



Mid Term (Odd) Semester Examination October 2024

Roll no. 2319680.....

Name of the Course and semester: B.Tech III

Name of the Paper: Introduction to Statistical Data Science

Paper Code: TCS 342

Time: 1.5 hour

Maximum Marks: 50

Note:

- (i) Answer all the questions by choosing any one of the sub questions
- (ii) Each question carries 10 marks.
- (iii) Please specify COs against each question.

Q1. (10 Marks)

a. Define modern data analysis and explain its applications. (CO1)

OR

b. From a lot of 10 items containing 3 defectives, a sample of 4 items is drawn at random without replacement. (CO1)

- (i) Find the probability distribution of X for the no. of defectives.
- (ii) Compute F(x)
- (iii) Find $P(X \leq 1)$, $P(X \geq 1)$, $P(0 < X < 2)$

Q2. (10 Marks)

a. Define sample space and probability and solve the given question. (CO1)

Twelve balls are distributed at random among three boxes. What is the probability that the first box will contain 3 balls.

OR

b. Define Poisson distribution. Fit a Poisson distribution to the following data and calculate theoretical frequencies. (CO1)

x	0	1	2	3	4
f	122	60	15	2	1

Q3. (10 Marks)

a. The length of a species of fish is normally distributed with mean of 5.3 cm and standard deviation of 2.5. A fish is selected at random, what is the probability that its length is (CO1)

- (i) Less than 4.5 cm
- (ii) Greater than 6.4 cm
- (iii) Between 3.6cm and 5.9 cm
- (iv) Between 2.8cm and 4.9cm

OR

b. Find a covariance matrix for the following dataset. (CO2)

$$X = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad Y = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \quad Z = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}$$



Mid Term (Odd) Semester Examination October 2024

Q4.

(10 Marks)

- a. Two random variables X and Y have the following joint probability density function: (CO2)

$$f(x,y) = \begin{cases} 2 - x - y; & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

Find

- (i) Var(X) and Var (Y)
- (ii) Covariance between X and Y

OR

- b. What is outlier. Find the outlier for the following data set representing house prices in \$. (CO2)
Prices= {120, 130, 135, 140, 150, 155, 500}

Q5.

(10 Marks)

- a. Let the random variable X has the distribution:

$P(X=0) = P(X=2) = p; \quad P(X=1) = 1-2p \quad \text{for } 0 \leq p \leq \frac{1}{2}$
For what values of p is the Var (X) a maximum.

(CO2)

OR

- b. State and prove the addition and multiplication theorem of expectation.

(CO2)