PSG COLLEGE OF ARTS & SCIENCE
(AUTONOMOUS)

BSc DEGREE EXAMINATION DECEMBER 2023
(Second Semester)

Branch – COMPUTER SCIENCE WITH DATA ANALYTICS

STATISTICAL DATA ANALYSIS

Time: Three Hours

Maximum: 50 Marks

SECTION-A (5 Marks)
Answer ALL questions
ALL questions carry EQUAL marks   (5 x 1 = 5)

1 Diagrams and graphs are tools of
   (i) collection of data
   (ii) analysis of data
   (iii) presentation of data
   (iv) primary data

2 Rank Correlation found by
   (i) karl pearson
   (ii) Spearman
   (iii) Galton
   (iv) Fisher

3 The mean of Binomial Distribution is
   (i) npq
   (ii) np
   (iii) $\sqrt{npq}$
   (iv) $\sqrt{np}$

4 The standard deviation of the sampling distribution is called
   (i) Standard Error
   (ii) Sample Error
   (iii) Type I Error
   (iv) Type II Error

5 The degrees of freedom for contingency table are on the basis of
   (i) (n-1)
   (ii) (r-1)
   (iii) (c-1)
   (iv) (r-1) (c-1)

SECTION - B (15 Marks)
Answer ALL Questions
ALL Questions Carry EQUAL Marks   (5 x 3 = 15)

6 a Sketch a pie diagram from the following data.

Brand of computer	Dell	HB	Samsung	LG	Acer
Sales in percentage	23	30	15	15	17

OR

   b The profits made by 75 firms in a year are as follows. Calculate mean and median from the following data.

Profits (Rs. in '000)	10-20	20-30	30-40	40-50	50-60	60-70
No. of firms	5	15	20	20	10	5

7 a Apply the rank correlation between IQ's and marks scored in an examination.

Person	A	B	C	D	E	F
I.Q	100	110	140	160	120	130
Exam Mark	70	80	81	78	72	75

OR

   b Explain the properties of regression coefficient.

8 a State addition and multiplication theorems on probability.

OR

   b The number of mistakes counted in one hundred typed pages of a typist revealed that he made 2.8 mistakes on an average per page. Find the probability that in a page typed by him, (i) there is no mistake (ii) there are two or less mistakes. (Given $e^{-2.8}$ =0.061)

9   a   A sample of 400 families in an old city is selected randomly, and a sample of 500 families is randomly selected from several new colonies of the same city. A survey is conducted for the number of houses possessing television (TV) sets. The number of TV holders in the old city is 48 out of 400 selected families and 120 in new colonies out of 500 families. Then the hypothesis whether the proportion of TV holders in old city and in the new colonies is the same.

OR

    b   Explain the testing procedure of one sample t test.

10 a   Explain the procedure for testing the equality of two variances.

OR

    b   From the following data find out whether there is any association between shift and quality of parts produced? (Given $\chi^2_{(2,0.05)}$= 5.991)

Shift	Good	Bad	Total
Day	900	130	1030
Evening	700	170	870
Night	400	200	600
Total	2000	500	2500

SECTION -C (30 Marks)
Answer **ALL** questions
ALL questions carry **EQUAL** Marks   (5 x 6 = 30)

11 a   The monthly profits in rupees of 83 shops are distributed as follows:

Profits per shop	0-50	50-100	100-150	150-200	200-250	250-300
Number of shop	8	15	30	17	9	4

Construct an ogive curve and identify the median value from the graph.

OR

    b   From the following data on daily sales of computer sets, analyze standard deviation and coefficient of variation.

No. of Computer sets	5	7	10	11	15	25	30
No. of Days	1	3	7	6	5	2	1

12 a   A Company wanted to assess the impact of R and D expenditure (X) on annual profit (Y) following table presents data for past 8 years.

X	9	7	5	10	4	5	3	2
Y	45	42	41	60	30	34	25	20

Examine the correlation - coefficient.

OR

    b   Distinguish between correlation and regression.

13 a   A manufacturing firm produces pipes in two plants, I and II, with daily production of 1,500 and 2,000 pipes respectively. The fraction of defective pipes produced by the two plants are 0.006 and 0.008 respectively. If a pipe selected at random from the day's production is found to be defective, what is the probability that it has come from plant I, plant II?

OR

    b   Point out the properties of Binomial distribution.

14 a The following information related to purchase of bulbs from two manufactures A &B:

Manufacturer	Mean	S.D	N
A	2950 hours	100 hours	100
B	2970 hours	90 hours	100

Is there a significant difference in the mean life of two makes of bulbs?

OR

b To test the variability of a certain modification in typist desks, 9 typists were given two tests of as nearly as possible the same nature, one on the desk in use and the other on the new type. The following difference in the number of words typed per minute were recorded:

Typist	A	B	C	D	E	F	G	H	I
Increase in number of words	2	4	0	3	-1	4	-3	2	5

Do the data indicate that modification in desk promotes speed in typing?

(Given for $t_{8, 0.05} = 2.306$)

15 a To test the significance of the variation of the retail prices of a commodity in three principal cities, Bombay, Calcutta and Delhi, four shops were chosen at random in each city and prices observed in rupees were as follows:

Bombay	16	8	12	14
Calcutta	14	10	10	6
Delhi	4	10	8	8

Do the data indicate that the prices in the three cities are significantly different?
(Given $F_{0.05,2,9} = 4.26$)

OR

b Test whether the following two samples have been drawn from the same population using Mann Whitney U Test.

Sample I	134	146	104	119	124	161	107	113	94
Sample II	70	101	118	85	107	132	94	97	-

Z-Z-Z   END