

INDIRA UNIVERSITY, PUNE

SCHOOL OF INFORMATION TECHNOLOGY-BSC (DS)

Term End Examination (2025 Pattern) December – 2025 - Semester – I

Subject Name: Descriptive Statistics
Subject Code: 25BDS102T

Max. Marks: 25
Time: 1:30 Hrs.

Instructions

- All Questions are Compulsory.
- Use of Non programmable scientific calculator is allowed.

CO #	Cognitive Ability	Course Outcome
CO1	Remember	Summarize data visually and numerically.
CO2	Understand	Analyse a problem, identify methods in descriptive statistics.
CO3	Apply	Define the computing requirements appropriate to its solution.
CO4	Analyze	Identify correlation between variables.
CO5	Evaluate	Fit an equation for prediction and apply the same for real data and assess data-based models.
CO6	Create	Execute statistical analysis with any software tool.

Q1.	<p style="text-align: center;">Attempt any 5 out of 7. (1 mark each) (5 Marks)</p> <p>a. The number of times a value occurs in a data set is called: i) Frequency ii) Class limit iii) Variable iv) Attribute</p> <p>b. The coefficient of variation (C.V.) is used to: i) compare averages . ii) compare consistency or stability between the data sets . iii) find the highest frequency . iv) calculate correlation.</p> <p>c. The measure of skewness based on moments uses: i) Central moments ii) Raw moments iii) Quartiles iv) Range</p> <p>d. Spearman’s rank correlation is used when: i) data is qualitative or ranked ii) data is quantitative iii) data is grouped iv) data has no relationship</p> <p>e. The model $Y = a + bX + cX^2$ is known as i) Linear regression ii) Polynomial regression iii) Logistic regression iv) Poisson regression</p> <p>f. Data collected by the researcher for the first time is called i) secondary data ii) internal data iii) primary data iv) derived data</p> <p>g. The fourth central moment is used for measuring i) Skewness ii) Kurtosis iii) Dispersion iv) Mean</p>	CO1
------------	---	-----

<p>Q2.</p>	<p>Attempt any 2 out of 4. (5 marks each) (10 Marks)</p> <p>a. The following table shows the marks obtained by 40 students in a statistics test. Construct a histogram to represent the data.</p> <table border="1" data-bbox="440 289 1170 363"> <tr> <td>Marks</td> <td>0-10</td> <td>10-20</td> <td>20-30</td> <td>30-40</td> <td>40-50</td> </tr> <tr> <td>Number of students</td> <td>5</td> <td>8</td> <td>12</td> <td>10</td> <td>5</td> </tr> </table> <p>b. Determine the standard deviation for the following grouped data:</p> <table border="1" data-bbox="440 432 1092 506"> <tr> <td>class interval</td> <td>0-10</td> <td>10-20</td> <td>20-30</td> <td>30-40</td> <td>40-50</td> </tr> <tr> <td>frequency</td> <td>5</td> <td>7</td> <td>8</td> <td>5</td> <td>5</td> </tr> </table> <p>c. Calculate Bowley's coefficient of skewness for the following data and interpret it: 2, 4, 6, 8, 10, 9, 7.</p> <p>d. From the following information, find regression equation of Y on X and estimate the production when the capacity utilization is 70%.</p> <table border="1" data-bbox="440 779 1276 921"> <tr> <td></td> <td>Average (Mean)</td> <td>S.D.</td> </tr> <tr> <td>Capacity Utilization (%) (X)</td> <td>88</td> <td>8.5</td> </tr> <tr> <td>Production (in lakh units) (Y)</td> <td>42</td> <td>12.5</td> </tr> </table> <p>Correlation Coefficient (r) = 0.72</p>	Marks	0-10	10-20	20-30	30-40	40-50	Number of students	5	8	12	10	5	class interval	0-10	10-20	20-30	30-40	40-50	frequency	5	7	8	5	5		Average (Mean)	S.D.	Capacity Utilization (%) (X)	88	8.5	Production (in lakh units) (Y)	42	12.5	<p>CO2</p>
Marks	0-10	10-20	20-30	30-40	40-50																														
Number of students	5	8	12	10	5																														
class interval	0-10	10-20	20-30	30-40	40-50																														
frequency	5	7	8	5	5																														
	Average (Mean)	S.D.																																	
Capacity Utilization (%) (X)	88	8.5																																	
Production (in lakh units) (Y)	42	12.5																																	
<p>Q3.</p>	<p>(A) Attempt all questions. (10 Marks)</p> <p>a. Explain the concept of Central Tendency. Which are the different measures of central tendency. Define Median.</p> <p>b. Fit a second-degree polynomial curve $Y=a+bX+cX^2$ to:</p> <table border="1" data-bbox="415 1150 889 1232"> <tr> <td>X</td> <td>5</td> <td>6</td> <td>7</td> <td>8</td> <td>9</td> </tr> <tr> <td>Y</td> <td>30</td> <td>48</td> <td>70</td> <td>96</td> <td>126</td> </tr> </table> <p style="text-align: center;">OR</p> <p>(B) Attempt the following question. (10 Marks)</p> <p>Calculate Karl Pearson's coefficient of correlation (r) for the following data and interpret it:</p> <table border="1" data-bbox="371 1470 829 1551"> <tr> <td>X</td> <td>10</td> <td>20</td> <td>30</td> <td>40</td> <td>50</td> </tr> <tr> <td>Y</td> <td>30</td> <td>40</td> <td>50</td> <td>60</td> <td>70</td> </tr> </table>	X	5	6	7	8	9	Y	30	48	70	96	126	X	10	20	30	40	50	Y	30	40	50	60	70	<p>CO3</p>									
X	5	6	7	8	9																														
Y	30	48	70	96	126																														
X	10	20	30	40	50																														
Y	30	40	50	60	70																														
