

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -2 EXAMINATION- 2025

B.Sc.-V Semester (M&C)

COURSE CODE (CREDITS): 24BS1MA511 (3)

MAX. MARKS: 25

COURSE NAME: Statistical Methods for Data Analysis

COURSE INSTRUCTORS: BKP

MAX. TIME: 1 Hour 30 Min

Note: (a) All questions are compulsory.

(b) The candidate is allowed to make suitable numeric assumptions wherever required for solving problems.

(c) Non-scientific calculators are allowed and statistical tables are given at the end of paper.

Q. No.	Question	CO	Marks																																	
Q1.	A group of 10 B. Sc. Computer Science students participated in a coding efficiency training program aimed at improving their programming speed. The time taken (in minutes) by each student to complete a standard coding task before and after the training is recorded as follows: <table><tr><th>Student</th><th>Before Training (X_1)</th><th>After Training (X_2)</th></tr><tr><td>1</td><td>52</td><td>48</td></tr><tr><td>2</td><td>47</td><td>44</td></tr><tr><td>3</td><td>55</td><td>50</td></tr><tr><td>4</td><td>49</td><td>46</td></tr><tr><td>5</td><td>58</td><td>54</td></tr><tr><td>6</td><td>53</td><td>50</td></tr><tr><td>7</td><td>51</td><td>47</td></tr><tr><td>8</td><td>50</td><td>46</td></tr><tr><td>9</td><td>54</td><td>50</td></tr><tr><td>10</td><td>56</td><td>52</td></tr></table> Use a suitable test at a 5% significance level to determine whether the training	Student	Before Training (X_1)	After Training (X_2)	1	52	48	2	47	44	3	55	50	4	49	46	5	58	54	6	53	50	7	51	47	8	50	46	9	54	50	10	56	52	CO-2	6
Student	Before Training (X_1)	After Training (X_2)																																		
1	52	48																																		
2	47	44																																		
3	55	50																																		
4	49	46																																		
5	58	54																																		
6	53	50																																		
7	51	47																																		
8	50	46																																		
9	54	50																																		
10	56	52																																		

	program significantly improved the students' coding performance.																													
Q2.	A study was conducted among Computer Science students to examine whether the choice of programming language depends on the year of study. The data collected from 90 students is summarized below: <table><tr><td>Year of Study</td><td>Python</td><td>C++</td><td>Java</td><td>Total</td></tr><tr><td>2nd Year</td><td>20</td><td>10</td><td>15</td><td>45</td></tr><tr><td>4th Year</td><td>10</td><td>15</td><td>20</td><td>45</td></tr><tr><td>Total</td><td>30</td><td>25</td><td>35</td><td>90</td></tr></table> Use a Chi-Square test for independence at a 5% level of significance to determine whether the preference for programming language is associated with the year of study.	Year of Study	Python	C++	Java	Total	2nd Year	20	10	15	45	4th Year	10	15	20	45	Total	30	25	35	90	CO-2	6							
Year of Study	Python	C++	Java	Total																										
2nd Year	20	10	15	45																										
4th Year	10	15	20	45																										
Total	30	25	35	90																										
Q3.	The following data show the engine load (X, in kg) and the corresponding fuel consumption rate (Y, in liters/hour) for a test conducted on a diesel engine in a Mechanical Engineering laboratory: <table><tr><td>Observation</td><td>Engine Load (X)</td><td>Fuel Consumption (Y)</td></tr><tr><td>1</td><td>10</td><td>2.5</td></tr><tr><td>2</td><td>20</td><td>3.0</td></tr><tr><td>3</td><td>30</td><td>3.6</td></tr><tr><td>4</td><td>40</td><td>4.1</td></tr><tr><td>5</td><td>50</td><td>4.7</td></tr><tr><td>6</td><td>60</td><td>5.0</td></tr><tr><td>7</td><td>70</td><td>5.6</td></tr><tr><td>8</td><td>80</td><td>6.2</td></tr></table> Use Karl Pearson's Coefficient of Correlation (r) to determine whether there is a linear relationship between engine load and fuel consumption rate.	Observation	Engine Load (X)	Fuel Consumption (Y)	1	10	2.5	2	20	3.0	3	30	3.6	4	40	4.1	5	50	4.7	6	60	5.0	7	70	5.6	8	80	6.2	CO-3	6
Observation	Engine Load (X)	Fuel Consumption (Y)																												
1	10	2.5																												
2	20	3.0																												
3	30	3.6																												
4	40	4.1																												
5	50	4.7																												
6	60	5.0																												
7	70	5.6																												
8	80	6.2																												
Q4.	The following two equations represent the lines of regression between variables X and Y: $8X - 10Y + 66 = 0$ $40X - 18Y = 214$ You are required to: (a). Calculate the means of X and Y. (b). Identify which equation represents Y on X and which represents X on Y. (c). Find the regression equations of Y on X and X on Y in their standard forms.	CO-3	7																											

t-Distribution Table (two-tailed, $\alpha = 0.05$)		Chi-Square (χ^2) Distribution Table ($\alpha = 0.05$)	
Degrees of Freedom (df)	Critical t-value ($\pm t_{0.025}$)	Degrees of Freedom (df)	χ^2 Critical Value (0.05)
1	12.706	1	3.841
2	4.303	2	5.991
3	3.182	3	7.815
4	2.776	4	9.488
5	2.571	5	11.070
6	2.447	6	12.592
7	2.365	7	14.067
8	2.306	8	15.507
9	2.262	9	16.919
10	2.228	10	18.307
15	2.131	12	21.026
20	2.086	15	24.996
25	2.060	20	31.410
30	2.042		
40	2.021		
60	2.000		
∞ (large sample, Z)	1.960		