

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY, WAKNAGHAT

TEST -3 EXAMINATION- 2025

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

COURSE CODE (CREDITS): 24BS1MA511 (3)

MAX. MARKS: 35

COURSE NAME: Statistical Methods for Data Analysis

COURSE INSTRUCTORS: BKP

MAX. TIME: 2 Hours

Note: (a) All questions are compulsory.

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

(c) Calculator is allowed and statistical tables are given at the end of paper.

Q. No.	Question	CO	Marks																																
Q1.	<p>Fifty students took up a test. The result of those who passed the test is given below:</p> <table border="1"> <tr> <td>Marks</td> <td>4</td> <td>5</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> </tr> <tr> <td>No. of Students</td> <td>8</td> <td>10</td> <td>9</td> <td>6</td> <td>4</td> <td>3</td> </tr> </table> <p>If the average of all the 50 students was 5.16 marks, Find the average of those who failed.</p>	Marks	4	5	6	7	8	9	No. of Students	8	10	9	6	4	3	CO-1	7																		
Marks	4	5	6	7	8	9																													
No. of Students	8	10	9	6	4	3																													
Q2.	<p>A small manufacturing unit wants to estimate weekly production output (Y) using two factors:</p> <table> <thead> <tr> <th>Variable</th> <th>Meaning</th> </tr> </thead> <tbody> <tr> <td>X_1 = Number of workers</td> <td>Continuous variable</td> </tr> <tr> <td>(X_2) = Machine hours used</td> <td>Continuous variable</td> </tr> <tr> <td>(Y) = Units produced per week</td> <td>Target variable</td> </tr> </tbody> </table> <p>The collected data for 5 weeks is:</p> <table> <thead> <tr> <th>Week</th> <th>Workers (X_1)</th> <th>Machine Hours (X_2)</th> <th>Output(Y)</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>4</td> <td>40</td> <td>56</td> </tr> <tr> <td>2</td> <td>6</td> <td>45</td> <td>68</td> </tr> <tr> <td>3</td> <td>5</td> <td>50</td> <td>65</td> </tr> <tr> <td>4</td> <td>7</td> <td>55</td> <td>80</td> </tr> <tr> <td>5</td> <td>6</td> <td>52</td> <td>75</td> </tr> </tbody> </table> <p>Fit a multiple linear regression model:</p> $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2$	Variable	Meaning	X_1 = Number of workers	Continuous variable	(X_2) = Machine hours used	Continuous variable	(Y) = Units produced per week	Target variable	Week	Workers (X_1)	Machine Hours (X_2)	Output(Y)	1	4	40	56	2	6	45	68	3	5	50	65	4	7	55	80	5	6	52	75	CO-4	7
Variable	Meaning																																		
X_1 = Number of workers	Continuous variable																																		
(X_2) = Machine hours used	Continuous variable																																		
(Y) = Units produced per week	Target variable																																		
Week	Workers (X_1)	Machine Hours (X_2)	Output(Y)																																
1	4	40	56																																
2	6	45	68																																
3	5	50	65																																
4	7	55	80																																
5	6	52	75																																

Q3.	<p>A survey was conducted among engineering students to determine whether the mode of study preparation is associated with their exam performance level. The responses collected from 120 students are displayed below:</p> <table border="1" data-bbox="311 422 1218 590"> <thead> <tr> <th>Exam Performance</th><th>Online Notes</th><th>Classroom Notes</th><th>Self-Study</th></tr> </thead> <tbody> <tr> <td>High</td><td>25</td><td>20</td><td>15</td></tr> <tr> <td>Low</td><td>15</td><td>25</td><td>20</td></tr> </tbody> </table> <p>Using the Chi-Square Test for Independence at a 5% level of significance, determine whether the exam performance of students is associated with their mode of preparation.</p>	Exam Performance	Online Notes	Classroom Notes	Self-Study	High	25	20	15	Low	15	25	20	CO-5	7												
Exam Performance	Online Notes	Classroom Notes	Self-Study																								
High	25	20	15																								
Low	15	25	20																								
Q4.	<p>A company wants to compare the mean productivity of workers trained using three different training programs. Five workers were randomly selected from each training method, and their daily output (units produced) was recorded.</p> <table data-bbox="498 923 1028 1185"> <thead> <tr> <th></th> <th>Training A</th> <th>Training B</th> <th>Training C</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>48</td> <td>52</td> <td>50</td> </tr> <tr> <td>2</td> <td>45</td> <td>55</td> <td>53</td> </tr> <tr> <td>3</td> <td>47</td> <td>54</td> <td>52</td> </tr> <tr> <td>4</td> <td>46</td> <td>53</td> <td>51</td> </tr> <tr> <td>5</td> <td>50</td> <td>56</td> <td>54</td> </tr> </tbody> </table> <p>Using One-Way ANOVA at 5% significance, test whether all three training programs have the same average productivity, or if at least one program is significantly better.</p>		Training A	Training B	Training C	1	48	52	50	2	45	55	53	3	47	54	52	4	46	53	51	5	50	56	54	CO-5	7
	Training A	Training B	Training C																								
1	48	52	50																								
2	45	55	53																								
3	47	54	52																								
4	46	53	51																								
5	50	56	54																								
Q.5	<p>A manufacturing company produces ball bearings in two different machines, Machine A and Machine B. To check whether both machines have equal variability in diameter, the following sample data (in mm) were collected:</p> <table border="1" data-bbox="472 1507 1049 1619"> <tbody> <tr> <td>Machine A</td> <td>10.2</td> <td>10.4</td> <td>10.1</td> <td>10.3</td> <td>10.5</td> </tr> <tr> <td>Machine B</td> <td>10.5</td> <td>10.7</td> <td>10.6</td> <td>10.8</td> <td>10.9</td> </tr> </tbody> </table> <p>Assume the populations are normally distributed.</p> <p>(a) State the null and alternative hypothesis.</p> <p>(b) At 5% level of significance, determine whether the variability in diameters produced by both machines is the same.</p>	Machine A	10.2	10.4	10.1	10.3	10.5	Machine B	10.5	10.7	10.6	10.8	10.9	CO-5	7												
Machine A	10.2	10.4	10.1	10.3	10.5																						
Machine B	10.5	10.7	10.6	10.8	10.9																						

F Table for $\alpha = 0.05$

<i>f</i>	$df_1=1$	2	3	4	5	6	7	8	9
df₂=1	161.4476	199.5000	215.7073	224.5832	230.1619	233.9860	236.7684	238.8827	240.5433
2	18.5128	19.0000	19.1643	19.2468	19.2964	19.3295	19.3532	19.3710	19.3848
3	10.1280	9.5521	9.2766	9.1172	9.0135	8.9406	8.8867	8.8452	8.8123
4	7.7086	6.9443	6.5914	6.3882	6.2561	6.1631	6.0942	6.0410	5.9988
5	6.6079	5.7861	5.4095	5.1922	5.0503	4.9503	4.8759	4.8183	4.7725
6	5.9874	5.1433	4.7571	4.5337	4.3874	4.2839	4.2067	4.1468	4.0990
7	5.5914	4.7374	4.3468	4.1203	3.9715	3.8660	3.7870	3.7257	3.6767
8	5.3177	4.4590	4.0662	3.8379	3.6875	3.5806	3.5005	3.4381	3.3881
9	5.1174	4.2565	3.8625	3.6331	3.4817	3.3738	3.2927	3.2296	3.1789
10	4.9646	4.1028	3.7083	3.4780	3.3258	3.2172	3.1355	3.0717	3.0204
11	4.8443	3.9823	3.5874	3.3567	3.2039	3.0946	3.0123	2.9480	2.8962
12	4.7472	3.8853	3.4903	3.2592	3.1059	2.9961	2.9134	2.8486	2.7964
13	4.6672	3.8056	3.4105	3.1791	3.0254	2.9153	2.8321	2.7669	2.7144
14	4.6001	3.7389	3.3439	3.1122	2.9582	2.8477	2.7642	2.6987	2.6458
15	4.5431	3.6823	3.2874	3.0556	2.9013	2.7905	2.7066	2.6408	2.5876

Chi-Square (χ^2) Distribution Table ($\alpha = 0.05$)

Degrees of Freedom (df)	χ^2 Critical Value (0.05)
1	3.841
2	5.991
3	7.815
4	9.488
5	11.070
6	12.592
7	14.067
8	15.507
9	16.919
10	18.307
12	21.026
15	24.996
20	31.410