P.B. SIDDHARTHA COLLEGE OF ARTS & SCIENCE

Siddhartha Nagar, Vijayawada – 520 010
Autonomous -ISO 9001 – 2015 Certified
Inferential Statistics

Offered to: B.SC (AI&ML)/ 22STAT37
Course Type: Core (Theory)
Year of Introduction: 2022   **Percentage of Revision: Nil**
Semester: III   **Paper No. : III**   **Credits: 4**
Hours Taught: 60 periods. per Semester   **Max. Time: 3** Hours
Course Prerequisites (if any): Student required basic knowledge in Probability and Distribution Theory

Course Description:
This course helps the students to familiarize with the ways in which we talk about uncertainty and estimate their situations in which probability arises. Also this course aims at providing basic knowledge about theoretical and application to test according to situations.

Course Objectives:
1)   To understand the problem of statistical inference with specific reference to point estimation and interval estimation.
2)   To differentiate between large and small samples and apply apt testing procedures.

   Learning Outcomes: At the end of the course, the student will
1)   Students will understood the distinguish between the parametric and Non Parametric situations.
2)   The parameters describe an underlying physical setting in such a way that their value affects the distribution of the measured data..

S. No	Program Outcomes
PO1.	**Effective Communication**: Speak, read, write and listen clearly in person and through electronic media in English and in one Indian language, and make meaning of the world by connecting people, ideas, books, media and technology
PO2.	**Effective Citizenship**: Demonstrate empathetic social concern and equity centered national development, and the ability to act with an informed awareness of issues and participate in civic life through volunteering.
PO3.	**Ethics:** Recognize different value systems including your own, understand the moral dimensions of your decisions, and accept responsibility for them.
PO4.	**Environment and Sustainability**: Understand the issues of environmental contexts and sustainable development
PO5.	**Critical Thinking**: Take informed actions after identifying the assumptions that frame our thinking and actions, checking out the degree to which these assumptions are accurate and valid, and looking at our ideas and decisions (intellectual, organizational, and personal) from different perspectives.
PO6:	**Specialized Skills / Transferable Skills:** Acquisition of communication and soft, analytical and technological skills that aid in enhancing
PO7.	**Self-directed and Life-long Learning:** Acquire the ability to engage in independent and life-long learning in the broadest context socio-technological changes

Course Outcomes:

Course Outcome	Upon successful completion of this course, students should have the knowledge and skills to:	Program Outcomes Mapping
CO 1	knowledge of point and interval estimation procedures and different methods of point estimation	PO - 5
CO 2	various basic concepts on sampling distributions and large sample tests based on normal distribution	PO - 6
CO3	Obtain the knowledge on various testing hypothetical statements and finding Uniformly Most Powerful Test	PO - 6
CO 4	a fundamental understanding of Parametric models for developing relevant inferences on associated parameters large and small samples.	PO - 6
CO 5	To obtain the knowledge and to know the applications of various Non-Randomized tests	PO - 6

Syllabus

Course Details

Unit	Learning Units	Lecture Hours
I	**Theory of Estimation:** Parameter, Statistic, Standard Error of the statistic, concept of bias and mean square error of an estimate, Criteria of good estimator - unbiasedness, consistency, efficiency, and sufficiency. Maximum Likelihood estimator (MLE) . ML estimates of μ & σ^2 . Concepts of confidence interval	12
II	**Testing of Hypothesis** Statistical hypotheses, critical region, level of significance and power of a test, types of errors. Neyman Pearson lemma (Statement only) and its applications.	12
III	**Exact Sampling distributions** Student's t-distribution, Chi-square distribution, Snedecor's F-distribution – definitions, properties and applications. **Small Sample tests - I** Chi-square test for goodness of fit and independence of attributes. t-test for single mean, difference of means and paired t-test.	12
IV	**Large sample Tests** Procedure for testing of hypothesis - Test for single mean and difference of two means, test for single proportion and difference of proportions. **Small Sample tests - II** F-test for equality of two population variances, ANOVA I- way and II-way classifications	12
V	**Non - Parametric methods** Definition, advantages and disadvantages. Advantages and Disadvantages, Measurement scales - Nominal, Ordinal, Interval and Ratio. One sample test- Sign test, Run test Two independent sample tests: Median test, Wilcoxon- Mann Whitney U - test, Kruskal Wallis test - Simple Problems	12

Text Book:

Fundamentals of Mathematical Statistics, 11th Edition, 2010, S. C. Gupta and V. K. Kapoor, Sultan Chand & Sons, New Delhi

Reference Books:

1. B.A/B.Sc. Second Year Statistics(2010) , Telugu Akademi, Hyderabad.
2. Mathematical Statistics with Applications, 2009, K.M.Ramachandran and Chris P.Tsokos Academic Press(Elsevier), Haryana .
3. Probability and Statistics, Volume I & II, D. Biswas, New central book Agency (P) Ltd, NewDelhi.
4. An outline of Statistical theory, Volume II,3rd Edition,2010(with corrections) A.M.Goon,M.K. Gupta, B.Dasgupta ,The World Press Pvt.Ltd., Kolakota.
   Sanjay Arora and Bansi Lal:. New Mathematical Statistics, Satya Prakashan , New Delhi.

***

Max.: 70 Marks   **Min. Pass : 28 Marks**

Statistical Inference

22STAT37

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

1.   a. Explain the criteria of good estimator. (10M)   (Co-1, L-2)

      b. Define the terms population and parameter with examples. (5M)   (Co-1, L- 1)

                                        (OR)

      c. Derive the MLE of the parameter λ in Poisson distribution. (10M)   (Co-1, L-2)

      d. Define the terms sample and statistics with examples. (5M)   (Co-1, L-1)

2.   a. Define Chi-square distribution and write its applications. (5M)   (Co-2, L-1)

      b. Out of 8,000 graduates in a town 800 are females, out of 1,600 graduate employees 120
      are females. Use χ^2 to determine if any distinction is made in appointment the basis of
      sex. (10M)   (Co-2, L-3)

                                        (OR)

      c. Define F-distribution and write its applications. (5M)   (Co-2, L-1)

      d. 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 the other sample of 10 observations it was 102.6.
      Test whether this difference is significant at 1% level. (10M)   (Co-2, L-3)

3.   a. In order to test a coin is perfect or unbiased it is tossed 5 times the null hypothesis of
      perfectness is rejected if and only if more than 4 heads are obtained then calculate (i)
      Critical region (ii) Probability of Type I error (iii) Probability Type II error (when the
      corresponding probability of getting head is 0.2)   (15M)   (Co-3, L-4)

                                        (OR)

      b. Explain Critical region, Types of errors. (5M)   (Co-3, L-2)

      c. Let p be the probability that a coin will follow head in a single toss. In order to test

      $H_0 : p = \dfrac{1}{2}$ *against* $H_1 : p = \dfrac{3}{4}$, the coin is tossed five times. H$_0$ is rejected if more than

      three heads appeared. Find the probability of Type I error and Type II error. (10M)

                                                                                   (Co-3, L-4)

4.   a. The marketing manager of a consumer product company wanted to know whether it is
      worth investing money and efforts in designing different sizes of package design with
      different color. He was wondering if the factors color and size of package could enhance
      the sale significantly. He performed the following experiment. The data matrix
      containing the response variable in 1000 is given below.

	Size of Package		
Color	Large	Medium	Small
Blue	90	96	116
Red	90	110	126
Pink	98	125	149

Perform the two-way ANOVA and test whether the mean sales are influenced by package size and color. What are your findings? (15M)   (Co-4, L-4)

(OR)

b. A sales manager of a large company conducted a sample survey in states A and B taking 400 and 500 samples respectively. The results were

	State A	State B
Average Sales	Rs. 2500	Rs. 2200
Standard Deviation	Rs. 400	Rs. 550

Test Whether the average sales is the same in the 2 states at 1% level.(8M) (Co-4, L-4)

c. A filling machine is expected to fill 5kg of powder into bags. A sample of 10 bags gave the weights 4.7, 4.9, 5.0, 5.1, 5.4, 5.2, 4.6, 5.1, 4.6 and 4.7. test whether the machine is working properly. (7M)   (Co-4, L-4)

5.   a. Explain the Non-Parametric methods also write its merits and demerits.(7M) (Co-5, L-4)

b. The number of defective items produced from two machines are observed as follows.

Machine 1	26, 27, 31, 26, 19, 21, 20, 25, 30
Machine 2	23, 28, 26, 24, 22, 19

Test whether these two samples are drawn from the same population by using median test. (8M)   (Co-5, L-4)

(OR)

c. Define nominal, ordinal, interval and ratio data. (5M)   (Co-5, L-1)

d. From a company trainers are selected randomly and divided into 3 groups and each group containing 10 members and there are given a course in the management skills by three different methods. At the end of the training period scores are as follows.

Method A	99	64	101	85	79	88	97	95	90	100
Method B	83	102	125	61	91	96	94	89	93	75
Method C	89	98	56	105	87	90	87	101	76	89

By using Kruskalwallis test to determine if the three methods are equally effective (or) not at 5% level. (10M)   (Co-5, L-4)

****

Statistical Data Analysis using SPSS-II

Offered to: B.SC (AI&ML)/ 22STAL37
Course Type: Core (Practical)
Year of Introduction: 2022   **Percentage of Revision: Nil**
Semester: III   **Paper No. III**   **Credits: 1**
Hours Taught: 30 periods. per Semester   **Max.Time: 2** Hours
Course Prerequisites (if any): Student required basic knowledge in computers
Course Description:
This course gives a working knowledge of SPSS software to students with the aim of getting to use data analysis. Students will be able to apply appropriate statistical tool for given data set using SPSS Software and get the output and report the finding.
Course Objectives
1)   To train students in SPSS Software
2)   To expose the students to the analysis of statistical data.
Learning Outcomes: At the end of the course, the student will
1)   able to do data analysis using SPSS
2)   known to choose the data to test various types.

S. No	Programme Outcomes
PO1.	**Effective Communication**: Speak, read, write and listen clearly in person and through electronic media in English and in one Indian language, and make meaning of the world by connecting people, ideas, books, media and technology
PO2.	**Effective Citizenship**: Demonstrate empathetic social concern and equity centred national development, and the ability to act with an informed awareness of issues and participate in civic life through volunteering.
PO3.	**Ethics:** Recognize different value systems including your own, understand the moral dimensions of your decisions, and accept responsibility for them.
PO4.	**Environment and Sustainability**: Understand the issues of environmental contexts and sustainable development
PO5.	**Critical Thinking**: Take informed actions after identifying the assumptions that frame our thinking and actions, checking out the degree to which these assumptions are accurate and valid, and looking at our ideas and decisions (intellectual, organizational, and personal) from different perspectives.
PO6:	**Specialized Skills / Transferable Skills:** Acquisition of communication and soft, analytical and technological skills that aid in enhancing
PO7.	**Self-directed and Life-long Learning:** Acquire the ability to engage in independent and life-long learning in the broadest context socio-technological changes

Course Outcomes:

Course Outcome	Upon successful completion of this course, students should have the knowledge and skills to:	Programme Outcomes Mapping
CO 1	To Apply statistical analysis that can test hypotheses under parametric approaches.	PO –6
CO 2	To Apply statistical analysis that can test hypotheses under non-parametric approaches.	PO –6

List of practical's

1. Small sample test (t-test): One Sample, Independent Sample and Paired Sample.

2. Large sample tests: One Sample, Independent Sample, Paired Sample (Using SPSS)

3. Analysis of variance: One-way and Two- way classification (Using SPSS)

4. Chi square Test: Test for Independence of Attributes

5. Chi square Test: Goodness of fit

6. Chi square Test: Test of Independence, 2X2, 3X3,…, mXn Cross tabulation (Using SPSS)

7. Non Parametric Tests: Mann Whitney U test and Wilcoxon Signed ranks test

8. Non Parametric Tests: Kruskal Wallis Test and Friedman test (Using SPSS)

****

Structure of Practical Paper

Total Marks: 50 Marks

(i) For Continuous Evaluation   :   15 marks (Internal Evaluation)
(ii) For semester end Practical Examination   :   35 marks (External Evaluation)