O.P.Code: 23HS0839	R23	H.T.No.									

SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)
B.Tech. II Year II Semester Supplementary Examinations December-2025
STATTISTICAL METHODS FOR DATA SCIENCE
CSE(Artificial Intelligence & DataScience)

Time: 3 Hours **Max. Marks: 70**

PART-A
(Answer all the Questions **10 x 2 = 20** Marks)

1	a	Define Mathematical Expectation with an example	CO1	L1	2M
	b	The Mean and Variance of a Binomial Distribution are 4 and 4/3 respectively. Find n, p and q.	CO1	L1	2M
	c	What is the formula for the modified minimum Chi-Square method	CO2	L1	2M
	d	Define Estimation, Estimate and Estimator	CO2	L1	2M
	e	Define Confidence Interval and Confidence Coefficient	CO3	L1	2M
	f	Find the Lower, Upper Confidence Limits and also Confidence Coefficient for $P[0\leq\theta\leq1.5]=0.90$	CO4	L3	2M
	g	State Neyman-Pearson Fundamental Lemma.	CO5	L1	2M
	h	Define Null Hypothesis.	CO5	L1	2M
	i	What is the difference between Small sample and Large Sample	CO6	L1	2M
	j	Write the Formula for F- Test.	CO6	L1	2M

PART-B
(Answer all Five Units **5 x 10 = 50** Marks)

UNIT-I

2		Two dice are thrown. Let X assign to each point (a,b) in S the maximum of numbers a,b. i.e, X(a,b) = max(a,b). Find the Probability Distribution of X. Also, find the mean and variance of the distribution.	CO1	L3	10M

OR

3		In a Normal distribution, 7% of the items are under 35 and 89% are under 63. Determine the mean and variance of the distribution.	CO1	L5	10M

UNIT-II

4		Suppose that the random sample has a normal distribution $N(\mu,\sigma^2)$. Determine the Maximum Likelyhood Estimator (i). for μ when $\sigma^2=1$ (ii). for σ^2 when $\mu=0$ is known	CO2	L3	10M

OR

5		The following data represents the body weight Y (in kg), body length X1 (in cm) and body breadth X2 (in cm) of 12 randomly selected sea fish.	CO2	L2	10M

X1	12	20	14	25	18	16	10	18	18	20	16	12
X2	4	7	6	12	10	8	4	8	9	10	7	4
Y	0.5	0.8	0.7	2	1.2	0.9	0.4	0.9	1.4	1.5	0.8	0.6

Assume the linear model of Y, X1 and X2 as $Y = a + bX1 + c\,X2$. Estimate the parameters a, b & c by the method of least squares.

UNIT-III

6	a	A factory is producing 50,000 pairs of shoes daily from a sample of 500 pairs 2% were found to be substandard quality. Estimate the no. of pairs that can be reasonably expected to be spoiled in the daily production and assign limits at 95% confidence level.	CO3	L2	5M
	b	A sample of 7 boxes of a certain cereals with the following weights: 775, 780, 781, 795, 803, 810, 823. Find 95% confidence interval for variance.	CO4	L3	5M

OR

7		From the lots of transparent polyester film sheets produced by two machines samples are taken from both the machines and thickness values of the film sheets in millimicrons are measured. The results are given below.	CO4	L5	

Machine	Sample Size	Mean	Sample Variance
1	10	115	25
2	12	112	9

(i). Compute 95% confidence intervals for the difference of population means assuming that the unknown population variances are equal.
(ii). Compute 99% confidence intervals for the difference of population means assuming that the unknown population variances are not equal.

UNIT-IV

8	a	Explain about Likelihood Ratio Test.	CO5	L2	
	b	Obtain the best critical region for testing null hypothesis $H_0: \lambda=\lambda_0$ against $H_1: \lambda=\lambda_1$ for an exponential distribution $f(x,\lambda)=\lambda e^{-\lambda x}$	CO5	L3	

OR

9		Suppose $x_1, x_2, x_3, \ldots, x_n$ a random sample from a normal distribution with mean μ and variance 16. Find the best critical region with a sample size of n=16 and a significance level $\alpha = 0.05$ to test the simple null hypothesis H_0: $\mu=10$ against a simple alternative hypothesis H_1: $\mu=15$.	CO5	L3	

UNIT-V

10	a	A Sample of 26 Bulbs gives a Mean Life of 990 Hours with a Standard Deviation of 20 Hours. The Manufacturer Claims that the Mean Life of Bulbs is 1000 Hours. Is the Sample not up to the Standard?	CO6	L4	
	b	Find the Maximum Difference that we can expect with Probability 0.95 between the Mean of Samples of Sizes 10 and 12 from a Normal Population if their Standard Deviations are Found to be 2 and 3 Respectively.	CO6	L3	

OR

11		The Nicotine in Milligrams of two Samples of Tobacco were Found to be as Follows.	CO6	L2	

Sample A	24	27	26	21	25	---
Sample B	27	30	28	31	22	36

Can it be said that the two Samples have come from the same Normal Population.

*** END ***