PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION MAY 2025
(Second Semester)

Branch – **COMPUTER SCIENCE WITH DATA ANALYTICS**

STATISTICAL DATA ANALYSIS

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	A sector diagram is also called _____. (a) Bar diagram (b) Pie diagram (c) Histogram (d) frequency polygon	K2	CO1
	2	Which one of the methods is not a primary data collection method? (a) Questionnaire method (b) Data collected from published sources (c) Local correspondent method (d) Indirect investigation	K1	CO1
2	3	_____ is the statistical tool that studies the degree of all the relationships between variables. (a) mean (b) range (c) correlation (d) mean deviation	K1	CO2
	4	In the regression equation $Y = 21 - 3X$, the slope is _____. (a) 21 (b) -21 (c) 3 (d) -3	K2	CO2
3	5	In binomial distribution, if n is a number of trials and p is the probability of success, then the value of the mean is _____. (a) -np (b) npq (c) n(1-p) (d) np	K1	CO3
	6	Normal Distribution is symmetric is about _____ (a) Variance (b) Mean (c) Standard deviation (d) Covariance	K2	CO3
4	7	In testingthe difference between two proportions, _____ test is used. (a) Z-test (b) t-test (c) Fisher test (d) F-test	K1	CO4
	8	The level of significance is generally denoted by _____. (a) β (b) δ (c) γ (d) α	K2	CO4
5	9	In a chi-square goodness-of-fit test, theoretical frequencies are also called _____ frequencies. (a) actual (b) expected (c) empirical (d) observed	K1	CO5
	10	The Kruskal-Walli's test is used for _____. (a) comparing two independent groups (b) comparing two related groups (c) comparing more than two independent groups (d) testing the association between two 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.	Describe qualitative and quantitative data.	K2	CO1
		(OR)		
	11.b.	Find mean deviation about median and its coefficient for the following data: C-I: 0-10 | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 Frequency: 7 | 12 | 18 | 25 | 16 | 14 | 8		
2	12.a.	Calculate the correlation coefficient between the height of father (X) and son (Y) from the data given below: X in inches: 64 | 65 | 66 | 67 | 68 | 69 | 70 Y in inches: 66 | 67 | 65 | 68 | 70 | 68 | 72	K3	CO2

Cont...

2		(OR)	K3	CO2
	12.b.	Find the rank correlation coefficient for the percentage of marks secured by a group of 8 students in Economics and Statistics: Marks in Economics: 50, 60, 65, 70, 75, 40, 70, 80 Marks in Statistics: 80, 71, 60, 75, 90, 82, 70, 50		
3	13.a.	If a random variable X follows Poisson distribution such that $P(X=1) = P(X=2)$. Find (i) mean of the distribution (ii)$P(X=0)$. $(e^{-2} = 0.13354)$	K2	CO3
		(OR)		
	13.b.	State the properties of Normal distribution.		
4	14.a.	A machine is designed to produce insulting washers or electrical devices of average thickness of .025 cm. A random sample of 10 washers was found to have an average thickness of 0.024 cm. with a S.D. of .002 cm. Test the significance of the deviation. Value of 't' for 9 degrees of freedom at 5 percent level is 2.262.	K4	CO4
		(OR)		
	14.b.	A sample of 1000 students from Bombay university was taken and their average weight was found to be 112 lbs. with a standard deviation of 20 lbs. Could the mean weight of the students in the population be 120 pounds?		
5	15.a.	In an infantile paralysis epidemic 500 persons contracted the disease 300 received no serum treatments and of these 75 became paralyzed. Of those who did receive serum treatment 65 became paralyzed. Was the serum treatment effective.	K3	CO5
		(OR)		
	15.b.	Two types of plastic, each produced by different process, were tested for strength. The measurement are presented in the following table. Do the data present the evidence of difference between locations of distribution of strength of two plastics? Use Mann Whitney-U test with $\alpha = 0.01$. Plastic type-I: 15.5, 18.7, 22.3, 17.6, 19.1, 14.8 Plastic type-II: 21.2, 22.4, 18.3, 19.3, 17.1, 27.7		

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	Scores obtained by two teams are given below: Team A: 15, 10, 7, 5, 3, 2 Team B: 20, 10, 5, 4, 2, 1 Calculate coefficient of variation and state which team is more consistent.	K4	CO1
2	17	Calculate two regression lines and find the age of husband, when wife's age is 30. Further calculate the age of wife when husband's age is 25. Age of Husband's: 22, 23, 23, 24, 26, 27, 27, 28, 30, 30 Age of Wife's: 18, 20, 21, 20, 21, 22, 23, 24, 25, 26	K4	CO2
3	18	The atoms of radioactive element are randomly disintegrating. If every gram of this element, on average emits 3.9 alpha articles per second, what is the probability that during the next second the number of alpha particles emitted from 1 gram is (i) atmost 6 (ii) atleast 2 (iii) atleast 3 and atmost 6.$(e^{-3.9} = 0.020)$.	K4	CO3
4	19	An I.Q. test was administered to 5 persons before and after they were trained. The results are given below: Candidates: I, II, III, IV, V Before Training: 110, 120, 123, 132, 125 After Training: 120, 118, 125, 136, 121 Test whether there is any change in I.Q. after the trainingprogram.$(t_{0.01}(4) = 4.6)$	K4	CO4
5	20	The following table gives the yields of 15 samples of plot under three varieties of seed: A: 20, 21, 23, 16, 20 B: 18, 20, 17, 15, 25 C: 25, 28, 22, 28, 32 Test using analysis of variance whether there is significant difference in the average yield of seed.	K4	CO5

Z-Z-Z END