PARVATHANENI BRAHMAYYA
SIDDHARTHA COLLEGE OF ARTS & SCIENCE
Autonomous
Siddhartha Nagar, Vijayawada–520010
Re-accredited at 'A+' by the NAAC

Course Code				**23STMIL233**			
Title of the Course				**Random Variables and Probability Distributions**			
Offered to:				**B.B.A (Honours) Business Analytics**			
L	**4**	**T**	**0**	**P**	**0**	**C**	**3**
Year of Introduction:		**2024-25**		**Semester:**		**3**	
Course Category:		**Minor**		**Course Relates to:**		**Local, Regional, National**	
Year of Revision:		**NA**		**Percentage:**		**NA**	
Type of the Course:				**SKILL DEVELOMENT**			
Crosscutting Issues of the Course :				**NA**			
Pre-requisites, if any				**23STMIL123**			

Course Description:

This course provides an in-depth understanding of probability theory and its applications in statistical analysis. Students will explore the concepts of discrete and continuous random variables, probability distributions, and moments, along with advanced topics such as generating functions and expectation theorems. The course also covers important discrete and continuous probability distributions, including the Binomial, Poisson, Normal, Exponential, and Uniform distributions, as well as exact sampling distributions like the Chi-square, t, and F distributions. Through practical examples and applications, students will develop a strong foundation for conducting hypothesis testing and statistical modeling.

Course Objectives:

S. No	**COURSE OBJECTIVES**
1	Understand the fundamental concepts of discrete and continuous random variables and their associated probability functions.
2	Explore mathematical expectations, moments, and generating functions for random variables and apply these in solving problems.
3	Study important discrete and continuous probability distributions and their properties.
4	Apply distribution theory to real-world problems and interpret the results using statistical software.
5	Analyze exact sampling distributions (Chi-square, t, and F distributions) and understand their applications in hypothesis testing.

Course Outcomes

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

NO	COURSE OUTCOME	BTL	PO	PSO
CO1	define and distinguish between discrete and continuous random variables, and interpret their probability mass functions (PMF) and probability density functions (PDF).	K2	1	1
CO2	compute and interpret moments, skewness, and kurtosis for given PMFs and PDFs, and analyze their properties in practical scenarios.	K3	1	1
CO3	use generating functions (moment, cumulant, probability generating functions) to solve problems involving moments and distributions, and understand the Weak Law of Large Numbers (WLLN).	K3	1	1
CO4	evaluate real-world problems using the Binomial, Poisson, Uniform, Normal, and Exponential distributions, including calculating probabilities, means, variances, and recurrence relations.	K4	1	1
CO5	analyze and interpret sampling distributions (Chi-square, t, and F distributions), and apply these distributions to hypothesis testing in practical applications.	K5	1	1

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	**2**							**2**	
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

Course Structure:

Unit – 1 **Univariate Random Variables** **(12 Hours)**

Definition, Discrete and Continuous random variables -Probability mass function and Probability density function with illustrations. Distribution function and its properties. Calculation of moments, coefficient of skewness and kurtosis for a given probability mass function (PMF) and Probability Density function (PDF)

Examples/Applications/Case Studies:

1. A random experiment of tossing a coin can be modeled using a random variable that represents the outcome of getting a head.
2. Rolling a die - Number of dots on the upper face

Exercises/Project:

1. Simulation

a. Introduce random number generators and simulation tools.
b. Provide a random variable scenario (e.g., flipping a coin, rolling a die, generating random numbers).
c. Have students use technology to simulate the experiment and collect data.

2. Probability Distributions

a. Introduce different types of probability distributions (discrete and continuous).
b. Provide examples of random variables.
c. Divide students into groups.
d. Assign each group a random variable.
e. Ask them to determine the possible values and their corresponding probabilities.
f. Create a probability distribution table or graph.
g. Discuss the characteristics of the distribution.

Specific Resources: (web):

https://www.khanacademy.org/math/statistics-probability/random-variables-stats-library

Unit – 2 Mathematical Expectations and Generating functions (12 Hours)

Definition, Mathematical expectation of a random variable, Properties of expectations. Moments and covariance using mathematical expectation and their properties. Addition and Multiplication theorems on expectation of two and n variables. Variance and its properties

Generating functions: Moment Generating Function, Cumulates Generating Function, Characteristic Function, Probability Generating Function (p.g.f.) Chebyshev's inequality, Weak Law of Large Numbers (W.L.L.N.) (statements only), simple problems.

Examples/Applications/Case Studies:

1. **Fair Games:** Used in gambling and game theory to determine the expected payoff or loss.
2. **Decision Making:** Helps in evaluating options by calculating expected values of different choices.

Exercises/Project:

1. Lottery Game Simulation

a. Create a simple lottery game with different prize values and corresponding probabilities.
b. Divide students into groups.
c. Each group should simulate playing the lottery multiple times, recording their winnings.

2. Insurance Premium Calculation

a. Present a hypothetical insurance scenario (e.g., car insurance).

b. Provide data on the probability of different types of accidents and corresponding claim amounts.
c. Calculate the expected claim amount per policyholder.

Specific Resources: (web):

https://ocw.mit.edu/courses/18-05-introduction-to-probability-and-statistics-spring-2022/resources/lecture-notes/

Unit–3: Discrete Probability Distributions

Discrete Uniform distribution – definitions, mean, variance.

Binomial distribution – Definition, Moments, Mode and Recurrence relation for probabilities. Poisson distribution – Definition, Moments, Mode and Recurrence Relation for probabilities simple problems.

Examples/Applications/Case Studies:

1. Rolling a fair die. Each number (1-6) has an equal probability of appearing.
2. Random number generation, simulations, load balancing in computer networks.

Exercises/Projects:

1. Simulation**:** Simulate the tossing of a fair coin 1000 times and compare the observed frequency of heads to the theoretical probability.
2. Gambling: Analyze the probability of winning different types of lottery games using binomial distribution

Specific Resources: (web) https://stattrek.com/

2. https://stattrek.com/probability-distributions/poisson.aspx

Unit–4: Continuous Probability Distributions

Continuous Uniform distribution – Definition, mean, variance and mean deviation about mean, Moment generating function and Characteristic function (statements only) and simple problems. Exponential distribution – Definition, mean and variance, Moment generating function and Characteristic function. Memory less property (statements only) and simple problems. **Normal distribution –** Definition, Properties of normal distribution, importance of normal distribution, Moment generating function, Characteristic function, Cumulant generating function of normal distribution. (Statements only) and simple problems.

Examples/Applications/Case Studies:

1. **Random Number Generation:** Many programming languages and statistical software use the continuous uniform distribution to generate random numbers within a specified range.

   Exponential distribution

2. **Customer Service:**

Scenario: A call center receives an average of 20 calls per hour.

Analysis: The time between calls follows an exponential distribution. This can be used to model wait times, staffing requirements, and system performance.

Exercises/Projects:

Project 1: Simulating Dice Rolls
Goal: Simulate the rolling of a fair six-sided die using the continuous uniform distribution.
Steps:

1. **Generate random numbers:** Use the run if() function in R to generate a large number of random numbers between 1 and 6.
2. **Round to integers:** Round the generated numbers to the nearest integer to simulate a die roll.

Analyze results: Calculate the frequency of each outcome, compare it to the expected frequency (1/6 for each outcome), and visualize the results using a histogram.

Queuing System Simulation

Exponential distribution

Project Goal: Simulate a queuing system (e.g., a call center, a grocery store checkout) using the exponential distribution to model inter-arrival and service times.
Steps:

1. **Define parameters:** Determine the average arrival rate and service rate for the system.
2. **Generate inter-arrival and service times:** Use the exponential distribution to generate random inter-arrival and service times for customers.
3. **Simulate the system:** Create a simulation model to track customer arrivals, waiting times, and service times.
4. **Analyze performance:** Calculate metrics like average waiting time, system utilization, and queue length.

Normal Distribution

Project 1: Analyzing Stock Returns
Goal: Analyze the distribution of stock returns and test for normality.
Steps:

1. **Collect data:** Gather historical stock price data for a specific stock or index.
2. **Calculate returns:** Calculate the daily or weekly returns of the stock or index.
3. **Visualize the distribution:** Create a histogram and a normal probability plot to assess the normality of the returns.
4. **Test for normality:** Use a statistical test, such as the Shapiro-Wilk test or the Kolmogorov-Smirnov test, to formally test for normality.
5. **Analyze results:** If the returns are normally distributed, you can use statistical methods based on the normal distribution, such as hypothesis testing or risk analysis. If the returns are not normally distributed, you may need to consider alternative models or techniques.

Specific Resources: (web):
https://www.pnw.edu/wp-content/uploads/2020/03/lecturenotes5-10.pdf

Unit – 5 Exact Sampling Distributions (12Hrs)

Definitions of population, sample, parameter, statistic, sampling distribution. Standard Error and its uses**, χ2 - Distribution**– Definition, Properties and applications. **Student 's t-**

distribution– Definition, Properties and applications. **F – Distribution** – Definition, Properties and applications.

Examples/Applications/Case Studies:

Case Study 1: Sampling Distribution of Sample Mean

Scenario: A researcher wants to estimate the average height of adult males in a city. They randomly sample 100 adult males and measure their heights.

Steps:

1. **Collect data:** Collect the heights of the 100 adult males.
2. **Calculate sample mean:** Calculate the mean height of the sample.
3. **Repeat sampling:** Repeat steps 1 and 2 many times (e.g., 10,000 times) to create a distribution of sample means.
4. **Analyze sampling distribution:** Analyze the distribution of sample means. It should be approximately normal with a mean equal to the population mean and a standard deviation equal to the population standard deviation divided by the square root of the sample size (standard error of the mean).
5. **Make inferences:** Use the sampling distribution to make inferences about the population mean height. For example, construct a confidence interval or conduct a hypothesis test.

Exercises/Projects:

Project 1: Simulating the Sampling Distribution of the Sample Mean

Goal: Simulate the sampling distribution of the sample mean from a normally distributed population.

Steps:

1. **Define population parameters:** Specify the population mean (μ) and population standard deviation (σ) of the normal distribution.
2. **Generate random samples:** Draw random samples of a specified size (n) from the normal distribution using the rnorm() function in R.
3. **Calculate sample means:** Calculate the mean of each sample.
4. **Repeat sampling:** Repeat steps 2 and 3 a large number of times (e.g., 10,000) to create a distribution of sample means.
5. **Visualize sampling distribution:** Plot a histogram or density plot of the distribution of sample means.
6. **Compare with theory:** Compare the observed sampling distribution to the theoretical sampling distribution, which is a normal distribution with mean μ and standard deviation $\sigma/\sqrt{n}$.

Project 2: Estimating Population Mean with Confidence Intervals

Goal: Estimate the population mean from a sample and construct confidence intervals using the sampling distribution of the sample mean.

Steps:

1. **Collect data:** Collect a random sample of size n from the population.
2. **Calculate sample mean and standard deviation:** Calculate the sample mean ($\bar{x}$) and sample standard deviation (s) of the data.
3. **Calculate standard error:** Calculate the standard error of the mean using the formula $SE = s/\sqrt{n}$.
4. **Construct confidence intervals:** Construct confidence intervals for the population mean using the t-distribution with n-1 degrees of freedom. For example, a 95% confidence interval can be calculated as $\bar{x} \pm t(\alpha/2, n-1) * SE$, where $t(\alpha/2, n-1)$ is the t-value corresponding to the desired confidence level and degrees of freedom.
5. **Interpret results:** Interpret the confidence intervals and make inferences about the population mean.

Specific Resources: (web):

https://egyankosh.ac.in/bitstream/123456789/14027/1/Unit-4.pdf

Note: No proofs of statements of unit 1 to 5.

Textbook:

1. S. C. Gupta, Fundamentals of Statistics, 8th Edition, 2023, Himalaya Publishing House Pvt. Ltd 'Ramdoot', Dr. Bhalerao Marg, Girgaon, Mumbai – 400 004, Maharashtra, India

References Books:

1. **Business Statistics A First Course, 8e Paperback – 30 October 2022, David.** Levine (Author)
2. Business Statistics: Problems & Solutions by J.K. Sharma (Author), Vikas Publishing House Pvt Ltd. Noida, UP, India

PARVATHANENI BRAHMAYYA
SIDDHARTHA COLLEGE OF ARTS & SCIENCE
Autonomous
Siddhartha Nagar, Vijayawada–520010
Re-accredited at 'A+' by the NAAC

23STMIL233: Random Variables and Probability Distributions

BBA Honours Business Analytics **Max. Marks: 70** **Semester III**

Max. Time: 3Hrs

Section - A

Answer the following questions **5 X 4M = 20M**

1. a. Define the Random variables and state its types.(CO1-K1)

(OR)

b. Define distribution function and state its properties. (CO1-K1)

2. a. Write the statements of Addition and Multiplication theorems of Mathematical Expectation. (CO2-K1)

(OR)

b. Define M.G.F and write its properties. (CO2-K1)

3. a. Define binomial distribution and write its applications. (CO4-K2)

(OR)

b. Define Poisson distribution and write its properties. (CO4-K2)

4. a. Write the properties of Normal distribution (CO4-K1)

(OR)

b. Define exponential distribution. (CO4-K1)

5. a. Define a) sample b) statistic c) sampling distribution. (CO5-K1)

(OR)

b. Define student-t distribution. (CO5-K1)

Section - B

Answer the following questions **5 X 10M = 50M**

6 a. A random variable has the following probability distribution(CO1-K3)

x	0	1	2	3	4	5	6	7	8
P(X= x)	a	3a	5a	7a	9a	11a	13a	15a	17a

(i) Determine 'a'

(ii) Find $P(X<3)$, $P(X \geq 3)$ and $P(0<X<5)$

(iii) Find the distribution function of X.

(OR)

b. The diameter of an electric cable, say X, is assumed to be a continuous random variable with p. d. f. $f(x) = 6x(1-x),\ 0 \leq x \leq 1$. (CO1-K3)

i) Check that f(x) is p. d. f.

ii) Determine a number **b** such that P(X <**b**) = P(X >**b**).

7 a. A symmetric die is thrown 600 times. Find the lower bound for the probability of getting 80 to 120 sixes. (CO3-K3)

(OR)

b. Given the following table: (CO3-K3)

X = x	-3	-2	-1	0	1	2	3
P(X = x)	0.05	0.1	0.3	0	0.3	0.15	0.1

Compute (i) E(X), (ii) E(2X+3), (iii) V(X) and (iv) V(2X+3)

8 a. If a Poisson distribution such that 3P(x=1) = 2P(x=3). Find P $(2 \leq X \leq 5)$ (CO4-K3)

(OR)

b. The probability of a men hitting a target is ¼ (i) If he fires 7 times what is the probability of his hitting the target at least twice. (ii) how many times must he fire so that the probability of his hitting the target at least once is greater than 2/3? (CO4- K3)

9. a Telephone calls arrive at a switchboard following an exponential distribution with parameter $\lambda = 12$ per hour. If we are at the switchboard, what is the probability that the waiting time for a call is i) at least 15 minutes ii) not more than 10 minutes. (CO4-K3)

(OR)

b. In an examination it is laid down that a student passed if he secured per cent or more marks. He is placed in the first, second or third division according has he secures 60% or more marks, between 45% and 60%marks and marks between 30% and 45% respectively. He gets distinction in case he secures 80% or more marks. It is noticed from the result that 10% of the students failed in the examination, whereas 5% of them obtained distinction. Calculate the percentage of the students placed in the second division. (CO4 -K3)

10 a. Write a short note on Chi-square distribution and write its applications. (CO5-K2)

(OR)

b. Write properties and applications of F-statistic. (CO5-K2)

***