2/6 **1801E1198** Candidate's Seat No :_____

MSc AIML (Rep) Sem.-2 Examination

Statistical Foundation

Time : 3-00 Hours] **January-2025** **[Max. Marks : 100**

Instructions:

- Write both the Sections in the separate answer book.
- Both Sections have equal weightage

SECTION - I

Q.1 Attempt any 2 (20)

a) Out of 100 applicants, 40 had work experience and 30 had a certificate. 20 had work experience and a certificate. What is the probability that an applicant had work experience or a certificate?

b) Explain any sampling method to measure the heights of 50 students found in the gym during basketball tournament.

c) Prepare a frequency distribution table for the following data:

H	H	C	H	O	C	C	H	C	O
O	H	C	H	H	C	H	H	O	H

Q.2 Attempt any 2 (20)

a) The lifetime of a component follows a normal distribution with a mean of 2,000 hr and a standard deviation s of 200 hr. Find the probability that a component will last between 2,000 and 2,400 hr

b) The SAT scores of 100 students have a mean of 975 and a standard deviation of 105. The GPAs of the same 100 students have a mean of 3.16 and a standard deviation of 0.22. Which one has a larger coefficient of variation?

c) An oil company wants to determine the factors affecting consumer choice of service stations, and therefore has obtained the names and addresses of all the registered car owners residing in that area. Describe how a sample could be obtained using random and convenience sampling.

Q.3 Attempt any 2 (10)

a) Wealth(W) can be generated from Employment(E) and Investment(I). Health(H) and Wealth(W) are bring Joy(J) to a person. Wealth(W) can be donated to charitable(C) Institutions. Create a Bayesian network and identify any one trail

b) Explain the difference between normal and standard normal distribution.

c) A population has a mean of 230 and a standard deviation of 41. Using Chebyshev's theorem, find at least what percentage of the observations fall within 2 standard deviations of the mean

(P.T.O)

E 1198-2

SECTION - II

Q.4 Attempt any 2 (20)

a) Suppose we have a single training set T and a single validation set V. How can we use Binomial test to test the hypothesis that the classifier makes a misclassification error?

b) The data set lists the number of cities in 10 different countries

2	1	1	2	9	1	1	2	2	4

Calculate the mean, median, mode.

c) Suppose that two dice were rolled and the sum of the two numbers was odd. Determine the probability that the sum of the two numbers was less than 8.

Q.5 Attempt any 2 of the following: (20)

a) Suppose that 20% of all patients undergoing MRI testing require sedation. If 5 patients are selected, find the probability that exactly 2 patients require sedation.

b) A college wants to test if the mean number of hours spent working per week by students who hold jobs is different from 20 hours. Write the null and alternative hypotheses. Determine if it is a two-tailed, a left-tailed, or a right-tailed test.

c) Explain Null hypothesis and Alternative hypothesis

Q.6 Attempt any 2 (10)

a) The frequency distribution of 50 days is given. Calculate the mean and the variance.

Number of Orders	f
10–12	4
13–15	12
16–18	20
19–21	14

b) The time a student takes to complete an exam follows a uniform distribution and is between 30 minutes and 75 minutes. What is the probability that the student takes to complete the exam is between 42 and 63 minutes?

c) For a sample of 15 students, the following sales amounts arranged in ascending order are observed: 0.10, 0.10, 0.25, 0.25, 0.25, 0.35, 0.40, 0.53, 0.90, 1.25, 1.35, 2.45, 2.71, 3.09, 4.10.
Find range, Q1, Q3.