A

02000MAT256062201

Pages: 5

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth Semester B.Tech Degree Supplementary Examination June 2023 (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 Find the Binomial distribution with mean 3 and variance 2. 3

2 Derive the mean of Poisson distribution. 3

3 Suppose that the current measurements in a strip of wire are assumed to follow a normal distribution with a mean of 12 milliamperes and a variance of 9 milliamperes. What is the probability that a measurement exceeds 15 milliamperes? 3

4 If $f(x, y) = 1.5(2 - 2x - y)$; is defined in the region bounded by $y=0$, $x=0$ and $y=2-2x$ and *0 elsewhere*. Find the marginal distribution of X for $0<x<1$.

5 The amount of a particular impurity in a batch of a certain chemical product is a random variable with mean value 4.0 g and standard deviation 1.5 g. If 50 batches are independently prepared, what is the (approximate) probability that the sample average amount of impurity X is between 3.5 and 3.8 g? 3

6 A random sample of 500 oranges was taken from a large consignment and 65 were found to be bad. Find the standard error of the proportion of bad ones and probable limit of bad oranges in the consignment. 3

7 The manufacturer of a patent medicine claimed that it was 90% effective in relieving an allergy for a period of 8 hours. In a sample of 200 people who had the allergy, the medicine provided relief for 160 people. Find the P value of the test. 3

8 A laboratory technician is timed 20 times in the performance of a task, getting $\bar{x} = 6.9$ and s = 1.2 minutes. If the probability of a Type I error is to be at most 0.05, does this constitute evidence against the null hypothesis that the average time is less than or equal to 7.5 minutes? 3

9 Explain the least square method for fitting a straight line for a given data. 3

10 Calculate the correlation coefficient between the heights of fathers in inches (X) and their sons (Y). 3

X	65	66	57	67	68	69	70	72
Y	67	56	65	68	72	72	69	71

PART B

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

Module -1

11 a) Derive the mean and variance of binomial distribution. 7

b) A competitive game of skill is played by teams of two players, each of whom can score 0,1,2 and 3 points. The score of the teams is the sum of the points gained by the two players. For one of the teams the probability distribution of the scores of the players are as follows. 7

x	0	1	2	3
f(x)	0.1	0.3	0.4	0.2

y	0	1	2	3
g(y)	0.3	0.1	0.4	0.2

Assuming the independence of X and Y derive

1) The joint pdf of X and Y.
2) The pdf of Z=X+Y.

12 a) Sport stories and financial reports, written by algorithms based on artificial intelligence, have become common place. 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 7
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.

b) National security requires that defense technology be able to detect incoming projectiles or missiles. To make the defense system successful, multiple radar screens are required. Suppose that three independent screens are to be operated and the probability that any one screen will detect an incoming missile is 0.8. Obviously, if no screens detect an incoming projectile, the system is unworthy and must be improved. 7
1) What is the probability that an incoming missile will not be detected by any of the three screens?
2) What is the probability that the missile will be detected by only one screen?
3) What is the probability that it will be detected by at least two out of three screens

Module -2

13 a) Derive the mean and variance of exponential distribution. 7

b) A soft-drink machine is regulated so that it discharges an average of 200 milliliters per cup. If the amount of drink is normally distributed with a standard deviation equal to 15 milliliters, 7
1) what fraction of the cups will contain more than 224 milliliters?

2) what is the probability that a cup contains between 191 and 209 milliliters?
3) how many cups will probably overflow if 230-milliliter cups are used for the next 1000 drinks?
4) below what value do we get the smallest 25% of the drinks?

14 a) $f(x_1x_2) = \begin{cases} kx_1^2x_2^2; & 0 < x_1 < x_2 < 1 \\ 0; & otherwise \end{cases}$ 7

1) Determine k
2) Obtain the marginal distribution of X_1
3) Test the independence of X_1 and X_2

b) A continuous random variable X has the following density function 7

$$f(x) = \begin{cases} ax & 0 \le x \le 1 \\ a & 1 \le x \le 2 \\ -ax + 3a & 2 \le x \le 3. \end{cases}$$

1) Determine the constant a.
2) Determine the distribution function.
3) Sketch the graph of f(x) and F(x).

Module -3

15 a) Determine the size of the sample for estimating the true weight of the cereal containers for the universe with $N = 5000$ on the basis of the following information: 7
1) the variance of weight = 4 ounces on the basis of past records.
2) estimate should be within 0.8 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?

b) Give the definitions of the fundamental terms *Population, Sampling frame, Statisitc(s) and parameter, Sampling error, Precision and significance level* in sampling concepts. 7

16 a) Describe the steps involved in the determination of sample size through the approach based on bayesian statistics. 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 .90, what will be the upper and lower limits of the confidence interval for the mean number of accidents per intersection per year?

Module -4

17 a) 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. 7

b) Four different locations in the northeast were used for collecting ozone measurements in parts per million. Amounts of ozone were collected in 5 samples at each location. Is there sufficient information here to suggest that there are differences in the mean ozone levels across locations at 5% level of significance? 7

Location			
I	II	III	IV
0.09	0.15	0.10	0.10
0.10	0.12	0.13	0.07
0.08	0.17	0.08	0.05
0.08	0.18	0.08	0.08
0.11	0.14	0.09	0.09

18 a) The Edison Electric Institute has published figures on the number of kilowatt hours used annually by various home appliances. It is claimed that a vacuum cleaner uses an average of 46 kilowatt hours per year. If a random sample of 12 homes included in a planned study indicates that vacuum cleaners use an average of 42 kilowatt hours per year with a standard deviation of 11.9 kilowatt hours, does this suggest at the 0.05 level of significance that vacuum cleaners use, on average, less than 46 kilowatt hours annually? Assume the population of kilowatt hours to be normal. 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) A study was made by a retail merchant to determine the relation between weekly advertising expenditures and sales. 7

1) Plot a scatter diagram.

2) Find the equation of the regression line to predict weekly sales from advertising expenditures.

3) Estimate the weekly sales when costs are $35

Cost ($)	40	20	25	20	30	50	40	20	50	40	25	50
Sales ($)	385	400	395	365	475	440	490	420	560	525	480	510

b) Compute and interpret the correlation coefficient for the following grades of 6 students selected at random. Test the hypothesis that $\rho=0$ against the alternative that $\rho\neq0$. Use a 0.05 level of significance. 7

Maths grade	70	92	80	74	65	83
English grade	74	84	63	87	78	90

20 a) Observations on the yield of a chemical reaction taken at various temperatures were recorded as follows:. 7

x (^0C)	150	150	200	250	250	300	150	200	200	250	300	300
y(%)	75.4	81.2	85.5	89	90.5	96.7	77.7	84.4	85.7	89.4	94.8	95.3

1) Plot the data.
2) Does it appear from the plot as if the relationship is linear?
3) Fit a simple linear regression and test for lack of fit.

b) A set of experimental runs was made to determine a way of predicting cooking time y at various values of oven width x_1 and flue temperature x_2. The coded data were recorded as follows: 7

y	X_1	X_2
6.40	1.32	1.15
15.05	2.69	3.40
18.75	3.56	4.10
30.25	4.41	8.75
44.85	5.35	14.82
48.94	6.20	15.15
51.55	7.12	15.32
61.50	8.87	18.18
100.44	9.80	35.19
111.42	10.65	40.40

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

***