Course Code		**23STMAP235**	
Title of the Course		**Descriptive Statistics and Probability Lab**	
Offered to: (Programme/s)		B.Sc. Honours -CSC (A, B, & C)	
Year of Introduction:	**2024-25**	**Semester:**	**3**
Course Category:	**MAJOR**	**Course Relates to:**	**Local, Regional, National, Global**
Year of Revision:	**2024 - 25**	**Percentage:**	**NA**
Type of the Course:		**Skill Development Course**	
Crosscutting Issues of the Course:		**NA**	
Pre-requisites, if any		**Basic Mathematics**	

Course Description:

The course on Descriptive Statistics, Probability, Random Variables, and Probability Distributions provides a comprehensive introduction to the fundamental concepts of statistics and probability theory. Students will learn to summarize and interpret data using descriptive statistics, including measures of central tendency and dispersion. The course covers key probability concepts, including random variables and their distributions, allowing students to understand the behavior of data in uncertain environments. Through practical applications, students will explore various discrete and continuous probability distributions, enabling them to model real-world scenarios and make informed decisions based on statistical analysis. Emphasis will be placed on problem-solving and the use of statistical software for data analysis.

Course Aims and Objectives:

S. No	COURSE OBJECTIVES
1	compute and interpret measures of central tendency (mean, median, mode) for various datasets, highlighting their significance in summarizing data.
2	calculating and analyzing measures of dispersion (range, variance, standard deviation) and to explore the concepts of moments, skewness, and kurtosis in understanding data variability.
3	define and illustrate discrete and continuous random variables, and to analyze their corresponding probability mass functions (PMF) and probability density functions (PDF).
4	calculate mathematical expectations and variances, and to apply generating functions (moment generating, probability generating) in solving probability-related problems.
5	identify and analyze both discrete (e.g., Binomial, Poisson) and continuous (e.g., Normal, Exponential) probability distributions, and to apply these distributions in real-world statistical modeling scenarios.

Course Outcomes

At the end of the course, the student will be able to…

NO	COURSE OUTCOME	BTL	PO	PSO
CO1	apply various measures of central tendency (mean, median, mode) using real datasets, and analyze how these measures reflect data characteristics.	**K3**	**2**	**2**
CO2	apply measures of dispersion (range, variance, standard deviation) to datasets, evaluate the impact of skewness and kurtosis on data distribution, and interpret moments to summarize data variability."	**K3**	**2**	**2**
CO3	illustrate the probability mass functions and density functions and analyze the properties of distribution functions through practical examples."	**K4**	**2**	**2**
CO4	compute mathematical expectations and variances using generating functions, apply the properties of moment generating functions, and analyze their relevance in probability theory."	**K3**	**2**	**2**
CO5	analyze probability distributions and apply them to solve real-world problems.	**K4**	**2**	**2**

For BTL: K1: Remember; K2: Understand; K3: Apply; K4: Analyze; K5: Evaluate; K6: Create

CO-PO-PSO MATRIX									
CO NO	**PO1**	**PO2**	**PO3**	**PO4**	**PO5**	**PO6**	**PO7**	**PSO1**	**PSO2**
CO1		3							3
CO2		3							3
CO3		3							3
CO4		3							3
CO5		3							3

Use the codes 3, 2, 1 for High, Moderate and Low correlation Between CO-PO-PSO respectively

Lab Course Structure

Lab Component 1: Computing Measures of Central Tendency (CO1-K3)

Objective: Compute and compare various mathematical and positional averages using different datasets.

Activities:

1. Calculate arithmetic mean, geometric mean, and harmonic mean for a given dataset. Discuss their properties and applications in different scenarios.
2. Compute the median and mode of the dataset and analyse how they differ from the mean in skewed data distributions.
3. Visualize the results by plotting histograms and marking the mean, median, and mode.
4. Discuss the implications of using different measures of central tendency, especially in skewed and multimodal datasets.
5. Software: R or Python for calculating the averages and generating plots.

Lab Component 2: Exploring Measures of Dispersion, Moments, Skewness, and Kurtosis**(CO2-K3)**

Objective: Understand how dispersion, skewness, and kurtosis describe data variability and shape.
Activities:
1. Compute the range, quartile deviation, mean deviation, and standard deviation for a dataset.
2. Calculate absolute and relative measures, including the coefficient of variation, and interpret the results.
3. Compute central and non-central moments and explore their interrelationships (statements only). Apply Sheppard's corrections to grouped data for central moments.
4. Calculate and interpret measures of skewness (e.g., Pearson's coefficient) and kurtosis for the dataset. Identify whether the dataset exhibits positive or negative skewness and if it is platykurtic, mesokurtic, or leptokurtic.
5. Visualize the dataset's distribution along with its skewness and kurtosis characteristics.
6. Software: R or Python for statistical computations and visualizations.

Lab Component 3: Exploring Random Variables and Distribution Functions (CO3-K4)
Objective: Understand the difference between discrete and continuous random variables, and compute probability mass functions (PMF) and probability density functions (PDF).
Activities:
1. Define a discrete random variable and a continuous random variable
2. For the discrete case, compute the PMF and verify that the probabilities sum to 1.
3. For the continuous case, compute the PDF and verify that the area under the curve is 1.
4. Graph both PMF and PDF for visualization.
5. Compute and graph the cumulative distribution function (CDF) for both types of random variables.

Lab Component 4: Mathematical Expectations and Generating Functions (CO4-K3)
Objective: Calculate expectations, variances, and covariance; explore generating functions and their properties.
Activities:
1. Calculate the mathematical expectations for given discrete and continuous random variables using PMF and PDF.
2. Use the addition and multiplication theorems of expectations to verify properties.
3. Compute variance and covariance for pairs of random variables.
4. **Software:** R or Python, with built-in statistical functions for generating random variables, computing expectations, and visualizing distributions.

Lab Component 5: Exploring Discrete Probability Distributions (CO5-K4)

Objective: Understand and apply the Binomial, Poisson, and Geometric distributions to solve real-world problems.

Activities:

(a)Binomial Distribution:

1.Define the Binomial distribution and its properties (e.g., probability mass function, mean, variance).

2.Fit a binomial distribution and calculate expected frequencies using direct and recurrence relation methods.

(b)Poisson Distribution:

1.Define the Poisson distribution and its properties (e.g., mean and variance are equal).

2.Fit a Poisson distribution and calculate expected frequencies using direct and recurrence relation methods.

(c)Geometric Distribution:

1.Define the Geometric distribution and discuss its properties.

2.Fit a geometric distribution and calculate expected frequencies using direct and recurrence relation methods.

Software: R or Python for generating random variables, calculating probabilities, and plotting PMFs.

Lab Component 6: Exploring Continuous Probability Distributions (CO5-K4)

Objective: Understand and apply the Uniform, Normal, and Exponential distributions to solve real-world problems.

Activities:

(a)Normal Distribution:

1.Define the Normal distribution and its properties (e.g., bell-shaped curve, mean, and standard deviation).

2. Apply it to real-world data (e.g., heights, test scores) and compute probabilities for ranges using the cumulative distribution function (CDF).

(b)Exponential Distribution:

1.Define the Exponential distribution and discuss its properties (e.g., mean and variance).

2. Fit a exponential distribution and calculate expected frequencies.

Software: R or Python for computing probabilities and visualizing the PDFs and CDFs.

Reference Book

1. Lab Manual: Descriptive Statistics and Probability , prepared by the Department of Statistics, PBSCAS, Vijayawada.

2. "Statistics with R: Solving Problems Using Real-World Data" by Sarah Stowell

Link: Statistics with R

3. "Introduction to Probability and Statistics Using Python" by Gönenç Sorguç & Thomas Haslwanter

https://www.springer.com/gp/book/9783030648908

23STMAP235: Descriptive Statistics and Probability Lab

Offered to B.Sc. Hons Computer Science A/B/C

Max. Marks: 50 **Max. Time: 3Hrs**

Pass. Min: 20

(A)	**Evaluation Procedure**	**35 Marks**
I	Experiments (Exam & Execution)	**30 Marks**
II	Viva	**3 Marks**
III	Record	**2 Marks**
(B)	**CONTINUOUS ASSESMENT(Internal)**	**15 MARKS**

15 marks for the continuous assessment (Day to day work in the laboratory shall be evaluated for 15 marks by the concerned laboratory teacher based on the regularity/ record/viva). Laboratory teachers are mandated to ensure that every student completes 80%-90% of the lab assessments.

TOTAL: (A)+(B) = 50 MARKS