

BOS	Computer Science
Class	F.Y. B.C.A
Semester	II
Course Name	Statistical Methods & Testing of Hypothesis
Course Code	PUSCA203
Type of Course	Core
Level of the Subject	Basic
Credit Points	2 Theory+1 Practical

Course Objectives:

1. Understand an interest in the concepts of ancient methods of learning Mathematics.
2. Inculcate interest in research through analyzing the data with the help of R.

Unit No.	Name of Unit	Topic No.	Contents	Hours
1	IKS and Sampling distributions	1.1	Squares and square roots , Cubes and cube roots, Divisibility, Strategies for Enhanced Mental Calculations- Nikhilam Sutra Nikhilam Sutra,Urdhva Tiryak Sutra,Ekadhikena Purvena Sutra, Anurupye Sutra, Yavadunam Tavadunikritya Varga Samam	15
		1.2	Introduction, Factors that influence sampling distribution, Types of distributions- Sampling distribution of mean/ proportion	
		1.3 ✓	Binomial Distribution- Properties and problems based on Binomial distribution Poisson Distribution- Properties and problems based on Poisson distribution	
		1.4	<u>Normal distribution-properties and problems based on Normal distribution,</u> Central limit theorem, Chi square distribution -definition and properties, t distribution - definition and properties , F distribution -definition and properties	
		2.1	Hypothesis- Null and Alternative, Types of error in hypothesis testing, level of significance, One tailed two-tailed test,	

2	Testing of Hypothesis (Parametric test)		critical region, p-value, Confidence interval for mean and proportion	15
		2.2	Large sample test (z test)-single mean, two means, single proportion, two proportions	
		2.3	Small sample test(t test) one sample mean, paired t test, unpaired t test	
3	ANOVA and Chi-Square test	3.1	Application and importance of ANOVA	15
		3.2	One Way ANOVA - procedure and examples	
		3.3	Chi-square test of goodness of fit , Chi-square test for association, Chi square test for independence of attributes, Yates correction	
Total Hours				45

Course Outcome:

1. Identify when to use a parametric method. Different parametric methods in estimation, testing, model fitting, and in analyses.
2. Develop the ability to analyze a problem and understand the appropriate statistical technique to analyze it.
3. Analyze the use of the inferential statistical tools to analyze a problem.
4. Apply Parametric statistical hypothesis testing to make a decision.
5. Explain the results obtained using statistical tools based on a problem scenario. and introduces ANOVA for analyzing a problem in higher level .
6. Understand the tricks to do the mathematical calculations with ease.

Reference:

1. Ross, S.M. (2006): A First course in probability. 6th Edⁿ Pearson
2. Kulkarni, M.B., Ghatpande, S.B. and Gore, S.D. (1999): Common statistical tests. Satyajeet Prakashan, Pune
3. Gupta, S.C. and Kapoor, V.K. (2002): Fundamentals of Mathematical Statistics, S. Chand and Sons, New Delhi
4. Gupta, S.C. and Kapoor, V.K. (4th Edition): Applied Statistics, S. Chand and Sons, New Delhi
5. Trivedi, K.S.(2009) : Probability, Statistics, Design of Experiments and Queuing theory, with applications of Computer Science, Prentice Hall of India, New Delhi
6. Statistical Methods , S.P. Gupta

Case Study:

1	<p>Testing the Mean Lifetime of Fluorescent Light Bulbs</p> <p>A company manufactures fluorescent light bulbs and claims that the mean lifetime of their bulbs is 1600 hours. However, there are concerns about the actual mean lifetime being less than the claimed value. As a data analyst, you have been tasked with conducting a hypothesis test to determine whether there is sufficient evidence to support the claim that the mean lifetime of the company's bulbs is indeed 1600 hours. A sample of 400 fluorescent light bulbs produced by the company has been selected for analysis. The sample has a mean lifetime of 1570 hours with a standard deviation of 150 hours.(Based on the data , the student has to find the interpretations)</p>
2	<p>Analyzing Customer Satisfaction Levels in a Restaurant Chain</p> <p>A restaurant chain wants to assess the satisfaction levels of its customers across different locations. They have collected data on customer feedback regarding their dining experience, categorized into three satisfaction levels: "Satisfied," "Neutral," and "Dissatisfied." The restaurant chain aims to determine if there is a significant difference in customer satisfaction levels among its various locations. As a data analyst, you are tasked with conducting a Chi-square test to analyze the data and provide insights to the management.</p> <p>The dataset consists of customer feedback collected from five different restaurant locations. For each location, the number of customers falling into each satisfaction category (Satisfied, Neutral, Dissatisfied) is recorded. (Based on the data , the student has to find the interpretations)</p>

PRACTICALS

S.N.	Topic
1	R program on Binomial distribution a. Finding binomial probabilities b. Finding cumulative probabilities c. Plotting the graph of pmf and cdf d. Fitting of binomial distribution
2	R program on Poisson distribution a. Finding poisson probabilities b. Finding cumulative probabilities c. Finding quantile values
3	R program on Normal distribution a. Generating random numbers b. Finding normal probabilities c. Graphical sketch of standard normal and normal variate pdf and cdf
4	R program on confidence interval a. Confidence interval for population means for large sample

	b. Confidence interval for population proportion for large sample
5	R program on one sample and two sample mean Z test a. Large sample test for one population mean b. Large sample test for two population means
6	R program on one sample and two sample proportion Z test a. Large sample test for one population Proportion b. Large sample test for two population Proportions
7	R program on unpaired t test a. Small sample test for one population mean b. Small sample test for difference of means c. Paired t test
8	R program on Chi square test a. Chi square test for goodness of fit b. Chi square test for independence of attributes c. Yates correction
9	R program on one way Anova a. Case study method b. Real data analysis
10	R program on Two way Anova a. Case study method b. Real data analysis
	Total Hours: 30