A

02000MAT256062202

Pages: 5

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

B.Tech Degree S4 (S, FE) Examination December 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 Define a discrete random variable. What do you mean by its i) probability distribution ii) expectation? 3

2 A biased die is thrown thirty times and the number of sixes seen is eight. If the die thrown a further twelve times find the probability that a six will occur exactly twice. 3

3 If the distribution function of a random variable is given by 3

$F(x) = \begin{cases} 1 - \frac{4}{x^2}; & x > 2 \\ 0; x \leq 2 \end{cases}$ find the probabilities that this random variable will take on a value (1) less than 3; (2) between 4 and 5.

4 Examine the independence of random variables X and Y with $f(x, y) = 6(x - y); 0 < y < x < 1$.

5 Define sampling distribution with an example. 3

6 A random sample of size 16 has 53 as mean. The sum of the squares of the deviations taken from mean is 150. Obtain 95% confidence limits of the population mean. 3

7 A test of the breaking strengths of 6 ropes manufactured by a company showed a mean breaking strength of 7750 lb and a standard deviation of 145 lb, whereas the manufacturer claimed a mean breaking strength of 8000 lb. Can we support the manufacturer's claim at a level of significance of 0.05. 3

8 A commonly prescribed drug for relieving nervous tension is believed to be only 60% effective. Experimental results with a new drug administered to a random sample of 100 adults who were suffering from nervous tension show that 70 received relief. Is this sufficient evidence to conclude that the new drug is superior to the one commonly prescribed? Use a 0.05 level of significance. 3

9 Fit a straight line using the method of least squares for the following data 3

x	1	3	5	7	8	10
y	8	12	15	17	18	20

10 Write down the most important properties of correlation coefficient. 3

PART B

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

Module -1

11 a) For health reasons, homes need to be inspected for radon gas which decays and produces alpha particles. One device counts the number of alpha particles that hit its detector. To a good approximation, in one area, the count for the next week follows a Poisson distribution with mean 1.3. Determine 7
1) the probability of exactly one particle next week.
2) the probability of one or more particles next week.
3) the probability of at least two but no more than four particles next week.
4) the variance of the Poisson distribution.

b) Prove that The Poisson distribution is a limiting case of the binomial distribution 7

12 a) $f(xy) = \frac{2x+3y}{72}; x = 0,1,2; y = 1,2,3$ is the joint pdf of (X,Y). 7

1) Find the distribution of X+Y.
2) Examine whether X and Y are independent.

b) Given the following probability distribution 7

1) Find c
2) Find $P(x \geq 6)$, $P(x < 3)$ and $P\ (0 < X < 5)$ 3) If $P(x \leq k) > \frac{1}{2}$ find the minimum value of k.

x	0	1	2	3	4	5	6	7
P(x)	0	c	2c	2c	3c	c^2	$2c^2$	$7c^2+c$

Module -2

13 a) The length of satisfactory service (years) provided by a certain model of laptop computer is a random variable having the probability density 7

$$f(x) = \begin{cases} \frac{1}{4.5} e^{\frac{-x}{4.5}}; & x > 0 \\ 0; & x \leq 0 \end{cases}$$

Find the probabilities that one of these laptops will provide satisfactory service for
1) at most 2.5 years;
2) anywhere from 4 to 6 years;
3) at least 6.75 years.

b) If the joint probability density of two random variables is given by $f(x_1x_2) = \begin{cases} 6e^{-2x_1-3x_2}; & x_1 > 0, x_2 > 0 \\ 0; & elsewhere \end{cases}$ find the probabilities that 7

1) The first random variable will take on a value between 1 and 2 and the second random variable will take on a value between 2 and 3;

2) The first random variable will take on a value less than 2 and the second random variable will take on a value greater than 2.

3) Find the joint cumulative distribution function.

4) Check whether the two random variables are independent.

14 a) Derive the mean and variance of uniform distribution. Find the distribution function of a random variable having a uniform distribution on (0, 1). 7

b) A lawyer commutes daily from his suburban home to his midtown office. The average time for a one-way trip is 24 minutes, with a standard deviation of 3.8 minutes. Assume the distribution of trip times to be normally distributed. 7

1) What is the probability that a trip will take at least 1/2 hour?

2) If the office opens at 9:00 A.M. and the lawyer leaves his house at 8:45 A.M. daily, what percentage of the time is he late for work?

3) If he leaves the house at 8:35 A.M. and coffee is served at the office from 8:50 A.M. until 9:00 A.M., what is the probability that he misses coffee?

4) Find the length of time above which we find the slowest 15% of the trips.

Module -3

15 a) Describe about Sandler's A test. 7

b) In a random selection of 64 of the 2400 intersections in a small city, the mean number of scooter accidents per year was 3.2 and the sample standard deviation was 0.8. 7

1) Make an estimate of the standard deviation of the population from the sample standard deviation.

2) Work out the standard error of mean for this finite population.

3) If the desired confidence level is 0.90, what will be the upper and lower limits of the confidence interval for the mean number of accidents per intersection per year?

16 a) What are the important points to be considered by the researcher while determining the sample size in a sampling analysis? 7

b) The foreman of a certain mining company has estimated the average quantity of ore extracted to be 34.6 tons per shift and the sample standard deviation to be 2.8 7

tons per shift, based upon a random selection of 6 shifts. Construct 95% as well as 98% confidence interval for the average quantity of ore extracted per shift.

Module -4

17 a) The specimen of copper wires drawn form a large lot have the following breaking strength (in kg. weight): 578, 572, 570, 568, 572, 578, 570, 572, 596, 544. Test (using Student's t-statistic) whether the mean breaking strength of the lot may be taken to be 578 kg. weight. Test at 5 per cent level of significance. 7

b) Set up an analysis of variance table for the following per acre production data for three varieties of wheat, each grown on 4 plots and state if the variety differences are significant at 5% level of significance. 7

Variety of wheat		
I	II	III
6	5	5
7	5	4
3	3	3
8	7	4

18 a) Newborn babies are more likely to be boys than girls. A random sample found 13,173 boys were born among 25,468 newborn children. The sample proportion of boys was 0.5172. Is this sample evidence that the birth of boys is more common than the birth of girls in the entire population? 7

b) An experiment was designed to study the performance of 4 different detergents for cleaning fuel injectors. The following "cleanness" readings were obtained with specially designed equipment for 12 tanks of gas distributed over 3 different models of engines. Looking at the detergents as treatments and the engines as blocks, obtain the appropriate analysis of variance table and test at the 0.01 level of significance whether there are differences in the detergents or in the engines. 7

	Engine I	Engine II	Engine III	Total
Detergent A	45	43	51	139
Detergent B	47	46	52	145
Detergent C	48	50	55	153
Detergent D	42	37	49	128
Total	182	176	207	565

Module -5

19 a) Heavy metals can inhibit the biological treatment of waste in municipal treatment plants. Monthly measurements were made at a state-of-the-art treatment plant of the amount of chromium ($\mu g/l$) in both the influent and effluent. 7

influent	250	290	270	100	300	410	110	130	1100
effluent	19	10	17	11	70	60	18	30	180

1) Make a scatter plot.

2) Make a scatter plot after taking the natural logarithm of both variables.

3) Calculate the correlation coefficient, r, in part (1) and part (2).

4) Comment on the appropriateness of r in each case.

b) Given the data 7

x	0	1	2	3	4	5	6	7	8	9
y	9.1	7.3	3.2	4.6	4.8	2.9	5.7	7.1	8.8	10.2

fit a regression curve of the form $\mu Y|x = \beta_0 + \beta_1 x + \beta_2 x^2$ and then estimate $\mu Y|2$.

20 a) The following data were obtained in a study of the relationship between the weight and chest size of infants at birth 7

Weight	2.75	2.15	4.41	5.52	3.21	4.32	2.31	4.30	3.71
Chest size	29.5	26.3	32.2	36.5	27.2	27.7	28.3	30.3	28.7

1) Calculate r.

2) Test the null hypothesis that $\rho = 0$ against the alternative that $\rho > 0$ at the 0.01 level of significance.

3) What percentage of the variation in infant chest sizes is explained by difference in weight?

b) An experiment was conducted to determine if the weight of an animal can be predicted after a given period of time on the basis of the initial weight of the animal and the amount of feed that was eaten. The following data, measured in kilograms, were recorded: 7

Final Wt. y	Initial Wt. x_1	Feed Wt. x_2
95	42	272
77	33	226
80	33	259
100	45	292
97	39	311
70	36	183
50	32	173
80	41	236
92	40	230
84	38	235

1)Estimate the multiple linear regression equation $\mu Y|x_1, x_2 = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.

***