KUMARAGURU
college of technology
character is life

B.E/B.TECH DEGREE EXAMINATIONS: APRIL / MAY 2023

(Regulation 2018)

Second Semester

ARTIFICIAL INTELLIGENCE AND DATA SCIENCE

U18MAI2203: Probability and Statistics

COURSE OUTCOMES

CO1: Understand and apply the concept of probability and random variables and predict probabilities of events in models following normal distribution.

CO2: Apply the concepts of two-dimensional random variables, central limit theorem and estimation, which lay the foundation for Machine Learning and Data Science.

CO3: Perform hypothesis testing and interpret the results which will form the basis for Data Analysis.

CO4: Understand the principles of design of experiments and perform analysis of variance which will help in Data Analysis.

CO5: Learn and apply multivariate analysis necessary for Principal Component Analysis.

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

Answer all the Questions:-
PART A (10 x 2 = 20 Marks)
(Answer not more than 40 words)

1. State Baye's theorem.   CO1 [K1]

2. Find C, if $P[X=n] = C\left(\dfrac{2}{3}\right)^{n;}$; $n = 1,2,3,...$ CO1 [K1]

3. The joint probability density function of the random variable (X,Y) is given by $f(x,y) = kxye^{-(x^2+y^2)}$, $x > 0, y > 0$. Find the value of k.   CO2 [K2]

4. The regression equations of two variables x and y are $x + 6y = 14$ and $2x + 3y = 1$, find the mean values of x and y.   CO2 [K2]

5. What is Type I and Type II error?   CO3 [K1]

6. Give the formula for the ψ^2 test of independence for attributes?   CO3 [K1]

a	b
c	d

7. Define Analysis of variance.   CO4 [K1]

8. Write down the ANOVA table for completely randomized design.   CO4 [K1]

9. Define random vector.   CO5 [K1]

10. Describe random matrix.   CO5 [K1]

Answer any FIVE Questions:-
PART B (5 x 16 = 80 Marks)
(Answer not more than 400 words)

11. a) In a bolt factory machines A, B, C manufacture respectively 25%, 35% and 40% of the total. Of their output 5%, 4% and 2% are defective bolts. A bolt is drawn at random from the product and is found to be defective. What are the probabilities that it was manufactured by machines A, B and C.   8   CO1 [K₃]

b) In a certain factory turning razar blades there is a small chance of $\frac{1}{500}$ for any blade to be defective. The blades are in packets of 10. Use Poisson distribution to calculate the approximate number of packets containing (i) no defective, (ii) one defective, (iii) 2 defective blades respectively in a consignment of 10,000 packets.   8   CO1 [K₃]

12. a) The joint probability mass function of (X,Y) is given by P(x, y) = K(2x + 3y), x = 0,1,2, y = 1,2,3. (i) Find all the marginal and conditional probability distributions. (ii) Find the probability distribution of (X+Y) and P[X+Y>3].   8   CO2 [K₃]

b) The following table shows the ages (X) and blood pressure (Y) of 8 persons.   8   CO2 [K₃]

X	52	63	45	36	72	65	47	25
Y	62	53	51	25	79	43	60	33

Obtain the regression equation of Y on X and find the expected blood pressure of a person who is 49 years old.

13. a) Two independent samples of sizes 9 and 7 from a normal population had the following values of the variables.   8   CO3 [K₃]

Sample I	18	13	12	15	12	14	16	14	15
Sample II	16	19	13	16	18	12	15		

Do the estimates of the population variance differ significantly at 5% level of significance?

b) Two sample polls of votes for two candidates A and B for a public office are taken one form among residents of rural areas. The results are given below. Examine whether the nature of the area is related to voting preference in the election.   8   CO3   [K₃]

Area / Votes for	A	B	Total
Rural	620	380	1000
Urban	550	450	1000
Total	1170	830	2000

14. The following are the number of mistakes made in 5 successive days by 4 technicians working for a photographic laboratory: 16   CO4   [K₄]

Technician I (X₁)	Technician II (X₂)	Technician III (X₃)	Technician IV (X₄)
6	14	10	9
14	9	12	12
10	12	7	8
8	10	15	10
11	14	11	11

Test at a level of significance $\propto = 0.01$, whether the difference among the 4 sample means, can be attributed to chance.

15. Suppose the random variables X1, X2 , X3 have the covariance matrix   16   CO5   [K₅]

$$\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$ Determine the population principal components.

16. A set of data involving four "Four tropical feed stuffs A,B,C,D" tried on 20 chicks is given below. All the twenty chicks are treated alike in all respects except the feeding treatments and each feeding treatment is given to 5 chicks. Analyze the data. Weight gain of baby chicks fed on different feeding materials composed of tropical feed stuffs. 16 CO4 [K4]

						Total
A	55	49	42	21	52	219
B	61	112	30	89	63	355
C	42	97	81	95	92	407
D	169	137	169	85	154	714
					Grand Total	G=1695

****