PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc DEGREE EXAMINATION MAY 2025
(Second Semester)

Branch – **COMPUTER SCIENCE**

STATISTICS FOR COMPUTER SCIENCE

Time: Three Hours Maximum: 75 Marks

SECTION-A (10 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** marks (10 × 1 = 10)

Module No.	Question No.	Question	K Level	CO
1	1	The value of perfect positive correlation coefficient is: a) –1 b) 1 c) 0 d) any positive value	K1	CO1
	2	Which of the following methods is commonly used to minimize error in linear regression? a) K-means b) Mean Squared Error (MSE) c) Root Mean Squared Error (RMSE) d) Least Squares	K2	CO1
2	3	If a coin is tossed twice, the probability of getting atleast one head is a) 1/2 b) 1/4 c) 3/4 d) none	K1	CO2
	4	The probability of an impossible event is: a) 1 b) 0 c) 0.5 d) Depends on the event	K2	CO2
3	5	If a random variable X has an expected value E(X)=5 and a constant c=3, what is the expected value of E(cX)? a) 5 b) 15 c) 3 d) 8	K1	CO3
	6	The mean of a Poisson distribution is denoted by λ. What does it represent? a) The maximum possible value b) The average rate of occurrence in a given interval c) The variance of the distribution d) The cumulative probability of the events	K2	CO3
4	7	A paired t-test is used when: a) Comparing two independent samples. b) Comparing two related or matched samples. c) Testing for variance in large samples. d) Testing correlation between two variables.	K1	CO4
	8	If a researcher uses a z-test for a sample of 100 observations, what distribution does the sample mean follow? a) Exponential b) Chi-square c) Uniform d) Normal	K2	CO4
5	9	In a contingency table for a Chi-square test of independence with 3 rows and 4 columns, what is the degrees of freedom? a) 12 b) 6 c) 7 d) 2	K1	CO5
	10	One-way ANOVA is used to test: a) Differences between means of three or more independent groups b) Differences between variances of two samples c) Differences between means of two groups d) Relationships between two categorical variables	K2	CO5

SECTION - B (35 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** Marks (5 × 7 = 35)

Module No.	Question No.	Question	K Level	CO
1	11.a.	Discuss the difference between Correlation and Regression.	K4	CO1
		(OR)		

Cont...

1	11.b.	Compute the rank correlation coefficient from the following data: Applicant: A B C D E F G H I J Rank R1: 1 2 3 4 5 6 7 8 9 10 Rank R2: 3 4 10 7 8 5 1 2 6 9	K4	CO1
2	12.a.	Discuss the various definitions of probability.	K4	CO2
		(OR)		
	12.b.	i) State Addition theorem. ii) A bag contains 8 red and 6 blue balls. Two drawing of each 2 balls are made. Compute the probability that the first drawing gives two red balls and second drawing gives 2 blue balls, if the balls drawn first are replaced before the second draw.		
3	13.a.	Describe Normal distribution and its applications	K5	CO3
		(OR)		
	13.b.	Discuss the expectation and variance of the random variable with its properties		
4	14.a.	A cigarette manufacturing firm claims that its brand A line of cigarettes outsells its brand B by 8%. If it is found that 42 out of a sample of 200 smokers prefer brand A and 18 out of another sample of 100 smokers prefer brand B then test whether the 8% difference is a valid claim.	K5	CO4
		(OR)		
	14.b.	Scores obtained in a shooting competition b 10 soldiers before and after intensive training are given below: Before: 67 24 57 55 63 54 56 68 33 43 After: 70 38 58 58 56 67 68 75 42 38 Test whether the intensive training is useful at 0.05 level of significance.		
5	15.a.	Given the following contingency table for hair colour and eye colour. Test the association between the two?	K6	CO5

		Hair Colour		
		Fair	Brown	Black
Eye colour	Blue	15	5	20
	Grey	20	10	20
	Brown	25	15	20

		(OR)		
	15.b.	Describe One-way ANOVA		

SECTION -C (30 Marks)

Answer **ANY THREE** questions

ALL questions carry **EQUAL** Marks (3 × 10 = 30)

Module No.	Question No.	Question	K Level	CO
1	16	Construct the regression equation 'x' on 'y' and 'y' on 'x' from the following data: x: 15 20 25 30 35 40 45 Y: 8 14 20 26 32 38 44	K4	CO1
2	17	i) State Baye's theorem ii) The chance that doctor A will diagnose a disease x correctly is 60%. The chance that a patient will die by his treatment after correct diagnosis is 40% and the chance of death by wrong diagnosis is 70%. A patient of doctor A, who had disease x died. Compute the chance that his disease was diagnosed correctly?	K4	CO2
3	18	A discrete random variable X has the following probability distribution: X: 0 1 2 3 4 5 6 7 8 P(X=x): k 3k 5k 7k 9k 11k 13k 15k 17k Compute i) the value of 'k' ii) $P(X\leq 3)$, $P(0<X<3)$, $P(X\geq 3)$ iii) Mean	K6	CO3
4	19	Elucidate the procedure for testing of hypothesis on z-test in two samples and z-test in single proportion.	K5	CO4
5	20	Elaborate the two-way ANOVA	K5	CO5

Z-Z-Z END