P.B. SIDDHARTHA COLLEGE OF ARTS & SCIENCE

Siddhartha Nagar, Vijayawada – 520 010

Autonomous -ISO 9001 – 2015 Certified

Probability Distributions and Testing of Hypothesis

Offered to: B.SC (CSCS-Computer Science with Cognitive Systems)/ 22STAT36
Course Type: Core (Theory)
Year of Introduction:2021 **Percentage of Revision: Nil**
Semester: III **Paper No. 2** **Credits: 4**
Hours Taught: 60periods per Semester **Max.Time: 3** Hours
Course Prerequisites: Students required basic knowledge in statistics and probability
Course Description: This course helps the students to familiarize students with the ways in which we talk about uncertainty and look at everyday situations in which probability arises. The imparts basic ideas about hypotheses testing procedures and Testing of Hypotheses focuses on solving practical statistical problems.

Course Objectives:

1) To understand the definitions of discrete, continuous, and joint random variables, compute the mean, variance and covariance of random variables, know the definition of mass (density) function and distribution function of a random variable and be able to find one from the other, and be able to find the marginal mass (density) function and distribution functions from the joint mass (density) function and distribution function.
2) To introduce the concepts of hypothesis testing.
3) To differentiate between large and small samples and apply apt testing procedures.

Learning Outcomes: At the end of the course, the student will

1) able to understand to identify the nature of the data for various distributions
2) get the knowledge to test the data using various testing procedures for quantitative and qualitative variables

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:** Acquisitionof 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	Acumen to apply Random Variable and expectation to data standard discrete probability distribution to different situations.	PO - 5
CO 2	knowledge of important discrete distributions such as Binomial, Poisson, Geometric distributions and relations with some other distributions	PO - 6
CO3	knowledge of important continuous distributions such as Uniform, Normal, Exponential and Gamma and relations with some other distributions	PO – 6
CO 4	Demonstrate the computation skills to estimate the parameters in point and interval forms and also getting the knowledge of formulating different hypothesis	PO - 7
CO 5	Testing the Qualitative and Quantitative factors in case of one and two samples using standard normal variate, student's t ,F-statistic and chi square test statistic and Quantitative factors in case of more than two samples using ANOVA	PO - 7

Syllabus

Course Details

Unit	Learning Units	Lecture Hours
I	**Random Variables & Expectations** Univariate – Definition and types. Bivariate Random variables – Definition. Probability mass function and probability density function with illustrations. Distribution function and its properties - Simple problems. Mathematical Expectations: Definition, Properties. Variance – Definition, Properties. Generating Functions statements of their properties with applications.	**12**
II	**Discrete Distributions** Discrete Probability Distributions – Binomial, Poisson and Geometric distributions - Definitions, properties and application, simple problems.	**12**
III	**Continuous Distributions** Continuous Probability Distributions – Uniform distribution (rectangular), Exponential and normal distributions - Definitions, properties and application, simple problems.	**12**
IV	**Exact Sampling distributions** Chi – square, Student's t distributions and Sendecor's F - definition, properties and applications, Problems based on small sample tests - Single mean, Difference of means, Paired t-test and difference of variances. Problems based on chi-square tests – Goodness of fit and Independence of attributes. Analysis of Variances – One Way and Two Way classifications. Testing - Goodness of fit, Independence of attributes.	**12**
V	**Testing of Hypothesis** Definitions of Parameter, Statistic, Standard Error of the statistic- mean and proportion, Concepts of statistical hypotheses – types of hypothesis, Critical region, types of errors, level of significance, power of a test and p-value. Procedure for testing of hypothesis, Problem based on Large samples tests - Single proportion, difference of proportions, single mean and difference of means.	**12**

Note: Proofs and derivations of theorems are excluded

TEXT BOOK:

S.C. Gupta, (2019), Seventh Edition, Fundamentals of Statistics, Mumbai: Himalaya Publishing House.

REFERENCE BOOKS

1. Sharma, J. K. (2013), *Business statistics*, New Delhi: Pearson Education
2. Levine, D.M., Berenson, M. L. & Stephan, D. (2012), *Statistics for managers using Microsoft Excel*, New Delhi: Prentice Hall India Pvt.
3. Aczel, A. D. &Sounderpandian, J. (2011), *Complete Business Statistics*, New Delhi: Tata McGraw Hill.

4. Anderson, D., Sweeney, D., Williams, T., Camm, J., & Cochran, J. (2013), *Statistics for Business and Economics*, New Delhi: Cengage Learning.
5. Davis, G., &Pecar, B. (2014), *Business Statistics using Excel*, New Delhi: Oxford University Press.

Websites of Interest: http://onlinestatbook.com/rvls/index.html

Co-Curricular Activities in the class:

1. Pictionary
2. Case Studies on topics in field of statistics
3. Snap test and Open Book test
4. Architectural – To be build the procedures
5. Extempore – Random concept to students
6. Interactive Sessions
7. Teaching through real world examples

***

Model Question Paper Structure for SEE

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

Section - A

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

1. a) Define Distribution function and write its properties.(CO_1,L_1)

(OR)

b) Find expected number of heads in tossing three coins. (CO_1,L_1)

2. a) Define Geometric distribution. State its properties. (CO_2,L_2)

(OR)

b) A hospital switch board receivers an average of 4 emergency calls in a 10minute interval. What is the probability that there are exactly 3 emergency calls in a 10minute interval? (CO_2,L_2)

3. a) Define Normal distribution. State its properties. (CO_3,L_2)

(OR)

b) Define Exponential distribution. (CO_3,L_2)

4. a) Explain the concepts of the two types of errors. (CO_4,L_4)

(OR)

b) What do you mean by statistical hypothesis explain with examples. (CO_4,L_4)

5. a) A random sample of 200 tins of coconut oil gave an average weight of 4.95 kgs with a standard deviation of 0.21 kg. Do we accept the hypothesis of net weight 5kgs per tin at 1% level? (CO_5,L_3)

(OR)

b) Define Parameter, Statistic, Power of a test, p-value. (CO_5,L_3)

Section -B

Answer the following **5x10M = 50 marks**

6. a) A random X has the following probability distribution.

X=x	-2	-1	0	1	2	3
P(X=x)	0.1	k	0.2	2k	0.3	3k

Find (i) k, (ii) P (X < 2), (iii) P (X ≥ 2), (iv) P (-2 < X < 2) (CO_1,L_1)

(OR)

b) The joint p.d.f of X and Y is given by $f(x,y) = 4xye^{-(x^2+y^2)}\ , x \geq 0, y \geq 0.$ Test whether X and Y are independent. (CO_1,L_1)

7. a) the probability of a man 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? (CO_2,L_5)

(OR)

b) A manufacturer, who produces medicine bottle, finds that 0.1% of the bottles are defective.

The bottles are packed in boxes containing 500 bottles. A drug manufacturer buys 100 boxes from the producer of bottles. Using Poisson distribution, find how many boxes will contain: (i) no defective, and (ii) at least two defectives. (CO_2, L_5)

8. a) If X is uniformly distributed with mean 1 and variance 4/3.

Find (i) $P(X<0)$, (ii) $P(-1 \leq X \leq 2)$ (CO_3, L_1)

(OR)

b) In a distribution exactly normal, 10.03% of the items are under 25-kilogram, weight and 89.97% of the items are under 70-kilogram weight. What are the mean and standard deviations of the distribution? (CO_3, L_1)

9. a) below are given the gains in weights (in kgs.) of pigs fed on two diets A and B.

Diet A: 25, 32, 30, 34, 24, 14, 32, 24, 30, 31, 35, 25

Diet B: 44, 34,22,10, 47, 31, 40, 30, 32, 35, 18, 21, 35, 29, 22

Test, if the two diets differ significantly as regards their effect on increase in weight. (CO_3, L_3)

(OR)

b) 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? (CO_3, L_3)

10. a) A random sample of 500 pineapples was taken from a large consignment and 65 were found to be bad. Show that the S.E of the proportion of bad ones in a sample of this size is 0.015 and deduce that the percentage of bad pineapples in the consignment almost certainly lies between 8.5 and 17.5. (CO_5, L_4)

(OR)

b) A sample of 900 members has a mean 3.4 cms. and s.d. 2.61 cms. Is the sample from a large population of mean 3.25cms. and s.d.2.61cms.?(CO_5, L_4)

****