PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

MSc(SS) DEGREE EXAMINATION DECEMBER 2025
(Second Semester)

Branch – SOFTWARE SYSTEMS (five years Integrated)

PROBABILITY AND STATISTICS

Time: Three Hours   Maximum: 75 Marks

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

Question No.	Question	K Level	CO
1.	If $P(A\|B) = 0.5$ and $P(B) = 0.4$, what is $P(A \cap B)$? a) 0.9 b) 0.2 c) 0.8 d) 0.4	K2	CO1
2.	If events A and B are independent, $P(A \cap B) =$? a) $P(A) + P(B)$ b) $P(A) \times P(B)$ c) $P(A) - P(B)$ d) $P(A) / P(B)$	K1	CO1
3.	Which distribution is often used to model the number of occurrences of an event in a fixed interval? a) Normal distribution b) Binomial distribution c) Poisson distribution d) Exponential distribution	K2	CO2
4.	What does the variance of a random variable measure? a) The central tendency of the distribution b) The probability of an event occurring c) The average of the data d) The spread of the distribution	K2	CO2
5.	Which of the following is used to test the hypothesis about a population mean when the sample size is large? a) Z-test b) t-test c) Chi-square test d) F-test	K2	CO3
6.	A p-value of 0.03 indicates that: a) There is no evidence to reject the null hypothesis b) The null hypothesis is always true c) There is strong evidence against the null hypothesis d) The evidence against the null hypothesis is weak	K2	CO3
7.	The F-distribution is used in ANOVA because it tests for: a) Differences in population means b) Differences in variances c) Normality of the data d) Correlation between variables	K1	CO4
8.	Which of the following is the correct assumption for the Chi-square test for goodness of fit? a) The sample data must be paired b) The data should follow a normal distribution c) The data must be continuous d) The expected frequency for each category should be greater than 5	K2	CO4
9.	In linear regression, the dependent variable is: a) The variable that is predicted b) The variable used to make predictions c) Independent of the predictor variable d) Not influenced by the independent variable	K1	CO5
10.	In simple linear regression, the regression equation is given by: a) $Y = \beta_0 + \beta_1 X$ b) $Y = \beta_0 + \beta_1 X + \varepsilon$ c) $Y = X + \beta_0$ d) $Y = \beta_0 + \varepsilon$	K1	CO5

Cont...

SECTION - B (35 Marks)
Answer ALL questions
ALL questions carry EQUAL Marks

$(5 \times 7 = 35)$

Question No.		Question	K Level	CO
11.a)		Derive the formula for conditional probability.	K2	CO1
		(OR)		
11.b)		In a city, 80% of residents are aware of a public health campaign, and 60% of those aware of the campaign are expected to take action. Only 20% of those who are unaware of the campaign will take action. If a resident takes action, what is the probability that they were aware of the campaign?	K4	
12.a)		Explain the concept of a probability mass function (PMF) and provide an example.	K2	CO2
		(OR)		
12.b)		A fair six-sided die, define the random variable X as the outcome of a roll. Calculate the probability mass function and cumulative distribution function for X.	K3	
13.a)		Explain the concept of hypothesis testing. What are the null and alternative hypotheses?	K2	CO3
		(OR)		
13.b)		What is a Type I error and a Type II error in hypothesis testing?		
14.a)		Describe the Chi-square test for goodness of fit with suitable examples.	K2	CO4
		(OR)		
14.b)		Differentiate between the F-test and t-test in hypothesis testing.		
15.a)		Explain the concept of linear regression and how the least squares method is used?	K2	CO5
		(OR)		
15.b)		Describe the significance of the coefficient of determination (R-squared) in regression analysis.	K3	

SECTION -C (30 Marks)
Answer ANY THREE questions
ALL questions carry EQUAL Marks

$(3 \times 10 = 30)$

Question No.	Question	K Level	CO
16.	Derive Bayes' Theorem and illustrate its uses with an example.	K3	CO1
17.	Derive the Constants of Binomial distribution through MGF.	K3	CO2
18.	Explain the concept of large sample test for a single population in the concept of testing of hypothesis.	K3	CO3
19.	What is the ANOVA test and how is it used to compare the means of three or more groups? Provide an example and discuss its assumptions.	K2	CO4
20.	Discuss the issues of multicollinearity in multiple regression. How can it be detected and addressed?	K3	CO5

Z-Z-Z   END