4 /41

1406E499

Candidate's Seat No :_____

M.Sc. AIML & AIML (DS) Semester 2 Examination

Statistical Foundation

June-2024

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

SECTION - I

Q. 1 Attempt any two (20)

a) Of 100 individuals who applied for systems analyst positions with a large firm during the past year, 40 had some prior work experience (W), and 30 had a professional certificate (C). However, 20 of the applicants had both work experience and a certificate, and thus are included in both of the counts.

i) What is the probability that a randomly chosen applicant had either work experience or a certificate (or both)?

ii) What is the probability that a random chosen applicant had either work experience or a certificate but not both?

iii) Determine the conditional probability that a randomly chosen applicant has a certificate given that he has some previous work experience.

iv) What is the probability that a randomly chosen applicant has neither work experience nor a certificate?

v) Show that W and C are dependent events?

b) Of 12 accounts, four contain a procedural error in posting account balances.

i) If an auditor randomly selects two accounts (without replacement), what is the probability that neither account will contain a procedural error?

ii) If the auditor samples three accounts, what is the probability that none of the accounts includes the procedural error?

iii) If an auditor samples one account randomly, what is the probability that it will contain the error?

iv) If an auditor samples two accounts randomly, what is the probability that at least one will contain the error?

v) If an auditor samples three accounts randomly, what is the probability that at least one will contain the error?

c) An analyst estimates that the probability is 0.30 that a new company plans to offer competitive services within the next three years, and 0.70 that the firm does not. If the new firm has such plans, a new manufacturing facility would definitely be built. If the new firm does not have such plans, there is still a 60 % chance that a new manufacturing facility would be built for other reasons.

i) What is the probability that that the new firm has in fact begun work on a new manufacturing facility.

ii) Suppose we observe that the new firm has in fact begun work on a new manufacturing facility. Given this information, what is the probability that the firm has decided to offer competitive telecommunications services?

(P.T.O)

EA99-2

Q. 2

a) An oil company wants to determine the factors affecting consumer choice of gasoline service stations in a test area, and therefore has obtained the names and addresses of and available personal information for all the registered car owners residing in that area. Describe how a sample of this list could be obtained using random, systematic, stratified, cluster and convenience sampling. (10)

b) The lifetime of an electrical component is known to follow a normal distribution with a mean of 2,000 hr and a standard deviation s of 200 hr. Find the probability that a randomly selected component will last between 2,000 and 2,400 hr (5)

OR

Q. 2

a) Because of economic conditions, a firm reports that 30 % of its accounts receivable from other business firms are overdue. If an accountant takes a random sample of five such accounts, determine the probability of each of the following events by use of the formula for binomial probabilities: (i) none of the accounts is overdue, (ii) exactly two accounts are overdue, (iii) exactly one of the accounts is overdue. (9)

b) What are the characteristics of normal distribution? (6)

Q. 3

a) A manufacturer contemplating the purchase of new toolmaking equipment has specified that, on average, the equipment should not require more than 10min of setup time per hour of operation. The purchasing agent visits a company where the equipment being considered is installed; from records there the agent notes that 40 randomly selected hours of operation included mean of 11.25 min of setup time, and the standard deviation of setup time per hour was 3.0 min. Based on this sample result, can the assumption that the equipment meets setup time specifications be rejected at the 1 percent level of significance? Explain the Type I error and Type II error that can occur in this case. (8)

Ho: $\mu \leq 10$, Ha: $\mu > 10$

b) 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? (7)

OR

Q. 3

a) Wealth(W) can be generated from Employment(E) and Investment(I). Health(H) and Wealth(W) are responsible for bringing Joy(J) to a person. Wealth(W) can be donated to charitable(C) Institutions (10)

i. Create a Bayesian network using the information given above
ii. Name any one causal trail
iii. Name any one common cause trail
iv. Name any one common effect trail
v. Is C independent of H? Show the active/inactive trail.

b) Explain the difference between normal and standard normal distribution. (5)

E 499-3

SECTION - II

Q. 4

a) What is the difference between discrete and continuous probability distribution? (10)

b) For two common stock issues in the electronics industry, the daily mean closing market price during a one-month period for stock A was $150 with a standard deviation of $5. For stock B, the mean price was $50 with a standard deviation of $3, which data set is more variable? (10)

Q. 5

a) The required times to complete a sample assembly task for 30 employees who have applied for a promotional transfer to a job requiring precision assembly. Prepare a frequency, relative frequency distribution for the data using a class interval of 2.0 for all classes and setting the lower stated limit of the first class at 9 min. Prepare a cumulative frequency distribution for the frequency distribution of assembly times (6)

10	14	15	13	17	16	12	14	11	13
15	18	9	14	14	9	15	11	13	11
12	10	17	16	12	11	16	12	14	15

b) Given table presents the joint probability table of voter reactions to a new property tax plan according to party affiliation. (9)

Party Affiliation	**Reaction**		
	In Favor (F)	**Neutral (N)**	**Opposed(O)**
Democratic (D)	0.30	0.05	0.05
Republican (R)	0.125	0.075	0.15
Independent(I)	0.125	0.025	0.10

Determine the following probabilities: (i) P(O), (ii) P(R and O), (iii) P(I), (iv) P(I and F), (v) P(OjR), (vi) P(R or D) (vii) P(D or F).

OR

Q. 5

For a sample of 15 students at an elementary-school snack bar, the following sales amounts arranged in ascending order of magnitude 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. (15)

i) Determine the mean, median, and mode for these sales amounts.
ii) How would you describe the distribution from the standpoint of skewness?
iii) Suppose that you are asked to determine the typical purchase amount only for this particular group of students. Which measure of average would you report? Why?
iv) Find range, Q1, Q3 and IQR
v) Draw the boxplot.
vi) Calculate the variance and standard deviation.

(P.T.O)

E499-4

Q. 6

a) Table given below presents voter reactions to a new property tax plan according to party affiliation. Test the null hypothesis that there is no relationship between party affiliation and voter reaction, using the 1 percent level of significance. (8)

Party Affiliation	Reaction		
	In Favor (F)	Neutral (N)	Opposed(O)
Democratic (D)	120	20	20
Republican (R)	50	30	60
Independent(I)	50	10	40

b) Why is it necessary to square the difference from the mean when computing the population variance? (7)

OR

Q. 6

a) 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? (5)

b) Explain ANOVA (10)