



VIT

Vellore Institute of Technology
(Deemed to be University under Section 3 of UGC Act, 1956)

REG.NO.:

SCHOOL OF ADVANCED SCIENCES
CONTINUOUS ASSESSMENT TEST - II
FALL SEMESTER 2024-2025

SLOT:B1

Programme Name & Branch :B.Tech CSBS
Course Code and Course Name :CBS 1009 Computational Statistics
Faculty Name(s) :Dr.B.Sripathy, Dr.V.Sujatha
Class Number(s) : VL2024250105573, 74
Date of Examination : 14.10.2024
Exam Duration : 90 minutes **Maximum Marks: 50**
General instruction(s): Answer All Questions

Q. No	Question	M																																								
1.	Let the response variables are the <u>sales</u> and <u>profit</u> of a factory and the predictor variables are the absenteeism of workers and the breakdown hours of machine. Construct the multivariate multiple linear regression model from the sample data given below. Calculate the residual. <table><tr><td>Sales</td><td>10</td><td>12</td><td>11</td></tr><tr><td>Profit(lakhs)</td><td>100</td><td>110</td><td>105</td></tr><tr><td>Absenteeism</td><td>9</td><td>8</td><td>7</td></tr><tr><td>Breakdown hours</td><td>62</td><td>58</td><td>64</td></tr></table>	Sales	10	12	11	Profit(lakhs)	100	110	105	Absenteeism	9	8	7	Breakdown hours	62	58	64	10																								
Sales	10	12	11																																							
Profit(lakhs)	100	110	105																																							
Absenteeism	9	8	7																																							
Breakdown hours	62	58	64																																							
2.	The following data represent the weight in(Kgs) of 20 students as follows: 56, 65, 77, 76, 87, 89, 95, 98, 78, 67, 88, 99, 74, 61, 59, 63, 45, 78, 79, 81. Calculate the autocorrelation lagged by 2.	10																																								
3.	The random variables X, Y, and Z have the following variance and covariance matrix $\Sigma = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 5 \end{bmatrix}$ Find the principal components Y_i, $i = 1, 2, \text{ and } 3$	10																																								
4.	Nine viewers rated two films they saw on three dimensions:1. Entertainment value 2. Writing quality 3. Action. The ratings are such that a score of 5 means they felt the film was exactly average of that dimension, with higher scores denoting higher quality and lower scores denoting lower quality. Using MANOVA, analyse the data and comment it. <table><tr><td>No.</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr><tr><td>Dimension</td><td>1</td><td>1</td><td>1</td><td>2</td><td>2</td><td>2</td><td>3</td><td>3</td><td>3</td></tr><tr><td>Film 1 Rating</td><td>4</td><td>2</td><td>1</td><td>5</td><td>6</td><td>5</td><td>6</td><td>8</td><td>8</td></tr><tr><td>Film 2 Rating</td><td>1</td><td>4</td><td>3</td><td>4</td><td>5</td><td>4</td><td>8</td><td>7</td><td>8</td></tr></table>	No.	1	2	3	4	5	6	7	8	9	Dimension	1	1	1	2	2	2	3	3	3	Film 1 Rating	4	2	1	5	6	5	6	8	8	Film 2 Rating	1	4	3	4	5	4	8	7	8	10
No.	1	2	3	4	5	6	7	8	9																																	
Dimension	1	1	1	2	2	2	3	3	3																																	
Film 1 Rating	4	2	1	5	6	5	6	8	8																																	
Film 2 Rating	1	4	3	4	5	4	8	7	8																																	
5.	Apply K -means clustering for the following data set: (2,4,10,12, 3, 20, 30, 11, 25) with a) 2 clusters b) with 3 clusters.	10																																								