07/02/2024 - FN

Reg. No. : | | | | | | | | | | | | |

Question Paper Code : 21286

B.E./B.Tech. DEGREE EXAMINATIONS, NOVEMBER/DECEMBER 2023

Third/Fourth Semester

Environmental Engineering

MA 3391 – PROBABILITY AND STATISTICS

(Common to : Artificial Intelligence and Data Science / Biotechnology and Biochemical Engineering / Computer Science and Business Systems and Plastic Technology)

(Regulations 2021)

Time : Three hours | Maximum : 100 marks

All Statistical tables are permitted and Graph sheet should be provided

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. A fair coin was tossed two times. Given that the first toss resulted in heads, what is the probability that both tosses resulted in heads?

2. A bag contains six blue balls and four red balls. Balls are randomly drawn from the bag, one at a time, until a red ball is obtained. If we assume that each drawn ball is replaced before the next one is drawn, what is the probability that the experiment stops after exactly five balls have been drawn?

3. A discrete random variable X has the probability distribution given below

x	0.5	1	1.5	2
$P(x)$	k	$2k$	$3k$	k

Find the value of K.

4. If X_1 has mean 4 and variance 9 while X_2 has mean –2 and variance 6, and the two are independent, find $E(2X_1 + X_2 - 5)$.

5. Suppose that Θ_1 and Θ_2 are estimators of the parameter θ. We know that $E(\Theta_1) = \theta$, $E(\Theta_2) = \frac{\theta}{2}$, $V(\Theta_1) = 10$, $V(\Theta_1) = 4$. Which estimator is best? In what sense it is best?

6. As a baseline for a study on the effects of changing electrical pricing for electricity during peak hours, July usage during peak hours was obtained for $n1=45$ homes with air-conditioning and $n2=55$ homes without. The July on-peak usage (kWh) is summarized as:

Sample Population	Size	Mean	Variance
With	45	204.4	13,825.3
Without	55	130.0	8,632.0

Obtain a 95% confidence interval for $\delta = \mu 1 - \mu 2$.

7. What are the advantages of nonparametric test?

8. Can we apply the Kolmogorov test for discrete distribution in nonparametric tests? – Justify.

9. Which distributions are followed while calculating the control chart for p-chart and c-chart?

10. What are lower control limits and upper control limits of c-chart?

PART B — (5 × 16 = 80 marks)

11. (a) (i) Only 1 in 1000 adults is affected with a rare disease for which a diagnostic test has been developed. The test is such that when an individual actually has the disease, a positive result will occur 99% of the time, whereas an individual without the disease will show a positive test result only 2% of the time. If a randomly selected individual is tested and the result is positive, what is the probability that the individual has the disease? (8)

(ii) Let X have the Poisson distribution with probability distribution $f(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, for $x = 0,1,....\infty$. Show that (8)

(1) $M(t) = e^{\lambda}(e^t - 1)$ for all t.

(2) $E(X) = \lambda$ and $Var(X) = \lambda$

Or

(b) (i) Sport stories and financial reports, written by algorithms based on artificial intelligence, have become common place, One company fed its algorithm with box scores and play-by-play information and created over one million on-line reports of little league games in 2011. They now write many stories on the Big Ten Network site. Suppose that the algorithm, or robot reporter, typically writes proportion 0.65 of the stories on the site. If 15 new stories are scheduled to appear on a web site next weekend, find the probability that

(1) 11 will be written by the algorithm

(2) at least 10 will be written by the algorithm

(3) between 8 and 11 inclusive will be written by the algorithm (8)

(ii) Find the probabilities that a random variable having the standard normal distribution will take on a value

(1) between 0.87 and 1.28

(2) between–0.34 and 0.62

(3) greater than 0.85

(4) greater than –0.65. (8)

12. (a) Let X and Y be discrete random variable with probability function

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

Where k is a constant.

(i) What is the value of k?

(ii) Find the marginal PMFs of X and Y

(iii) Are X and Y independent? (16)

Or

(b) The joint PDF of the random variables X and Y are defined as follows :

$$f(x,y) = \begin{cases} 25e^{-5y}, & 0 < x < 0.2,\ \ y > 0 \\ 0, & otherwise \end{cases}$$

(i) Find the marginal PDFs of X and Y.

(ii) What is the covariance of X and Y? (16)

13. (a) (i) Let $X_1, \ldots, X_n$ be a random sample of size n from an exponential distribution. Obtain the maximum likelihood estimators of λ. (6)

(ii) Suppose that X is a random variable with mean μ and variance σ^2, Let be a random sample of size n from the population represented by X. Show that the sample mean $\overline{X}$ and sample variance S^2 are unbiased estimators of μ and σ^2 respectively. (10)

Or

(b) (i) To reduce the amount of recycled construction materials entering landfills it is crushed for use in the base of roadways. Green engineering practices require that their strength, resiliency modulus (MPa), be accessed. Measurements on $n_1 = n_2 = 6$ specimens of recycled materials from two different locations produce the data (Courtesy of Tuncer Edil) (8)

Location 1 :	707	632	604	652	669	674
Location 2 :	552	554	484	630	648	610

Use the 0.05 level of significance to establish a difference in mean strength for materials from the two locations.

(ii) It is desired to determine whether there is less variability in the silver plating done by Company 1 than in that done by Company 2. If independent random samples of size 12 of the two companies' work yield $s_1 = 0.035$ mil and $s2 = 0.062$ mil, test the null hypothesis $\sigma_1^2 = \sigma_2^2$ against the alternative hypothesis $\sigma_1^2 < \sigma_2^2$ at the 0.05 level of significance. (8)

14. (a) (i) The following data constitute a random sample of 15 measurements of the octane rating of a certain kind of gasoline.

99.0 102.3 99.8 100.5 99.7 96.2 99.1 102.5 103.3

97.4 100.4 98.9 98.3 98.0 101.6

Test the null hypothesis $\mu = 98.0$ against the alternative hypothesis $\mu > 98.0$ at the 0.01 level of significance. Conducting a sign test to test the hypothesis. (8)

(ii) Suppose that in a study of sedimentary rocks, the following diameters (in millimeters) were obtained for two kinds of sand.

Sand I: 0.63 0.47 0.17 1.36 0.35 0.51 0.49 0.45 0.18 0.84 0.43
0.32 0.12 0.40 0.20

Sand II: 1.13 1.01 0.54 0.48 0.96 0.89 0.26 1.07 0.39 1.11
0.88 0.58 0.92 0.53

Use the U test at the 0.01 level of significance to test the null hypothesis that the two samples come from identical populations against the alternative hypothesis that the populations are not identical. (8)

Or

(b) (i) An engineer is concerned about the possibility that too many changes are being made in the settings of an automatic lathe, Given the following mean diameters (in inches) of 40 successive shafts turned on the lathe.

0.261	0.258	0.249	0.251	0.247	0.256	0.250	0.247	0.255	0.243
0.252	0.250	0.253	0.247	0.251	0.243	0.258	0.251	0.245	0.250
0.248	0.252	0.254	0.250	0.247	0.253	0.251	0.246	0.249	0.252
0.247	0.250	0.253	0.247	0.249	0.253	0.246	0.251	0.249	0.253

use the 0.01 level of significance to test the null hypotheses of randomness against the alternative that there is a frequently alternating pattern. (8)

(ii) It is desired to check whether pinholes in electrolytic tin plate are uniformly distributed across a plated coil on the basis of the following distances in inches of 10 pinholes from one edge of a long strip of tin plate 30 inches wide:

4.8 14.8 28.2 23.1 4.4 28.7 19.5 2.4 25.0 6.2

Test the null hypothesis at the 0.05 level of significance using the Kolmogorov– Smirnov test for uniformity. (8)

15. (a) A machine fills boxes with dry cereal. 15 samples of 4 boxes are drawn randomly. The weights of the sampled boxes are shown as follows. Draw the control charts for the sample mean and sample range and determine whether the process is in a state of control.

Sample Number	1	2	3	4	5	6	7
Weights of boxes (x)	10.0	10.3	11.5	11.0	11.3	10.7	11.3
	10.2	10.9	10.7	11.1	11.6	11.4	11.4
	11.3	10.7	11.4	10.7	11.9	10.7	11.1
	12.4	11.7	12.4	11.4	12.1	11.0	10.3

Sample Number	8	9	10	11	12	13	14	15
Weights of boxes (x)	12.3	11.0	11.3	12.5	11.9	12.1	11.9	10.6
	12.1	13.1	12.1	11.9	12.1	11.1	12.1	11.9
	12.7	13.1	10.7	11.8	11.6	12.1	13.1	11.7
	10.7	12.4	11.5	11.3	11.4	11.7	12.0	12.1

Or

(b) 10 samples each of size 50 were inspected and the number of defectives in the inspection were 2, 1, 1, 2, 3, 5, 5, 1, 2, 3. Construct a p-chart and np-chart for each sample. Assuming special causes for out-of control points, find the revised control limits.
