

**S.Y.B.Sc.(IT) SEM-IV**  
**COMPUTER ORIENTED STATISTICS**

**Total Marks :75**

**NOTE:**

- (i) All the questions are compulsory.
- (ii) All the questions carry equal marks.
- (iii) Simple calculator is allowed.

**Q-1 Attempt Any Three:****[15]**

[1] Find the Mean and Mode for the following data:

Age	20-25	25-30	30-35	35-40	40-45	45-50
No. of Employees	3	3	16	5	7	6

[2] Find the Median for the following data representing the heights of 50 students.

Height in cms	140-150	150-160	160-170	170-180	180-190	190-200
No. of students	7	3	7	12	15	6

[3] Find Quartile Deviation for the following data:

Rainfall (in Inches)	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No. of years	12	8	10	12	10	7	6

[4] Find the coefficient of variation for the following data:

Production	100-110	110-120	120-130	130-140	140-150
No. of workers	30	52	100	68	20

[5] The average monthly sales for the first few months for a salesman is Rs. 5000. For the remaining months of the year, his performance was improved with average sales of Rs. 7000 per months. This resulted in an average sales/ month for the entire year of Rs. 6500. Find no. of months for his monthly sales was Rs. 5000.

[6] A student Scores 60 marks in theory, 94 in practical and 24 in CW in Physics. Another student scores 65 in practical and 25 per classwork in Physics. Another scores 72 marks in theory, 81 in practical and 15 in CW in Physics. If theory, practical and CW in Physics are assigned weights 6,3,1. Find which student is better?

**Q-2 Attempt Any Three:****[15]**

[1] Compute the first four moments for the following distribution of marks.

Marks out of 20	5	6	7	8	9	10	11	12	13	14	15
No. of Students	1	2	5	10	20	51	22	11	5	3	1



[2]

(1) Find Pearson's coefficient of skewness

Mean = 70.5; Median = 80; Mode = 85; Standard deviation = 19.33.

(2) Find Bowley's coefficient of skewness

$Q_1 = 466.67$ ;  $Q_2 = 554.28$ ;  $Q_3 = 633.33$ .

[3] One box contains 5 red and 4 blue balls and the other box contains 4 red and 7 blue balls. A ball is selected at random from the first box and without noting the color put in the other. A ball is then drawn from the second box. What is the probability that it is blue?

[4]

(1)  $P(\bar{A}) = 2/3$ ,  $P(B) = 1/4$ ,  $P(A \cup B) = 5/12$ . Find  $P(A \cap B)$ ,  $P(A/B)$  and  $P(B/A)$ .

(2) If A and B are independent events and  $P(A) = 1/3$  and  $P(B) = 1/4$ , then find  $P(A \cap B)$ .

[5] Find Expected Mean and Variance from the following data:

No. of Petals	3	4	5	6	7	8	9
P(x)	0.05	0.10	0.20	0.30	0.25	0.075	0.025

[6] Define the following Terms:

1. Mutually Exclusive Events
2. Complementary Events
3. Addition Theorem of probability
4. Multiplication Theorem of Probability
5. Equally Likely Events

**Q-3 Attempt Any Three:****[15]**

[1] Define the following Term:

1. Central Limit Theorem
2. Sampling Distribution
3. Type I Error
4. Standard Error of the Mean (S.E.)
5. Null Hypothesis





- [2] A manufacturer developed a fishing rod and claims that it has a mean breaking strength of 8 kg. with a standard deviation of 0.5 Kg. If a random sample of 50 fishing lines is tested and found to have a mean breaking strength of 7.8Kg, test the manufacturer claim at  $\alpha = 0.01$ . [ $z_{\alpha} = 2.58$ ]
- [3] A sample of 8 items has a mean 1408 with a standard deviation 192. Construct a 95% confidence interval for the population mean. [ $t_{\frac{\alpha}{2}, n} = 2.365$ ]
- [4] The height of 10 soldiers selected at random had a mean height of 158cm and a variance of 39.0625cm. Assuming a significance level of 5%, test the hypothesis that the soldiers of the population are on average lesser than 162.5cm. tall. [ $t_{\alpha, v} = 1.833$ ]
- [5] In a study of television viewing habits, in order to obtain an interval estimate of the average number of hours per week that teenagers spend watching television programs, a random sample of 100 teenage children is taken. Sample investigation revealed a mean of 9.2 hours, with a standard deviation of 3.2 hours. Obtain the desired interval estimate with a confidence coefficient 0.99.
- [6] Find the P-Value for the following by using the following given area under standard Normal variate z: (i)  $P[z \geq 1.5]$  (ii)  $P[z \leq 2]$  (iii)  $P[-1.5 \leq Z \leq 2]$   
Given: Area between  $z = 0$  and  $z = 1.5$  is 0.4332  
Area between  $z = 0$  and  $z = 2$  is 0.4772

**Q-4 Attempt Any Three:**

[15]

- [1] A population of heights has  $\mu = 68$ . What is the probability of selecting a sample of size 25 that has a mean of 70 or greater and sample S.D. = 4. Take  $\alpha = 5\%$ .  
[Given that  $t_{\alpha, v} = t_{0.05, 24} = 1.711$ ]
- [2] In a study on cholesterol levels, a sample of 12 men and women was chosen. The plasma cholesterol levels (n mol/L) of the subjects were as follows: 6.0, 6.4, 7.0, 5.8, 6.0, 5.8, 5.9, 6.7, 6.1, 6.5, 6.3 & 5.8. We assume that these 12 subjects constitute a simple random sample of a population of similar subjects. We wish to estimate the variance of the plasma cholesterol levels with a 95% confidence interval.  
[Given  $\chi^2_{11, 0.975} = 21.920$  and  $\chi^2_{11, 0.025} = 3.816$ ]



[3] The yields of the three varieties of wheat using four kinds of fertilizers are given as:

Fertilizers	Varieties of wheat		
	V <sub>1</sub>	V <sub>2</sub>	V <sub>3</sub>
t <sub>1</sub>	64	72	74
t <sub>2</sub>	55	57	47
t <sub>3</sub>	59	66	58
t <sub>4</sub>	58	57	53

0.05 level of significance to test the following hypothesis Ho: There is no difference in the average yield of wheat when different kinds of fertilizers are used. [ $F_{\alpha, v_1, v_2} = 4.76$ ]

[4] A die is tossed 120 times with the following results:

X	1	2	3	4	5	6
f	20	22	17	18	19	24

Is this a balanced die? Take  $\alpha = 5\%$ .  $\chi^2_{0.05, 5} = 11.070$

[5] In an experimental study of the independence of hypertension on smoking habits, the following data are taken from 180 individuals.

	Non-Smokers	Moderate Smokers	Heavy Smokers	Total
Hypertension	21	36	30	87
No-hypertension	48	26	19	93
Total	69	62	49	180

Test the hypothesis at a 0.05% level of significance that the presence or absence of hypertension is independent of smoking habits. [ $\chi^2_{0.05, 2} = 5.991$ ]

[6] Write the properties of  $\chi^2$  (CHI-SQUARE) distribution.

**Q-5 Attempt Any Three:**

**[15]**

[1] The marks obtained by ten students are as follows. Find Spearman's rank correlation coefficient.

Roll No.	: 1	2	3	4	5	6	7	8	9	10
Marks in Accounts	: 90	88	90	76	88	62	98	90	70	76
Marks in Economics	: 61	58	64	73	73	78	58	82	58	67

[2] Calculate Spearman's rank correlation coefficient between the following data:

X :	81	72	60	33	29	11	56	42
Y :	75	56	42	15	30	20	60	80





[3] The table below gives the heights of fathers (X) and the heights of their sons (Y) respectively.

heights of fathers (inches) (X)	64	62	66	63	67	61	69	65	67	66
heights of sons(inches) (Y)	67	65	67	64	68	65	67	64	70	66

Find the regression equation X on Y

[4] Fit the trend by the method of least squares and estimate it for the year 2005.

Year:	1993	1994	1995	1996	1997	1998	1999	2000
Export (in millions):	8	10	12	11	13	15	14	17

[5] A survey conducted to study the relationship between expenditure on clothes & expenditure on entertainment in a locality gave the following results: -

	Average	Standard deviation
Expenditure on cloths	300.00	20.00
Expenditure on entertainment	100.00	15.00

**Coefficient of Correlation**  $r = 0.78$

Find the regression of expenditure on entertainment related to expenditure on clothes.

[6] If the two lines of regression are  $4y - 15x + 530 = 0$  and  $20x - 3y - 995 = 0$ . Find (i) the Mean values of X and Y and (ii) the Coefficient of correlation r.