SIDDHARTH INSTITUTE OF ENGINEERING & TECHNOLOGY:: PUTTUR
(AUTONOMOUS)
B.Tech.II Year II Semester Regular Supplementary Examinations July/August-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	Write Some Properties of Random Variables.	CO1	L1	2M
	b	Define Population and Sample with Examples.	CO1	L1	2M
	c	Write any Two Properties of Maximum Likelihood Estimation.	CO2	L1	2M
	d	Define Estimation, Estimate and Estimator.	CO2	L1	2M
	e	Define Interval Estimation with an Example.	CO3	L1	2M
	f	Find the Lower, Upper Confidence Limits and also Confidence Coefficient for $P[0 \le \theta \le 1.5] = 0.90$.	CO4	L3	2M
	g	What is Type-II Error?	CO5	L1	2M
	h	Define Alternative Hypothesis.	CO5	L1	2M
	i	What is the Difference between F-test and t-tests?	CO6	L1	2M
	j	Write the Formula for Chi-square Test for Goodness of Fit.	CO6	L1	2M

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

UNIT-I

2   A Random Variable X has the following Probability Function   l. CO1   L5   10M

x	0	1	2	3	4	5	6	7
P(x)	0	K	2K	2K	3K	K^2	$2K^2$	$7K^2+K$

Determine (i) K (ii) Evaluate $P(x \ge 6)$ and $P(0 < x < 5)$ (iii) Find the Minimum Value of K such that $P(x \le K) > 0.5$ (iv) Mean

OR

3   In a competition of 1000 cases, the mean of a certain test is 14 and   CO1   L1   10M
standard deviation 2.5. Assuming the distribution to be normal, find

      (i). How many students score between 12 and 15

      (ii). How many students score above 18

      (iii). How many students score below 8.

UNIT-II

4   In a watch repair shop, the service time in minutes is 14, 17, 27, 18,   CO2   L1   10M
12, 8, 22, 13, 19 and 12. Give a maximum likelihood estimate of
mean service time with the assumption that the service time follows
an exponential distribution with parameter λ.

OR

5   A training data set of 9 different values for mid semester (say X)   CO2   L2   10M
and end semester (say Y) values are given by

X	10	7	3	16	9	1	7	10	8
Y	42	39	32	50	44	55	43	37	43

Assuming a linear relationship Y on X, Estimate the parameters by
the method of least squares.

6	a	Out of 300 households in a town 123 have T.V sets, find 95% confidence limits to the true value of proportion of households with T.V sets in the whole town.	CO3	L3	5M
	b	In a certain newspaper, it was given that 5% of Canadians are illiterate and that 7% of Americans are illiterate. In a sample of 1005 Canadians and 1015 Americans, what should be the difference be at 95% level of significance.	CO3	L1	5M

OR

7	a	A random sample of size n=100 is taken from a population with $\sigma = 5.1$. Given that the sample mean . Construct 95% confidence interval for population mean μ.	CO4	L3	5M
	b	Among 11 patients in a certain study, the Standard Deviation of the property of interest was 5.8. In another group of 4 patients the S.D was 3.4. Construct a 95% confidence interval for the ratio of the variances of these two populations.	CO4	L5	5M

UNIT-IV

8	a	Let p be the probability that a coin will fall head in a single toss in order to test H_0: p = 0.5 against H_1: p = 0.75 .The coin is tossed 5 times and H_0 is rejected if more than 3 heads are obtained. Calculate the probability of type-1 error and power of test.	CO5	L3	5M
	b	Obtain best critical region for testing null hypothesis $H_0: \lambda = \lambda_0$ against $H_1: \lambda = \lambda_1$ for an exponential distribution $f(x,\lambda) = \lambda e^{-\lambda x}$			5M

OR

9		State and prove Neyman-Pearson Fundamental Lemma.	CO5	L3	10M

UNIT-V

10		A pair of dice are thrown 360 times and the frequency of each sum is indicated below:	CO6	L4	10M

Sum	2	3	4	5	6	7	8	9	10	11	12
Frequency	8	24	35	37	44	65	51	42	26	14	14

Would you say that the dice are fair on the basis of the chi-square test at 0.05 level of significant?

OR

11		Samples of two types of electrical light blubs were tested for length of life and following data were obtained.	CO6	L2	10M

	Type I	Type II
Sample Size	8	7
Sample Mean	1234 hrs	1036 hrs
Sample Standard Deviation	36 hrs	40 hrs

Is the difference in the means sufficient to warrant that Type I is superior to Type II regarding length of life.

*** END ***