A

02000MAT256062205

Pages: 7

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth Semester B.Tech Degree (S, FE) Examination January 2024 (2019 Scheme)

Course Code: MAT256

Course Name: PROBABILITY AND STATISTICAL MODELLING

Max. Marks: 100 Duration: 3 Hours

PART A

(Answer all questions; each question carries 3 marks) Marks

1 Consider whether the next person buying a computer at a certain electronics store buys a laptop or desktop model. Let 3

$$X = \begin{cases} 0, & \textit{if the customer buys a desktop} \\ 1, & \textit{if the customer buys a laptop} \end{cases}$$

If 20% of all purchasers during that week select a desktop, find the pmf of X.

2 A random variable X takes values 0, 1, 2, 3 with probabilities $\frac{8}{15}, \frac{1}{3}, \frac{1}{15}, \frac{1}{15}$. Find the mean and the variance. 3

3 A college professor never finishes his lectures before the end of the hour and always finishes his lectures within two minutes after the hour. Let X be the time that lapses between the end of the hour and the end of the lecture. Suppose that the pdf of X is 3

$$f(x) = \begin{cases} kx^2, & 0 \le x \le 2 \\ 0, & \textit{otherwise} \end{cases}$$

Find the value of k.

4 The random variable X is exponentially distributed with mean 3. Find $P(X > t+3 \mid X > t)$, where t is any positive real number. 3

5 Describe the difference between confidence level and significance level. 3

6 Briefly describe sampling theory. 3

7 Let the test statistic T have a t-distribution when H_0:$\mu=\mu_0$ is true. Give the significance level for the following data: 3

$H_a : \mu \neq \mu_0$, $n = 31$, rejection region : $t \geq 1.697$ or $t \leq -1.697$

8 A trucking firm is suspicious that the average lifetime of a certain tyre is at least 28,000 miles. To check the claim, the firm puts 40 of these tyres on its trucks and gets a mean lifetime of 27,463 miles with a standard deviation of 1,348 miles. What can we conclude if the probability of type I error is to be at most 0.01? 3

9 Write three properties of sample correlation coefficient. 3

10 A study to assess the capability of subsurface flow wetland systems to remove Biochemical Oxygen Demand (BOD) and various other chemical constituents resulted in the accompanying data on x = BOD mass loading (kg/ha/d) and y = BOD mass removal (kg/ha/d): 3

x	3	8	10	11	13	16	27	30	35	37	38	44	103	142
y	4	7	8	8	10	11	16	26	21	9	31	30	75	90

Values of the relevant summary quantities are:

$$n = 14, \sum x_i = 517, \sum y_i = 346, \sum x_i^2 = 39095, \sum y_i^2 = 17454,$$

$$\sum x_i y_i = 25825$$

Obtain the equation of the least squares line.

PART B

(Answer one full question from each module, each question carries 14 marks)

Module -1

11 a) In a city, 4% of all licensed drivers will be involved in at least one road accident in any given year. Use Poisson distribution to determine the probability that among 150 drivers, randomly chosen in this city, 7

1) Only five will be involved in at least one road accident in any given year.
2) At most three will be involved in at least one road accident in any given year.

b) Let X be a discrete random variable with mean 10 and variance 25. Find the positive values of α and β such that $Y = \alpha X - \beta$ has mean 0 and variance 1. 7

12 a) The joint probability distribution of X and Y is given by $f(x, y) = \frac{1}{27}(2x + y)$; $x = 0,1,2$ and $y = 0,1,2$. 7

1) Find the marginal distributions of X and Y
2) Are X and Y independent random variables?

b) Suppose that 20% of all copies of a particular textbook fail a certain binding strength test. Let X denote the number among 15 randomly chosen copies that fail the test. Using the table of Binomial distributions or by direct calculation. 7

1) Find the probability that at most 8 will fail the test.
2) Find the probability that exactly 8 will fail the test.
3) Find the probability that at least 8 will fail the test.
4) Find the probability that failure is between 4 and 7 (inclusive).

Module -2

13 a) A random variable X follows uniform distribution in (-3, 3). Find 7

1) $P(|X| < 2)$
2) $P(|X - 2| < 2)$
3) The value of k for which $P(X > k) = \frac{1}{3}$

b) The lifetime of a certain brand of tube light may be considered as a random variable with mean 1200 hours and standard deviation 250 hours. Using Central Limit Theorem, find the probability that the average lifetime of 60 bulbs exceeds 1250 hours. 7

14 a) Annie and Alvin have agreed to meet between 5 PM and 6 PM for dinner at a local health food restaurant. Let X and Y be Annie's and Alvin's arrival times respectively. Suppose that X and Y are independent with each uniformly distributed on the interval [5,6]. 7

1) What is the joint pdf of X and Y?
2) If the first one to arrive will wait only 10 minutes before leaving to eat elsewhere, what is the probability that they have dinner at the health food restaurant?

b) Let X be a continuous random variable with cdf 7

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{x}{4}(1 + \ln\frac{4}{x}), & 0 < x \leq 4 \\ 1, & x > 4 \end{cases}$$

1) Find $P(X \leq 1)$
2) Find $P(1 \leq X \leq 3)$
3) Find the probability density function of X

Module -3

15 a) A random sample of 700 units from a large consignment showed that 200 were damaged. Find (i) 95% and (ii) 99% confidence limits for the proportion of damaged units in the consignment. 7

b) Determine the size of the sample for estimating the true weight of the cereal containers for the universe with $N = 5000$ based on the following information: 7

1) The variance of weight is 4 ounces on the basis of past records.
2) The estimate should be within 8 ounces of the true average weight with 99% probability.

Will there be a change in the size of the sample if we assume infinite population in the given case? If so, explain by how much?

16 a) A random sample of 100 items taken from a large batch of articles contains 5 defective items. 7

1) Set up 96% confidence limits for the proportion of defective items in the batch.
2) If the batch contains 2,669 items, set up 95% confidence interval for the proportion of defective items.

b) A survey is proposed to be conducted to know the annual earnings of the old statistics graduates of Delhi University. How large should the sample be in order to estimate the mean monthly earning within $\pm$ Rs. 10,000 at 95% confidence level? The standard deviation of the annual earning of the entire population is known to be Rs. 30,000. 7

Module -4

17 a) A specimen of copper wires drawn from a large lot has the following breaking strengths- 578, 572, 570, 568, 572, 578, 570, 572, 596, 544. Test (Using Student's t-statistic or Sandler's A-statistic) whether the mean breaking strength of the lot may be taken to be 578 kg weight. Verify the inference. 7

b) The article "Origin of precambrian Iron formation" reports the following data on total Fe for four types of Iron formation (1=Carbonate, 2=Silicate, 3=Magnetite, 4=Hermatite): 7

1	20.5	28.1	27.8	27.0	28.0
	25.2	25.3	27.1	20.5	31.3
2	26.3	24.0	26.2	20.2	23.7
	34.0	17.1	26.8	23.7	24.9
3	29.5	34.0	27.5	29.4	27.9
	26.2	29.9	29.5	30.0	35.6
4	36.5	44.2	34.1	30.3	31.4
	33.1	34.1	32.9	36.3	25.5

Carry out an analysis of variance F test at significance level 0.01, and summarize the results in an ANOVA table.

18 a) A random sample of 150 recent donations at a certain blood bank reveals that 82 were type A blood. Does this suggest that the actual percentage of type A donations differs from 40%, the percentage of the population having type A blood? Carry out a test of the appropriate hypotheses using a significance level of 0.01. 7

b) 6 samples each of four types of cereal grown in a certain region were analyzed to determine the Thiamine content, resulting in the following data: 7

Wheat	5.2	4.5	6.0	6.1	6.7	5.8
Barley	6.5	8.0	6.1	7.5	5.9	5.6
Maize	5.8	4.7	6.4	4.9	6.0	5.2
Oat	8.3	6.1	7.8	7.0	5.5	7.2

Does this data suggest that at least two of the grains differ with respect to the average Thiamine content? Use a level $\alpha = 0.05$ test based on the $P-$ value method.

Module -5

19 a) The flow rate y (m^3 / min) in a device used for air-quality measurement depends on the pressure drop x (in. of water) across the device's filter. Suppose that for x values between 5 and 20, the two variables are related according to the simple linear regression model with true regression line $y = -0.12 + 0.095x$. 7

1) What is the expected change in flow rate associated with a 1 in. increase in pressure drop?
2) What is the expected flow rate for a pressure drop of 10 in.?
3) Suppose $\sigma = 0.025$ and consider a pressure drop of 10 in. What is the probability that the observed value of flow rate will exceed 0.835?

b) The following table gives a correlation analysis to investigate the relationship between maximal lactate level x and muscular endurance y. 7

x	400	750	770	800	850	1025	1200
y	3.80	4.00	4.90	5.20	4.00	3.50	6.30

x	1250	1300	1400	1475	1480	1505	2200
y	6.88	7.55	4.95	7.80	4.45	6.60	8.90

S_{XX}= *36.9839*, S_{YY}= *2,628,930.357*, S_{XY}= *7,377.704*. A scatter plot shows a linear pattern.

Test to see whether there is a positive correlation between maximal lactate level and muscular endurance in the population from which this data was selected.

20 a) Based on the following data, it is suggested that the thermal conductivity (y) of Polyethylene is a linear function of $10^4 \frac{1}{x}$, where x is the lamellar thickness (in Å): 7

x	240	410	460	520	590	745	8300
y	12.0	14.7	15.2	15.2	15.6	16.0	18.0

1) Estimate the parameters of the regression function and the regression function itself.

2) Predict the value of thermal conductivity when the lamellar thickness is 500Å

b) Physical properties of six flame-retardant fabric samples were investigated in the article "Sensory and Physical Properties of Inherently Flame-Retardant Fabrics" (Textile Research, 1984: 61–68). Use the accompanying data and a 0.05 significance level to determine whether a linear relationship exists between stiffness x (mg-cm) and thickness y (mm). Is the result of the test surprising in light of the value of r? 7

x	7.98	24.52	12.47	6.92	24.11	35.71
y	0.28	0.65	0.32	0.27	0.81	0.57

*****