SET-3

INDIRA UNIVERSITY, PUNE

SCHOOL OF INFORMATION TECHNOLOGY-BSC (AIML)

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

Subject Name: - Notion of Statistical Data Analysis **Max. Marks: 25**
Subject Code: 25AML102T **Time: 1:30 Hrs.**

Instructions

- All Questions are Compulsory.
- Use of Scientific calculator is allowed.

CO #	Cognitive Ability	Course Outcome
CO1	**Remember**	Recall and define fundamental statistical concepts.
CO2	**Understand**	Organize and manage data effectively for analysis.
CO3	**Apply**	Use statistical thinking to explore and interpret data.

Q1.	**Attempt any 5 out of 7. (1 mark each) (5 Marks)** a) What is the last less than cumulative frequency equal to? b) Define mode. c) Find the first central moment for data: 5, 6, 7, 8, 9. d) Name the three types of skewness. e) What does coefficient of correlation r = +1 indicate? f) If regression line is Y = 10 + 2X, find Y when X = 3. g) If variance = 9, find standard deviation.	CO1
Q2.	**Attempt any 2 out of 4. (5 marks each) (10 Marks)** a) Draw histogram for the following data and find mode using it. b) The ages (in years) of 8 employees are: 25, 26, 27, 29, 30, 31, 33, 35. Compute the variance and standard deviation. c) Explain the concept of r^{th} ordered raw and central moments. State the mathematical formula of r^{th} raw moment and r^{th} central moment for ungrouped data. d) Explain the concept of skewness and kurtosis with the help of suitable diagrams.	CO2

Class	0-10	10-20	20-30	30-40	40-50	50-60
Frequency	5	12	24	30	15	8

Q.3.	**Attempt all questions:** **(10 Marks)** a) The following information is available: $\Sigma X = 40$, $\Sigma Y = 50$, $\Sigma XY = 450$, $\Sigma X^2 = 200$, $\Sigma Y^2 = 300$, $n = 5$. Find the Karl Pearson's coefficient of correlation and comment on the nature of correlation. b) Explain how to apply the concept of coefficient of determination (r^2) in understanding the proportion of explained variation with illustration. **OR**	CO3
	The distribution of ages (in years) of 50 employees is given below: Apply Bowley's formula for coefficient of skewness based on quartiles and interpret whether the distribution is symmetric or skewed. **Age** | 20-25 | 25-30 | 30-35 | 35-40 | 40-45 | 45-50 **Frequency** | 5 | 10 | 20 | 8 | 5 | 2	CO3

*****