

7022

M.Com./Semester I (OC)

G

Paper 6102 : STATISTICAL ANALYSIS

Time : 3 Hours

Maximum Marks : 100

(Write your Roll No. on the top immediately on receipt of this question paper.)

Attempt All questions.

All questions carry equal marks.

1. (a) Explain Bayes' theorem and its significance in probability. 6
- (b) A sample of 500 respondents was selected in a large city to study consumer behavior. One of the questions asked was, "Do you enjoy shopping for clothing" ? Of 240 males, 136 answered 'yes' and of 260 females, 224 answered 'yes'. What is the probability that a respondent chosen at random :
 - (i) enjoys shopping for clothing
 - (ii) is a female and enjoys shopping for clothing
 - (iii) is a male or enjoys shopping for clothing
 - (iv) is a male or a female
 - (v) is male and does not enjoy shopping for clothing. 7

- (c) Researchers at a Medical College have determined that children under 2 years old who sleep with the lights on have a 42% chance of becoming myopic before they are 16. Children who sleep in darkness have a 21% probability of becoming myopic. A survey indicates that 28% of children under 2 sleep with some light on. Find the probability that a child under 16 is myopic. 7

Or

- (d) Distinguish between statistical dependence and statistical independence. Give examples also. 6
- (e) Three airlines serve a small city. Airline A has 50% of all the scheduled flights, airline B has 30% and airline C has the remaining 20%. Their on-time rates are 80%, 65% and 40% respectively. A plane has just left on time.
- (i) What is the probability that it was Airline A ?
- (ii) What is the probability that it was Airline C ? 7
- (f) The controls of an airplane have several backup systems or redundancies, so that if one fails the plane will continue to operate. Suppose that the mechanism that controls the flaps has two backups. If the probability that the

main control fails is .0001 and the probability that each backup will fail is .01, what is the probability that all three fail to operate ? 7

2. (a) Distinguish between simple random sampling and stratified sampling. 6

(b) The final marks in a course are normally distributed with a mean of 70 and standard deviation of 10. The professor must convert all marks to letter grades. She decides that she wants 10% A's, 30% B's, 40% C's, 15% D's and 5% E's. Determine the cutoffs each letter grade. 7

(c) The manufacturer of a sports car wants to estimate the proportion of people in a given income bracket who are interested in the model. The company wants to know the population proportion to within 0.10 with 99% confidence. Current company records indicate that the

proportion p may be around 0.25. What is the minimum required sample size for this survey ? 7

Or

- (d) Distinguish between discrete and continuous probability distributions. 6
- (e) A student specialising in finance is deciding on the number of firms to which he should apply. Given his work experience and grades, he can expect to receive a job offer from 80% of the firms to which he applies. The student decides to apply to only four firms. What is the probability that he received (i) no job offers ? (ii) at least two job offers ? 7
- (f) A car manufacturer wants to estimate the average miles-per-gallon highway rating for a new model. From experience with similar models, the manufacturer believes the miles-

per-gallon standard deviation is 4.6. A random sample of 100 highway runs of the new model yields a sample mean of 32 miles per gallon. Give a (i) 95% confidence interval and (ii) 99% confidence interval for the population average miles-per-gallon highway rating. 7

3. (a) What are non-parametric tests ? Explain the advantages and disadvantages of using non-parametric tests. 8
- (b) A test is being designed to compare the wearing quality of two brands of tyres. Six cars were randomly selected, equipped with one tyre of brand A and one tyre of brand B (the other two tyres are not part of the test) and driven for one month. The Table below gives the amount of wear (in thousandths of an inch) that took place in such a test :

Tyre brand	Car					
	1	2	3	4	5	6
A	125	64	94	38	90	106
B	133	65	103	37	102	115

Do the sample data provide sufficient evidence for us to conclude that the *two* brands show unequal wear, at 5% level of significance ?

12

Or

(c) Distinguish between *upper* tailed and *lower* tailed tests. 8

(d) A teacher in a school believes that students who finish exams more quickly than others have better exam scores.

The following set of data shows the score and order of finish for 12 students on an exam. Using rank correlation method, do these data indicate that the first students to complete an exam have higher grades ? ($\alpha = 0.05$): 12

Order of finish	Exam score
1	90
2	74
3	76
4	60
5	68
6	86
7	92
8	60
9	78
10	70
11	78
12	64

4. (a) What is meant by p-chart in quality control ? Under what conditions do we use a p-chart ? 8
- (b) An article reports that smaller firms seem to be hiring more than large ones as the economy picks up its pace.

The table below gives number of employers hired and those laid off, out of a random sample of 1032, broken down by firm size. Is there evidence that hiring practices are dependent on firm size ? Test at 5% level of significance. 12

	Small firm	Medium-sized firm	Large firm
Number hired	210	290	325
Number laid off	32	95	80

Or

- (c) Explain the concept of standard error of mean. 8
- (d) To study waiting times of customers in a bank, the following summary of data (in minutes) is available for subgroups of 4 people for a 10 day period : 12

Day	Mean	Range
1	13.6	3.5
2	14.3	4.1
3	15.3	5.0
4	12.6	2.8
5	11.8	3.7
6	12.9	4.8
7	17.3	4.5
8	13.9	2.9
9	12.6	3.8
10	15.2	4.6

(i) Compute control limits for the range.

(ii) Compute control limits for the mean.

(iii) Is there evidence of special cause variation in (i) ?

$$4+4+4=12$$

P.T.O.

5. (a) Distinguish between coefficient of determination and coefficient of correlation. 8
- (b) A real estate developer is considering investing in a shopping mall on the outskirts of a large metropolitan city. Three plots of land are being evaluated. Of particular importance is the income in the area surrounding the proposed mall. A random sample of four families is selected from the area near each proposed mall and the sample results are shown below. At 5% level of significance, can the developer conclude there is a difference in the mean income ? 12

Income (in Rs. '000)

Area A	Area B	Area C
64	74	75
68	71	80
70	69	76
60	70	78

Or

- (c) What is meant by 'level of significance' in hypothesis testing ? Explain. 8

- (d) A company keeps track of its expenditure on advertising (X_1) and its spending on in-store displays (X_2) over a period of 10 weeks. The resulting sales for each week are recorded as the dependent variable Y . Use the data given below to estimate the least squares equation relating these variables : 12

Y	X_1	X_2
72	12	5
76	11	8
78	15	6
70	10	5
68	11	3
80	16	9
82	14	12
65	8	4
62	8	3
90	18	10