

S.Y. B.Sc. (IT) : Sem. IV  
**Computer Oriented Statistical Techniques**



Time : 2½ Hrs.]

Prelim Question Paper

[Marks : 75

- N.B.:** (1) All questions are compulsory.  
 (2) Use of Non-programmable calculator is permitted.

1. Attempt any **THREE** of the following : **[15]**

- (a) Use the following frequency distribution of weekly wages to find the arithmetic mean of wage of employees at P & R company.

Weekly Wage (\$)	Number of employees
250.00 - 259.99	8
260.00 - 269.99	10
270.00 - 279.99	16
280.00 - 289.99	14
290.00 - 299.99	10
300.00 - 309.99	5
310.00 - 319.99	2

- (b) Find the median TV viewing time of the following data for the 400 students.

Viewing time (minutes)	Number of students
300 - 399	14
400 - 499	46
500 - 599	58
600 - 699	76
700 - 799	68
800 - 899	62
900 - 999	48
1000 - 1099	22
1100 - 1199	6

- (c) The final grades in Mathematics of students of a class are recorded in the following table. Find the mode.

Grades in Mathematics	Number of students
50 - 54	1
55 - 59	2
60 - 64	11
65 - 69	10
70 - 74	12
75 - 79	21
80 - 84	6
85 - 89	9
90 - 94	4
95 - 99	4

- (d) Cities A, B and C are equidistant from each other. A motorist travels from A to B at 30 mph, from B to C at 40 mph and from C to A at 50 mph. Determine his average speed.

- (e) Find the 10 - 90 percentile range of the height of students from the following data.

Height (inches)	Number of students
60 - 62	5
63 - 65	18
66 - 68	42
69 - 71	27
72 - 74	8

- (f) Find the standard deviation for the following data of daily wages of workers.

Daily wages (in Rs.)	number of workers
200 - 220	24
220 - 240	32
240 - 260	40
260 - 280	17
280 - 300	7

2. Attempt any **THREE** of the following: [15]

- (a) A population consists of three numbers 3, 7, and 15. Consider all possible samples of size 2 that can be drawn with replacement from this population. Show that average of the sample mean is the population mean.

- (b) Find the first four moments about mean for the following data.

X	12	14	16	18	20	22
f	1	4	6	10	7	2

- (c) For a data  $Q_1 = 268.25$ ,  $Q_3 = 290.75$ ,  $P_{10} = 258.125$ ,  $P_{90} = 301$  and median = 279. Find  
 (a) Quartile Coefficient of Skewness.  
 (b) 10 - 90 Percentile Coefficient of Skewness.

- (d) In how many ways 5 boys and 2 girls can be arranged in a row such that the girls occupy the ends.

- (e) If an unbiased die is rolled once, find the probability distribution of "number on the uppermost face of the die". Also find the mean of the probability distribution.

- (f) Find the probability that in 120 tosses of a fair coin proportion of heads is  
 (a) between 40% & 60%    (b) more than 5/8

3. Attempt any **THREE** of the following: [15]

- (a) A random sample of 50 mathematics grades out of a total of 200, showed a mean of 75 and a standard deviation of 10. Find 95% confidence limits for mean of 200 grades.

- (b) A sample of 150 light bulbs of brand A showed a mean life time of 1400 hours with a standard deviation of 120 hours. A sample of 200 light bulbs of brand B showed a mean life time of 1200 hours with a standard deviation of 80 hours. Find 98% confidence limits for the difference of mean life time of the populations of brands A & B.

- (c) The standard deviation of the lifetimes of a sample of 200 electric bulbs was computed to be 100 hours. Find 95% confidence limits for the standard deviation of all such electric bulbs.

- (d) The breaking strength of cables produced by a manufacturer has a mean of 1800 pounds (lb) and a standard deviation of 100 lb. By a new technique in the manufacturing process, it is claimed that the breaking strength can be increased. To test this claim, a sample of 50 cables is tested and it is found that the mean breaking strength is 1850 lb. Can we support the claim at the 0.01 significance level?

- (e) A pair of dice is tossed 100 times and it is observed that 7 appeared 23 times. Test the hypothesis that the dice are fair (i.e. not loaded) by using a two-tailed test at 5% significance level.

- (f) A sample of 100 electric light bulbs produced by manufacturer A showed a mean lifetime of 1190 h with a standard deviation of 90 h. A sample of 75 bulbs produced by manufacturer B showed a mean lifetime of 1230 h with a standard deviation of 120 h. Is there a difference between the mean lifetime of the two brands of bulbs at significance level of 0.01?

4. Attempt any **THREE** of the following:

[15]

- (a) A sample of 12 measurements of the breaking strength of cotton threads gave a mean of 7.38 g and a standard deviation of 1.24 g. Find 99% confidence limits for the actual breaking strength.
- (b) The mean life time of electric bulbs produced by a company has in the past been 1120 h. A sample of 8 electric bulbs recently chosen from the supply of newly produced bulbs showed a mean life time of 1070 h with a standard deviation of 125 h. Test the hypothesis that the mean life time has not changed, using significance level of 0.01.
- (c) In the past, the standard deviation of weights of certain 40.0-ounce package filled by a machine was 0.25 ounce (oz). A random sample of 20 packages showed a standard deviation of 0.32 oz. At 0.05 significance level is the apparent increase in variability significant?
- (d) The number of books borrowed from a public library during a particular week is given in the following table.

Day	Mon	Tue	Wed	Thu	Fri
Number of books borrowed	132	108	120	114	146

Test the hypothesis that the number of books borrowed does not depend on the day of the week, using significance level of 0.01.

- (e) In his experiment with peas, Mendel observed that 315 were round & yellow, 108 were round & green, 101 were wrinkled & yellow and 36 were wrinkled & green. According to his theory, the numbers should be in the proportion 9:3:3:1. Is there any evidence to doubt his theory at 0.05 significance level?
- (f) The following table shows the number of students in class A & B, who passed and failed in an examination. Using Yates' correction at 1% significance level, test the hypothesis that there is no difference between class A & class B.

	Passed	Failed
Class A	72	17
Class B	64	23

5. Attempt any **THREE** of the following:

[15]

- (a) A temperature of 100 degrees Celsius ( $^{\circ}\text{C}$ ) corresponds to 212 degrees Fahrenheit ( $^{\circ}\text{F}$ ), while a temperature of  $0^{\circ}\text{C}$  corresponds to  $32^{\circ}\text{F}$ . Assuming that a linear relationship exists between Celsius and Fahrenheit temperatures, find the equation connecting Celsius and Fahrenheit temperatures. Also find the Fahrenheit value for  $^{\circ}\text{C} = 200$ .
- (b) The following table gives the values of pressure (P) corresponding to various values of the volume (V). The relation between P & V is  $PV^{\alpha} = C$  where  $\alpha$  & C are constants.

V	54.3	61.8	72.4	88.7	118.6	194.0
P	61.2	49.2	37.6	28.4	19.2	10.1

Find the values of  $\alpha$  & C. Write equation connecting P & V.

- (c) Fit a least square parabola  $Y = a_0 + a_1X + a_2X^2$  to the following data.

X	0	1	2	3	4	5	6
Y	2.4	2.1	3.2	5.6	9.3	14.6	21.9

- (d) The following table shows the respective heights (in inches) X & Y of a sample of 12 fathers and their oldest sons.

X	65	63	67	64	68	62	70	66	68	67	69	71
Y	68	66	68	65	69	66	68	65	71	67	68	70

Find the correlation coefficient between X and Y.

- (e) A correlation coefficient based on a sample of size 18 was computed to be 0.32. Can we conclude at 5% significance level that the corresponding population correlation coefficient differs from zero?

(f) 12 pairs of (X, Y) values gave the regression equation of Y on X as  $Y = 35.82 + 0.476 X$ . The standard error of estimate is 1.28 and the standard deviation of X is 2.66. At the 0.05 significance level, test the null hypothesis that the population regression coefficient of the regression equation is 0.18 versus the alternative hypothesis that it exceeds 0.18.

**Table Values:**

1) A = Area under standard Normal Curve between 0 & Z

Z	1.645	1.96	2.28	2.33	2.58	2.65
A	0.45	0.475	0.4887	0.49	0.495	0.4960

2) p = Area under chi-square curve to the left of  $\chi_p^2$

Degree of freedom	1	3	4	19
$\chi_{0.95}^2$		7.81		30.1
$\chi_{0.99}^2$	6.63		13.3	

3) p = Area under t curve to the left of  $t_p$

Degree of freedom	7	10	11	16
$t_{0.95}$		1.81		1.75
$t_{0.99}$			2.72	
$t_{0.995}$	3.50			



**S.Y. B.Sc.IT Sem-IV: Paper Discussion Schedule**

Date	Day	Timing	Centre
18 April, 2018	Wednesday	5.00 p.m. to 7.00 p.m.	Andheri
19 April, 2018	Thursday	9.00 a.m. to 11.00 a.m.	Dadar
19 April, 2018	Thursday	5.00 p.m. to 7.00 p.m.	Borivali
19 April, 2018	Thursday	9.00 a.m. to 11.00 a.m.	Thane
19 April, 2018	Thursday	4.00 p.m. to 6.00 p.m.	Nerul