Course Code				**23STMAL234**			
Title of the Course				**Inferential Statistics**			
Offered to: (Programme/s)				**B.Sc.(Honors) - Statistics**			
L	**4**	**T**	**0**	**P**	**0**	**C**	**3**
Year of Introduction:	**2024-25**			**Semester:**			**3**
Course Category:	**MAJOR**			**Course Relates to:**		**Local, Regional, National, Global**	
Year of Revision	**NA**			**Percentage:**		**NA**	
Type of the Course:				**SKILL DEVELOMENT**			
Crosscutting Issues of the Course:				**NA**			
Pre-requisites, if any				**23STMAL121, 23STMAL122**			

Course Description:

This course provides a comprehensive introduction to statistical estimation, hypothesis testing, and both parametric and non-parametric tests. Students will learn fundamental concepts in parameter estimation, properties of estimators, and the theory behind maximum likelihood estimation. The course also covers hypothesis testing for both small and large sample sizes, as well as non-parametric methods for analyzing data without strict distributional assumptions. Real-world applications are explored through activities, such as estimating economic indicators and testing hypotheses in healthcare, business, and the social sciences.

Course Objectives:

S. No	COURSE OBJECTIVES
1	introduce the key concepts and properties of statistical estimators, including unbiasedness, consistency, efficiency, and sufficiency.
2	develop students' ability to formulate and test statistical hypotheses using parametric and non-parametric methods.
3	provide practical skills in conducting large and small sample tests, with applications across various fields.
4	teach the application of estimation and hypothesis testing in real-world datasets and decision-making scenarios.
5	compare parametric and non-parametric tests and highlight their advantages and disadvantages in different contexts.

Course Outcomes

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

NO	COURSE OUTCOME	BTL	PO	PSO
CO1	understand and apply various methods of parameter estimation, including maximum likelihood estimation, for common distributions.	**K2**	**2**	**1**
CO2	analyze the properties of estimators, including consistency, efficiency, and sufficiency, and evaluate the quality of different estimators.	**K4**	**2**	**1**
CO3	formulate and test statistical hypotheses for both large and small samples, including single means, proportions, and variances.	**K5**	**2**	**1**
CO4	apply non-parametric methods, such as the Wilcoxon test and Mann-Whitney U test, in situations where parametric assumptions are violated.	**K3**	**2**	**1**
CO5	iInterpret statistical results from estimation, hypothesis testing, and non-parametric methods to make informed decisions in applied fields such as business, healthcare, and social sciences.	**K5**	**2**	**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: Theory of Estimation **(12 hours)**

Estimation of a parameter, criteria of a good estimator – unbiasedness, consistency, efficiency, & sufficiency. Statement of Neyman's factorization theorem. Estimation of parameters by the method of moments and maximum likelihood (M.L), properties of MLE's. Rao – Cramer Inequality, properties. Binomial, Poisson & Normal population parameters estimate by MLE method. Confidence Intervals and limits- Normal distributions.

Applications :

1. Estimating economic indicators like GDP, inflation, and unemployment rates.
2. Estimating consumer preferences and demand for products

Activity :

Collect datasets on heights, weights, test scores.

Calculate the sample mean as an estimate of the population mean.

Discuss the concept of margin of error and confidence intervals.

Unit–2:Testing of Hypothesis (12 hours)

Concepts of Statistical hypotheses, null and alternative hypothesis, critical region, two types of errors, level of significance and power of a test. One and two tailed tests. P value. Steps in solving testing of hypothesis problem. Optimum test under different situations-Most powerful

test(MP Test) and uniformly most powerful test (UMP Test). Neyman and Pearson's Lemma. Examples in case of Binomial, Poisson, Exponential and Normal distributions.

Applications

Hypothesis testing is a powerful statistical tool used across various fields to make informed decisions based on data. Some Areas applications are

1. Medicine and Healthcare
2. Business and Economics

Activity

- Formulating Hypotheses
- Stating the null hypothesis (H_0).
- Stating the alternative hypothesis (H_1)
- Setting the Significance Level (α)
- Determining the Critical Value or P-value
- Finding the critical value.
- Calculating the p-value.

Unit–3:Large Sample Tests (12 hours)

Large sample tests - Single mean and difference of two means, confidence intervals for mean(s), Single proportion and difference of proportions, difference of standard deviations, Correlation Coefficients-Problems.

Applications

Large sample tests, typically employing the Z-test, are widely used in various fields due to their simplicity and robustness. Some application areas are

1. Business and Economics
2. Social Sciences
3. Healthcare

Activity 1:

Student satisfaction: Conduct a survey among a large sample of students to assess satisfaction with various campus services, facilities, or academic programs.

Student behavior : Investigate habits like study patterns, social media usage, or dietary preferences among a representative student sample

Unit – 4: Small Sample Tests (12 hours)

Student's t-test for single mean, difference of means, paired t-test and correlation coefficient. χ^2 test for goodness of fit, independence of attributes and significance of population variance. F test for equality of variances and problems.

Applications

Small sample tests are essential when dealing with limited data. They find applications in various fields where large datasets are difficult or impossible to obtain.

1. Research and Development
2. Quality Control

Activity 1 :

Psychology experiments: Conduct small-scale experiments on perception, memory, or cognitive abilities

Activity 2 :

Study habits: Investigate if there's an association between study habits (library, dorm, group

study) and exam scores

Unit–5:Non-Parametric Tests (12 hours)

Advantages and disadvantages of NP tests, comparison with parametric tests. One sample - Run test, Sign test and Wilcoxon – signed rank test . Two independent sample tests - Median test, Wilcoxon – Mann – Whitney U test, Wald Wolfowitz's runs test and problems.

Applications

These are statistical methods that don't rely on specific assumptions about the population distribution. They are used when parametric tests (like t-tests, ANOVA) cannot be applied due to violations of assumptions or data characteristics. Widely used fields are

1. Social sciences
2. Medicine
3. Biology

Activity 1:

Compare satisfaction levels between different academic programs using the Mann-Whitney U test.

Activity 2:

Analyze if there's a difference in social media usage (measured by time spent) between genders using the Wilcoxon rank-sum test.

Textbooks:

1. S. C. Gupta and V. K. Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand & Sons, New Delhi
2. A.M.Goon,M.K. Gupta, B.Dasgupta, An outline of Statistical theory, The World Press Pvt.Ltd., Kolakota

References

1. K.M.Ramachandran and Chris P.Tsokos, Mathematical Statistics with Applications, Academic Press(Elsevier), Haryana
2. D. Biswas, Probability and Statistics, New central book Agency (P) Ltd, NewDelhi

***

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

23STMAL234 : Inferential Statistics

Major 3 **B.Sc. Honours (Statistics)** **Semester : III**

Time: 3 hours **Maximum Marks: 70**

Section – A

Answer the following **5 x 4M = 20Marks**

1. (a) Find the unbiased estimator of mean for normal population. (CO1-K2)
   OR
   (b) Explain method of moments (CO1-K2)

2. (a) Write a short note on types of errors (CO2-K1)
   OR
   (b) Define null and alternative hypothesis with examples (CO2-K1)

3. (a) Explain the testing procedure (CO3-K2)
   OR
   (b) Explain the test procedure of testing single mean in large samples(CO3-K2)

4. (a) Explain the test procedure of testing paired mean in small samples(CO4-K2)
   OR
   (b) Explain the test procedure of testing equality of two population variances(CO3-K2)

5. (a) Explain sign test (CO4-K2)
   OR
   (b) Explain run test (CO4-K2)

Answer the following **5 x 10M = 50Marks**

6. (a) Explain the characteristics of a good estimator (CO1-K2)
   (OR)
   (b) Find Maximum likelihood estimator for μ and σ^2 in normal population. (CO1-K2)

7. (a) State and prove Neyman-Pearson's lemma. (CO2-K3)
   (OR)
   (b) If $x \geq 1$ is the critical region for testing $H_0: \theta=2$ vs $H_1: \theta=1$ on the basis of the single observation from an exponential distribution with probability density function $f(x, \theta)= \theta e^{-\theta x}$. Obtain the value of Type I and Type II errors. (CO2-K3)

8 (a) In a Survey of buying habits, 400 women shoppers are chosen at random on Supermarket 'A' located in a certain section of the city. Their average weekly food expenditure is Rs.250 with a S.D. of Rs. 40. For 400 women shoppers are chosen at random on Supermarket 'B' in another section of the city, the average weekly food expenditure is Rs.220 with a S.D. of Rs 55.Test at 1% level of significance whether the average weekly food expenditure of the populations of shoppers are equal. (CO3-K4)

   OR

   (b) A candidate for an election made a speech in city A but not in B. A sample of 500 voters from city A showed that 59.6% of the voters were in favour of him where as a sample of 300 voters from city B showed that 50% of the voters favoured him discuss whether his speech could produce any effect on

voters in city A. Use 5% level. (CO3-K4)

9. (a) The following data relate to the marks obtained by 11 students before and after intensive coaching . Do the data indicate that the students have benefited by the intensive coaching (CO3-K4)

Marks before coaching	19	23	16	24	17	18	20	18	21	19	20
Marks after coaching	17	24	20	24	20	22	20	20	18	22	19

OR

(b) In one sample of 8 observations the sum of squares of deviations of the sample values from the sample mean was 84.4 and in another sample of 10 observations it was 102.6 . Test the significant difference between the sample variances at 1% level of significance (CO3-K4)

10. (a) A sample of 48 tools produced by a machine shows the following sequence of good (G) and defective (D) tools :

G G G G G G D D G G G G G G G G G D D D D G G G G G D G G G G G G G G D D G G G G G D G G.

Test the randomness of the sequence at the 0.05 significance level (CO4-K4)

OR

(b) Find out whether the distributions of the two diets is same are not using median test. (CO4-K4)

Diet A: 16.3,10.1,10.7,13.5,14.9,11.8,14.3.10.2,12,14.7,23.6,15.1, 14.5,18.4,13.2,14,24.2,27.2,15.9,16.4,22.3

Diet B: 21.3,23.8,15.4,25.1,19.6,12.1,13.9,18.8,19.2,15.3,20.1, 14.8,18.9,20.7,21.1,26.4,18.6,15.8,16.2,17.8,23.6,24.3

***