Code No: R20A0027

MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY

(AUTONOMOUS INSTITUTION - UGC, GOVT. OF. INDIA)

MODEL QUESTION PAPER-I

III B.Tech I-SEMESTERS REGULAR/SUPPLEMENTARY NOV/DEC-2022 EXAMINATION

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

Note: This question paper consists of 5 sections. Answer **FIVE** Questions, choosing **ONE** question from each section and each question carries 14marks.

SECTION-I

Q.1) What do you mean by Statistics? Explain its importance in Data Science.

OR

Q.2) a) How can we represent categorical and numerical data graphically? Explain in brief.

b) States differ widely in the percentage of college students who are enrolled in public institutions. The National Centre for Education Statistics provided the accompanying data on this percentage for the 50 U.S. states for fall 2002.
Percentage of College Students Enrolled in Public Institutions:
86 96 66 86 80 78 62 81 77 81 77 76 73 69 76 90 78 82 70
83 46 80 78 93 66 91 76 86 58 81 91 57 81 88 71 86 84 57
44 82 79 67 86 75 55 75 80 80 85 69
Construct a frequency distribution table (include cumulative frequencies) and plot the data using appropriate graph. Also interpret the graph.

SECTION-II

Q.3) Calculate the first four moments about mean and also find $\beta_1, \beta_2, \gamma_1, \gamma_2$

Marks	5	10	15	20	25	30	35
Frequency	4	10	20	36	16	12	2

OR

Q.4) Briefly discuss about Monte-Carlo Approximation?

SECTION- III

Q.5) What do you mean by sampling? Discuss the different techniques of sampling.

OR

Q.6) Four brands of flashlight batteries are to be compared by testing each brand in five flashlights. Twenty flashlights are randomly selected and divided randomly into four groups of five flashlights each. Then each group of flashlights uses a different brand of battery. The lifetimes of the batteries, to the nearest hour, are as follows:-

BRAND A	BRAND B	BRAND C	BRAND D
42	28	24	20
30	36	36	32
39	31	28	38
28	32	28	28
29	27	33	25

Preliminary data analyses indicate that the independent samples come from normal populations with equal standard deviations. At the 5% significance level, does there appear to be a difference in mean lifetime among the four brands of batteries? $(F_{0.05,(df_1=3, df_2=16)} = 3.24)$

SECTION- IV

Q.7) Explain in brief about the different methods of estimation.

OR

Q.8) a) Let p be the probability that a coin will fall head in a single toss in order to test : $p = \frac{1}{2}$ against $H_0 : p = \frac{3}{4}$. The coin is tossed 5 times and H_0 is rejected if more than 3 heads are obtained. Find the probability of Type-I error and power of the test.

b) State and explain the double dichotomy relating to decision and hypothesis.

SECTION- V

Q.9) Write few lines about Stochastic Process and its application in data science.

OR

Q.10) a) Define Markov Chain and Transition Probability.

b) The Markov chain with states $\{E_1, E_2, E_3\}$ and the transition probability matrix-

$$P = \begin{bmatrix} 0.6 & 0.4 & 0 \\ 0 & 0.5 & 0.5 \\ 0.2 & 0.4 & 0.4 \end{bmatrix}$$

(i)   Draw the state transition diagram for P
(ii)   Prove that P is irreducible.
(iii) Find $P_{23}^{(2)}$ and $P_{23}^{(4)}$

MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY

(AUTONOMOUS INSTITUTION - UGC, GOVT. OF. INDIA)

MODEL QUESTION PAPER-II

III B.Tech I-SEMESTERS REGULAR/SUPPLEMENTARY NOV/DEC-2022 EXAMINATION

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

Note: This question paper consists of 5 sections. Answer **FIVE** Questions, choosing **ONE** question from each section and each question carries 14marks.

SECTION-I

Q.1) Describe three basic reasons to study Statistics with examples

OR

Q.2) a) Define stem-and leaf chart. When do you use a Stem-and-Leaf Display/Chart and how

to construct it?

b) The UNICEF report "Progress for Children" (April, 2005) included the accompanying data on the percentage of primary-school-age children who were enrolled in school for 19 countries in Northern Africa and for 23 countries in Central Africa.

Northern Africa

54.6, 34.3, 48.9, 77.8, 59.6, 88.5, 97.4, 92.5, 83.9, 96.9, 88.9,

98.8, 91.6, 97.8, 96.1, 92.2, 94.9, 98.6, 86.6

Central Africa

58.3, 34.6, 35.5, 45.4, 38.6, 63.8, 53.9, 61.9, 69.9, 43.0, 85.0,

63.4, 58.4, 61.9, 40.9, 73.9, 34.8, 74.4, 97.4, 61.0, 66.7, 79.6,

98.9.

Construct a comparative Stem-and-Leaf chart and interpret it.

SECTION-II

Q.3) State the Bayes' Theorem and its application. A box contains 3 blue and 2 red marbles while another box contains 2 blue and 5 red marbles. A marble drawn at random from one of the boxes turns out to be blue. What is the probability that it came from the first box?

OR

Q.4) a) What is the difference between conditional probability and combinatorial probability? Illustrate in brief with examples.

b) The joint density function of two continuous random variables X and Y is-

$$f(x,y) = \begin{cases} c - x - y & 0 \le x \le 1, 0 \le y \le 1 \\ 0 & otherwise \end{cases}$$

i.   Find the value of the constant c
ii. Conditional density function of X and Y.
iii. Find the variants of X and Y
iv. Find the marginal distribution functions of X and Y.
v.   Co-variance between X and Y

SECTION- III

Q.5) Differentiate between Observational and Experimental studies.

OR

Q.6) Three sets of five mice were randomly selected to be placed in a standard maze but with Different color doors. The response is the time required to complete the maze as seen below. Perform the appropriate analysis to test if there is an effect due to door color-
(Given $F_{0.01,(df1 = 2, df2 = 12)} = 6.93$)

Color	TIME				
Red	9	11	10	9	15
Green	20	21	23	17	30
Black	6	5	8	14	7

SECTION- IV

Q.7) Explain in brief about Maximum Likelihood Estimation and its properties. Also find the MLE for the parameter λ of the Poisson distribution and prove that it is **'sufficient'** for estimating the parameter λ.

OR

Q.8) How is Bayesian inference different from Frequentist inference? Discuss the application of Bayesian statistics in data science.

SECTION- V

Q.9) Write in detail about the different types of Stochastic Processes.

OR

Q.10) a) In a small town 90% of all sunny days are followed by sunny days and 80% of all cloudy days are followed by cloudy days. Use this information to model the small town's weather as Markov Chain and sketch the transition diagram for the same. Find the probability that it will be cloudy four days from now, given that it is sunny today
b) Prove that P is Ergodic-

$$P = \begin{bmatrix} 0 & 0.8 & 0.2 \\ 0.3 & 0.7 & 0 \\ 0.4 & 0.5 & 0.1 \end{bmatrix}$$

Code No: R20A0027

MALLA REDDY COLLEGE OF ENGINEERING & TECHNOLOGY

(AUTONOMOUS INSTITUTION - UGC, GOVT. OF. INDIA)

MODEL QUESTION PAPER-III

bIII B.Tech I-SEMESTERS REGULAR/SUPPLEMENTARY NOV/DEC-2022 EXAMINATION

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

Note: This question paper consists of 5 sections. Answer **FIVE** Questions, choosing **ONE** question from each section and each question carries 14marks.

SECTION-I

Q.1) Write in details the six steps involved in data analysis process.

OR

Q.2) a) What are the two types of boxplots? Explain their construction in brief.

b) The technical report "Ozone Season Emissions by State" (U.S. Environmental Protection Agency, 2002) gave the following nitrous oxide emissions (in thousands of tons) for the 48 states in the continental U.S. states:

```
76   22   40   7   30   5   6   136   72   33   0   89   136   39   92   40
13   27   1   63   33   60   27   16   63   32   20   2    15   36   19   39
130  40   4   85   38   7   68   151   32   34   0   6    43   89   34   0
```

Use these data to construct a boxplot that shows outliers and comment on any interesting features of the plot.

SECTION-II

Q.3) a) Define joint, marginal and conditional probability functions with examples.

b) Joint probability mass function of X, Y is given by-
  $P(x,y) = k(2x + 3y);$ $x = 0,1,2$ $y = 1,2,3$
  Find all the marginal and conditional probability and also find probability distribution of X+Y.

X Y	1	2	3	
0	3k	6k	9k	18k
1	5k	8k	11k	24k
2	7k	10k	13k	30k
	15k	24k	33k	

OR

Q.4) a) What do you mean by moments and moment generating function? Explain in brief.

b) A random variable X has the density function given by-

$$f(x) = \begin{cases} 4x(9 - x^2)/81 & 0 \le x \le 3 \\ 0 & otherwise \end{cases}$$

Find the first four moments about the origin and about the mean.

SECTION- III

Q.5) Discuss the biases in sampling.

OR

Q.6) The times required by three workers to perform an assembly-line task were recorded on five randomly selected occasions. Here are the times, to the nearest minute:

HANK	JOSEPH	SUSAN
8	8	10
10	9	9
9	9	10
11	8	11
10	10	9

Construct a one-way ANOVA table and interpret the results. ($F_{2,12}(0.05) = 3.8853$)

SECTION- IV

Q.7) Define Estimation. Write in brief the criteria of a good estimator.

OR

Q.8) Elaborate the general method of test construction given by Neyman and Pearson for testing hypothesis.

SECTION- V

Q.9) What do you mean by Markov Chain? Derive the Chapman-Kolmogorov equations.

OR

Q.10) a) Consider this transition probability matrix $\begin{bmatrix} 0.4 & 0.6 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.3 & 0.7 & 0 \\ 0 & 0 & 0.5 & 0.4 & 0.1 \\ 0 & 0 & 0 & 0.8 & 0.2 \end{bmatrix}$ and construct

the transition diagram. Write few lines on the transition of states of this matrix.

b) Illustrate with proper examples the terms- Transient, Absorbing, Recurrent, Irreducible, Periodic, Aperiodic and Ergodic