National University of Computer and Emerging Sciences

Islamabad Campus

Probability and Statistics (MT2005)

BSCS(All Sections)

Course Instructors

Muhammad Ali, Sara Aziz, Sehrish Hassan Shigri,
Tajammul Hussain

Sessional-I Exam

Total Time (Hrs): 1

Total Marks: 75

Total Questions: 5

Date: February 25, 2025

Student Signature

DO NOT WRITE BELOW THIS LINE

Attempt all the questions

CLO # 1: Understand concepts of Statistical methods for data analysis.

Question # 1 [25 marks]

(a) A software development team tracks the number of bugs reported per sprint. The average number of bugs per sprint is 20, with a standard deviation of 4 bugs. Using Chebyshev's theorem, determine the range in which at least 75% of the sprint bug counts will fall. (3)

(b) In a study analyzing the response times (in milliseconds) of two different database query optimization techniques, the following datasets represent the measured response times from seven test cases:

Dataset for Optimization Technique A:

$$X = \{12, 17, 21, 26\}$$

Dataset for Optimization Technique B:

$$Y = \{14, 19, 23, 27\}$$

i. Show that the Mean Absolute Deviation (MAD) $\leq$ Standard Deviation (SD) for dataset X. (4)

ii. Compute the correlation coefficient between the response times of the two optimization techniques and interpret the result. (3)

(c) A network administrator records the data transmission times (in milliseconds) for various network packets. The times are:

$$\{25, 30, 42, 55, 60, 120\}$$

i. Find the five-number summary (Minimum, Q_1, Median, Q_3, Maximum). (5)

ii. Compute the Interquartile Range (IQR). (2)

iii. Identify any outliers using the $1.5 \times IQR$ rule. (2)

iv. Perform EDA on the given data by constructing a Box-and-Whisker Plot. (3)

v. Classify the data as left-skewed or right-skewed using Box-and-Whisker Plot. (1)

vi. If a new data point, 130 ms, is added, how does it affect the Box-and-Whisker Plot? (2)

CLO # 2: Apply combinatorial techniques, including permutations and combinations, to compute probabilities in discrete probability models

---

Question # 2 [10 marks]

(a) If a group of 12 people is surveyed, how many possibilities are there where 5 choose coffee, 4 choose tea, and 3 choose juice? (3)

(b) A system administrator needs to select 4 test cases from the 7 available cases to analyze in detail. However, the administrator must include at least one test case from the first three test cases. How many ways can the administrator choose the test cases if the order does not matter? (4)

(c) Determine the coefficient of $x^9y^5z^6w^3$ in the expansion of (3)

$$(x + 2y + 3z + w)^{23}.$$

CLO # 3: Apply basic principles of probability including sample space, joint probability, conditional probability, Bayes rule, total probability and random variables, in solving computational and real-world problems.

Question # 3 [15 marks]

(a) Let $(\Omega, \mathcal{F}, P)$ be a probability space, where Ω is the sample space, $\mathcal{F}$ is an event space, and P is a probability measure on $\mathcal{F}$. Under what conditions, $\mathcal{F}$ is a σ-algebra set on Ω. (3)

(b) Two events A and B are independent if

$$P(A \cap B) = P(A)P(B)$$

Show that if A and B are independent, then A^c and B^c are also independent. (4)

(c) Suppose a biased coin has probability 1/3 of landing heads. The coin has been tossed twice. Define the sample space and determine whether the events

$$A = \{\text{first toss is heads}\}, \quad B = \{\text{second toss is heads}\}$$

are independent. (3)

(d) Prove the Generalized Bayes' Theorem using the Law of Total Probability and the Definition of Conditional Probability. (5)

Hint: Start with the definition of conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

and use the Law of Total Probability:

0350

$$P(B) = \sum_{i=1}^{n} P(B \mid H_i)P(H_i),$$

where $\{H_1, H_2, \ldots, H_n\}$ form a partition of the sample space.

LO # 3: Apply basic principles of probability including sample space, joint robability, conditional probability, Bayes rule, total probability and random variables, in solving computational and real-world problems.

Question # 4 [16 marks]

(a) A certain disease affects 1% of a population. A test for this disease has:

- True positive rate (sensitivity): 90%
- False positive rate: 5%

If a randomly selected person tests positive, what is the probability that they actually have the disease? (6)

(b) Suppose a factory produces 60% of its products from Machine A and 40% from Machine B.

- Machine A produces 5% defective items.
- Machine B produces 10% defective items.

If a randomly selected product is defective, what is the probability that it came from Machine B? (6)

(c) Suppose 36% of families own a dog, 30% of families own a cat, and 22% of the families that have a dog also have a cat. A family is chosen at random and found to have a cat. What is the probability they also own a dog? (4)

CLO # 3: Apply basic principles of probability including sample space, joint probability, conditional probability, Bayes rule, total probability and random variables, in solving computational and real-world problems.

Question # 5 [9 marks]

(a) A software system experiences different types of requests, categorized as low, medium, and high priority. The priority level X of an incoming request follows the probability distribution:

X	-1	0	1
$f(X)$	3c	3c	6c

i. Determine the value of c. (3)

ii. Find the cumulative probability distribution for $X < 1$. (3)

(b) A box contains 8 light bulbs, 2 of which are defective. A technician randomly selects 3 bulbs for testing. If X is the number of defective bulbs selected, find the probability distribution of X. (3)