MCA DEGREE EXAMINATIONS: DEC 2022

(Regulation 2020)

First Semester

MASTER OF COMPUTER APPLICATIONS

P20MAT1101: Probability and Statistics for Data Analysis
(Statistical tables Required)

COURSE OUTCOMES

CO1: Understand about data collection, represent data graphically using bar chart and pie chart. and compute various measures of central tendency and dispersion for analysis of data

CO2: Interpret the correlation between variables and predict unknown values using regression

CO3: Explore random variables and predict probabilities for situations following normal distribution.

CO4: Perform hypothesis testing using large sample tests and Chi square test and interpret the results, which will form the basis for data analysis

CO5: Understand the principles of design of experiments and perform analysis of variance.

Time: Three Hours   **Maximum Marks: 100**

Answer all the Questions:-

PART A (10 x 2 = 20 Marks)

1.	There are two branches of a company employing 280 and 320 persons respectively. If the A.M. of the salaries of the two branches are Rs. 750 and Rs.937.5 respectively, find the A.M. of the salaries of the employees of the company as a whole.	CO1	[K₃]
2.	The mean of 50 items is 25 and their standard deviation is 2. Find the sum and the sum of the squares of all the items.	CO1	[K₂]
3.	The equations of the two lines of regression for a bivariate data are $Y = 10(X – 5)$ and $X = 2.5(Y – 14)$. Find the arithmetic means of X and Y.	CO2	[K₂]
4.	Write the relationship between correlation coefficient and regression coefficient.	CO2	[K₂]
5.	The diameter of an electric cable X is a continuous R.V with p.d.f $f(x) = kx(1 - x)$, $0 \leq x \leq 1$. Find the value of k.	CO3	[K₁]
6.	State any two properties of Normal distribution.	CO3	[K₂]
7.	What do you mean by Type I error and Type II error?	CO4	[K₂]
8.	A survey of 150 randomly selected bricks showed that 1/5 of them were defective. Construct a 0.95 confidence interval estimate of the proportion p of defective bricks.	CO4	[K₃]
9.	Write a note on Basic principles of Experimental Design.	CO5	[K₂]
10.	Compare RBD and CRD.	CO5	[K₄]

PART B (6 x 5 = 30 Marks)

11. Calculate S.D for the following: CO1   [K₂]

Size:	6	7	8	9	10	11	12
Freq:	3	6	9	13	8	5	4

12. From the following data calculate the rank correlation coefficient   CO2   [K₂]

X:	48	33	40	9	16	16	65	54	16	57
Y:	13	13	24	6	15	4	20	9	6	19

13. A random variable X has the following probability distribution.   CO3   [K₁]

X	-2	-1	0	1	2	3
P(x)	0.1	k	0.2	2k	0.3	3k

     a) Find k   b) Evaluate P(X < 2) and P(-2 < X < 2)   c) Find the cdf of X.

14. The mean lifetime of a sample of 100 light bulbs produced by a company is found to be 1580 hours with standard deviation of 90 hours. Test the hypothesis that the mean lifetime of the tubes produced by the company is 1600 hours.   CO4   [K₃]

15. 1,000 students at college level were graded accounting to their I.Q. and the economic conditions of their homes. Use Chi square test to find out whether there is any association between economic condition at home and I.Q.   CO4   [K₃]

I.Q Economic Condition	High	Low	Total
Rich	460	140	600
Poor	240	160	400
Total	700	300	1000

     Given for Chi Square Table value v = 1 @ 5% level is 3.84.

16.
(a) Construct a Pie diagram to represent the following information of a company during a year.   3   CO1   [K₃]

     Item of cost   Rs(Crore)

     Labour Cost   10

     Overheads Cost   30

     Materials Cost   60

16.
(b) The probability function of an infinite discrete distribution is given by   2   CO3   [K₂]

     $P(X = j) = \frac{1}{2^j} (j = 1, 2, \ldots)$. Find Mean of X

Answer any FIVE Questions

PART C (5 x 10 = 50 Marks)

17. Calculate mean, median and mode for the following data: | 10 | CO1 | [K₂]

Less than : 10 20 30 40 50 60 70

Frequency: 7 14 28 45 60 68 70

18. Marks obtained by 10 students in Mathematics (x) and Statistics (y) are given below: — 10 — CO2 — [K₂]

X: 60 34 40 50 45 40 22 43 42 64

Y: 75 32 33 40 45 33 12 30 34 51

Find the correlation coefficient and two regression lines. Also find y when x = 55

19. The saving bank account of a customer showed an average balance of Rs.150 and a standard deviation of Rs.50. Assuming that the account balances are normally distributed. — 10 — CO3 — [K₃]

(i) What percentage of account is over Rs.200?
(ii) What percentage of account is between Rs.120 and Rs.170?
(iii) What percentage of account is less than Rs.75?

20 (a). A buyer of electric bulbs bought 100 bulbs each of two famous brands. Upon testing these he found that brand A had a mean life of 1500 hours with a standard deviation of 50 hours whereas brand B had a mean life of 1530 hours with a standard deviation of 60 hours. Can it be concluded at 5% level of significance, that the two brands differ significantly in quality? — 5 — CO4 — [K₃]

20 (b). A machine puts out 16 imperfect articles in a sample of 500. After the machine is overhauled, it puts out 3 imperfect articles in a batch of 100. Has the machine improved? — 5 — CO4 — [K₃]

21. The following data represents the number of units of productions per day turned out by different workers using four different types of machines. — 10 — CO5 — [K₄]

		Machine Type			
		A	B	C	D
	1	44	38	47	36
Workers	2	46	40	52	43
	3	34	36	44	32
	4	43	38	46	33
	5	38	42	49	39

Test whether the 5 men differ with respect to mean productivity and test whether the mean productivity is the same for the four different machine types.

| 22. (a) | In a sample of 1000 people in Mumbai, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state at 1% level of significance? | 3 | CO4 | [K₃] |

22. (a) In a sample of 1000 people in Mumbai, 540 are rice eaters and the rest are wheat eaters. Can we assume that both rice and wheat are equally popular in this state at 1% level of significance? 3 CO4 [K$_3$]

22. (b) A random sample is selected from each of 3 makes of ropes and their breaking strength are measured with the following results. 7 CO5 [K$_3$]

I   70   72   75   80   83
II 100 110 108   112 113 120 107
III 60   65   57   84   87   73

Test whether the breaking strength of the ropes differ significantly.

****