.2 & 4 \end{pmatrix}$. Suppose we select a laptop model at random, what is the probability that (a) the battery life of basic model is greater than 9 hours? (b) the battery life of basic model is greater than 9 hours, given that the life of battery of Pro is 11 and Ultra is 13 hours respectively? (c) Are $X_1 - 2X_2$ and $X_2 + X_3$ independent?	10	1	3
3.	A carpenter measures the length (X_1), width (X_2) and height (X_3) of 4 wooden blocks (in cm). The data is assumed to follow a trivariate normal distribution. The measurements are given below, where each column corresponds to one block:	10	1	2



VIT

Vellore Institute of Technology
(Declared as a University under section 3 of UGC Act, 1956)

SCHOOL OF ADVANCED SCIENCES
CONTINUOUS ASSESSMENT TEST - I
FALL SEMESTER 2025-2026

REG.NO.:

SLOT: A1

	$X = \begin{bmatrix} 5 & 7 & 6 & 8 \\ 3 & 5 & 4 & 6 \\ 2 & 4 & 3 & 5 \end{bmatrix}$																																							
	a). Find the maximum likelihood estimate of the mean vector. b). Find the maximum likelihood estimate of the covariance matrix. c). Based on your result, comment on whether longer blocks tend to be taller.																																							
4.	A property consultant wants to predict the monthly rent (Y in thousands) based on the area (X_1 in sq. m.) and the distance from city centre (X_2 in km). <table><tr><th>Apartment</th><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th></tr><tr><td>X_1</td><td>50</td><td>55</td><td>60</td><td>65</td><td>70</td><td>75</td><td>80</td><td>85</td></tr><tr><td>X_2</td><td>12</td><td>10</td><td>8</td><td>6</td><td>5</td><td>4</td><td>3</td><td>2</td></tr><tr><td>Y</td><td>15</td><td>18</td><td>20</td><td>22</td><td>25</td><td>27</td><td>28</td><td>30</td></tr></table> a). Using the above data, fit a multiple linear regression model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon$. b). Predict the rent for $X_1 = 72$ and $X_2 = 4$. Also, Identify which factor affects the rent more?	Apartment	1	2	3	4	5	6	7	8	X_1	50	55	60	65	70	75	80	85	X_2	12	10	8	6	5	4	3	2	Y	15	18	20	22	25	27	28	30	10	2	2
Apartment	1	2	3	4	5	6	7	8																																
X_1	50	55	60	65	70	75	80	85																																
X_2	12	10	8	6	5	4	3	2																																
Y	15	18	20	22	25	27	28	30																																
5.	A company record shows the details of sales (Y in thousands) based on two predictors: advertising spend (X_1 in lakhs) and the number of sales persons (X_2) <table><tr><td>Y</td><td>10</td><td>12</td><td>14</td><td>16</td><td>18</td><td>45</td></tr><tr><td>X_1</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>X_2</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td></tr></table> a). Check for collinearity between the advertising expenditure and number of salesmen. b). If the collinearity is perfect then fit a simple regression model using X_1 as the predictor.	Y	10	12	14	16	18	45	X_1	4	5	6	7	8	9	X_2	5	6	7	8	9	10	10	2	3															
Y	10	12	14	16	18	45																																		
X_1	4	5	6	7	8	9																																		
X_2	5	6	7	8	9	10																																		
