

TILAK MAHARASHTRA VIDYAPEETH, PUNE
MASTER OF COMPUTER APPLICATIONS
(Specialization in Artificial Intelligence & Machine Learning) /
(Specialization in Data Science)
EXAMINATION :MAY - 2024
SEMESTER - I
Sub: Probability & Statistics (MCAI 23-101/MCDS23-101)

Date :22/05/2024

Total Marks :60

Time: 2.00 pm to 4.30 pm

Instruction:

1. All questions are compulsory unless and otherwise stated.
2. Bold figures to the right of every question are the maximum marks for that question.
3. Candidates are advised to attempt questions in order.
4. Answers written illegibly are likely to be marked zero.
5. Use of scientific calculators, Log tables, Mollier Charts is allowed.
6. Draw neat and labelled diagram wherever necessary.

Q.1. Answer the following in 2-3lines. (Any 5) (10)

1. A bag contains 3 red, 6 white and 7 blue balls what is the probability that two balls drawn are white and blue
2. Define discrete random variable
3. Define Karl Pearson Coefficient of Correlation
4. If the values of two regression coefficients are $\frac{4}{5}$ **and** $\frac{9}{20}$, what will be the correlation coefficient?
5. State the measures of dispersion
6. Define types of error in hypothesis testing
7. What is the meaning of critical value in test of significance?

Q.2. Answer the following in short. (Any 4) (20)

1. If two dice are thrown, what is the probability that the sum is greater than 8.
2. Let X be the random variable with following probability distribution.
Find **$E(X)$ and $E(X^2)$**

x	2	4	8
P(X = x)	0.25	0.35	0.4

3. Calculate correlation coefficient between X and Y for the following data

X	65	66	67	67	68	69	70	72
Y	67	68	65	68	72	72	69	71

4. Calculate mean of the following frequency distribution

Class Interval	0-8	8-16	16-24	24-32	32-40	40-48
Frequency	8	7	16	24	15	7

5. Find the median wage of the following distribution

Wages in Rs	20-30	30-40	40-50	50-60	60-70
Number of laborers	3	5	20	10	5

6. What are the steps involved in hypothesis testing.

Q. 3. Answer the following in detail. (Any 3)

(30)

1. In a bolt factory machines A, B and C manufactures respectively 25%, 35% and 40% of the total. Of their output 5, 4, 2 percent are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by machines A, B and C?
2. Define Markov inequality, Chebyshev's inequality. Suppose it is known that the number of items produced in a factory during a week is a random variable with mean 50.
 - (a) What is the probability that this week's production will exceed 75?
 - (b) If the variance of a week's production is known to equal 25, then what is the probability that this week's production will be between 40 and 60?
3. In a partially destroyed laboratory record of an analysis of correlation data, the following results only are legible

$$\text{Variance on } X = 9$$

$$\text{Regression equations are } 8X - 10Y + 66 = 0 \text{ and } 40X - 18Y = 214$$

What were (i) the mean values of X and Y (ii) two regression coefficients (iii) Standard deviation of Y

4. Calculate the standard deviation and variance for the following data

Class Interval	20-30	30-40	40-50	50-60	60-70	70-80	80-90
Frequency	3	61	132	153	140	51	2

5. The means of two single large samples of 1000 and 2000 members are 67.5 inches and 68.0 inches respectively. Can the samples be regarded as drawn from the same population of standard deviation 2.5 inches? (Test at 5% level of significance). Take critical value of Z as 1.96
