

STATISTICS

(Paper II)

Time Allowed : Three Hours
Maximum Marks : 300

QUESTION PAPER SPECIFIC INSTRUCTIONS

(Please read each of the following instructions carefully before attempting questions)

There are EIGHT questions divided in two Sections.

Candidate has to attempt FIVE questions in all.

Question Nos. 1 and 5 are compulsory and out of the remaining, any THREE are to be attempted choosing at least ONE question from each Section.

The number of marks carried by a question/part is indicated against it.

Wherever any assumptions are made for answering a question, they must be clearly indicated.

Word limit in questions, wherever specified, should be adhered to.

SECTION 'A'

1. Attempt any five sub-parts of the following:

- (a) Describe classification of states in a Markov chain. Define n -step transition probability. Prove the following Chapman – Kolmogorov equation for the transition probabilities:

$$p_{ij}^{(n+1)} = \sum_k p_{ik} p_{kj}^{(n)} \quad 12$$

- (b) What are the goals achieved through sensitivity analysis? If in the following linear programming problem (L.P.P.):

$$\text{Maximize } z = 3x_1 + 2x_2 + 5x_3$$

subject to

$$[x_1 + 2x_2 + x_3 \leq 430 ; 3x_1 + 2x_3 \leq 460 ; x_1 + 4x_2 \leq 420 ; x_1, x_2, x_3 \geq 0]$$

the profit coefficient of x_1 is reduced to 2 from 3, determine whether the original solution remains optimal. 12

- (c) In a life test with replacement, 35 heaters were put into continuous operation and the first five failures occurred after 250, 380, 610, 980 and 1250 hours. (i) Assuming the exponential model, construct a 95 percent confidence interval for the mean life of this kind of heater. (ii) Test the manufacturer's claim that the mean life of these heaters is at least 5000 hours at 5% level of significance. Given $\chi_{0.025}^2(10) = 20.48$, $\chi_{0.975}^2(10) = 3.25$, $\chi_{0.95}^2(10) = 3.94$, $\chi_{0.05}^2(10) = 18.31$ 12

- (d) Describe a CUSUM control chart. Compare this chart with a Shewhart chart with respect to performance. How is a V-mask useful in CUSUM charts? 12

- (e) What is the effect of lengthening of a test on its (i) reliability and (ii) validity? A given test has a reliability coefficient of 0.8 and standard deviation 20. What is the estimated reliability coefficient of this test in a group in which standard deviation is 15? 12

- (f) Explain (i) feasible solution and (ii) basic solution in the context of linear programming problem (L.P.P.). Prove that the objective function of a L.P.P. assumes its minimum at an extreme point of the convex set generated by the set of all feasible solutions to the problem, and if it assumes its minimum at more than one extreme point then it takes on the same value for every convex combination of those particular points. 12

2. (a) Show that an assignment problem (A.P.) is the special case of a transportation problem. Describing the steps of the method you use to solve an A.P., find the solution of the following A.P.:

Job \ Lathe machines	1	2	3	4	5	6
----------------------	---	---	---	---	---	---

20

A	6.0	7.0	6.0	9.0	10.0	12.0
B	8.0	9.0	5.0	7.0	9.5	6.0
C	1.0	1.5	1.0	1.5	1.0	1.0
D	4.5	4.5	4.0	6.0	5.0	3.0
E	3.0	4.0	2.0	2.5	2.0	3.0
F	3.5	3.0	1.0	4.0	2.5	2.5

- (b) Define transient and persistent states in a Markov chain. Let $\{X_n, n \geq 0\}$ be a Markov chain with three states $\{0, 1, 2\}$ and with transition matrix

$$\begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 3/4 & 1/4 \end{bmatrix}$$

The initial probability distribution is

$$P\{X_0 = i\} = 1/3, i = 0, 1, 2$$

Find (i) $P[X_2 = 2, X_1 = 1 | X_0 = 2]$ and (ii) $P[X_3 = 1, X_2 = 2, X_1 = 1, X_0 = 2]$

20

- (c) What do you mean by acceptance sampling plans? Between a single sampling plan with $n=20, c=2$ and a double sampling plan with $n_1=10, c_1=0, n_2=10, c_2=2$, can it be said that the second inspection scheme is more economical than the first? Give reasons justifying your answer.

20

3. (a) What is replacement problem? How would you determine the replacement policy of items whose maintenance cost increases with time and money carries a rate of interest r per year? An auto owner finds from his past record that the cost per year of an auto whose purchase cost is Rs. 60,000 is as given below:

Year	Running cost (Rs.)	Resale value (Rs.)
1	10000	30000
2	12000	15000
3	14000	7500
4	18000	3750
5	23000	2000
6	28000	2000
7	34000	2000
8	40000	2000

Determine at what time is its replacement due.

15

- (b) Let $f(t) = \lambda^a t^{a-1} e^{-\lambda t}$ be the Weibull failure density function. Given the censored sample $t_{(1)} < t_{(2)} < \dots < t_{(d)}$ corresponding to a total sample of size n from Weibull distribution, show that the estimating equations for λ and a are respectively

$$\hat{\lambda} = d / (\sum_{i=1}^{d-1} t_{(i)}^{a-1} + (n-d+1) t_{(d)}^{a-1})$$

and

$$d/\hat{a} = \sum_{i=1}^d \log t_{(i)} = \lambda [\sum_{i=1}^{d-1} t_{(i)}^a \log t_{(i)} + (n-d+1) t_{(d)}^a \log t_{(d)}]$$

25

- (c) Derive an expression for finding optimum inventory level in a single-period probabilistic model without set-up cost with continuous demand and discrete replenishment unit. A baking firm makes a profit of Rs. 5.00 per kg of cakes sold on the same day it baked and disposes all unsold cakes at a loss of Rs. 1.20 per kg. The demand follows rectangular distribution between 2000 and 3000. Find optimum quantity to be baked daily.

20

4. (a) Describe simulation and Monte Carlo simulation. What is the role of random numbers in simulation? Describe how would you select a random sample from a population having exponential distribution.

15

- (b) Explain the principles which are used to find the solution of a rectangular game which does not possess a saddle point. Find the value of the game and mixed strategies of the players for the following game:

Player A \ Player B	B1	B2	B3	B4	B5	B6
A1	4	2	0	2	1	1

25

A2	4	3	1	3	2	2
A3	4	3	7	-5	1	2
A4	4	3	4	-1	2	2
A5	4	3	3	-2	2	2

- (c) Explain assignable and chance causes of variation. Describe the principles on which a control chart is based. Describe $(\bar{X}, \sigma)$ chart and mention its advantages and disadvantages over $(\bar{X}, R)$ chart. 20

SECTION 'B'

5. Attempt any five sub-parts of the following:

- (a) What do you mean by 'seasonal index' in the analysis of a time series? Discuss the different steps involved in the computation of the seasonal indices using link relative method. 12
- (b) Discuss the problem of autocorrelation and its consequences when one applies least square method for estimating the parameters. 12
- (c) Differentiate between total fertility rate (TFR), gross reproduction rate (GRR) and net reproduction rate (NRR). Does TFR strictly conform to our idea of a measure for reproduction? How does NRR indicate the growth of population? 12
- (d) Write a note on the collection, compilation and publication of demographic data by National Sample Survey Organization. 12
- (e) What do you mean by a life table? The values of l_x in a life table are given as follows:

age (x)	102	103	104	105	106	107	108
l_x	97	59	32	15	6	2	0

- Calculate remaining entries of the life table for $x \geq 102$ and find the probability that a person of exact age 102 years will die between ages 103 and 107. 12
- (f) Explain clearly the concepts of reliability and validity of scores in psychological and educational experiments. Explain the relation of validity coefficient to error of measurement. 12
6. (a) What do you understand by (i) agricultural statistics, (ii) trade and price statistics? What are the present Indian official statistical systems for collecting these data? 20
- (b) What do you understand by cost of living index number? Explain how would you construct it. The following table gives the group index numbers and the weights of different heads of expenditure in the calculation of cost of living index number except the index for the group 'fuel and lighting':

Group	Index	Weight
Food	221	35
Clothing	198	14
Fuel and Lighting	—	15
Rent	183	8
Miscellaneous	161	20

- If the overall cost of living index is 193, find the index number of fuel and lighting group. 20
- (c) Explain the terms (i) percentile score, (ii) summation score as a function of ability, (iii) Z-score and (iv) T-score. Describe how T-scores are to be found when we are given test scores in the form of a frequency table. 20
7. (a) Discuss the differences between direct and indirect methods of standardization of death rates. Calculate the crude and standardized death rate of the year 1990 by direct and indirect methods of standardization by taking 1980 population as standard:

Age group (years)	1980 Population	1980 Total deaths	1990 Population	1990 Age specific death rates per thousand population

25

0 – 10	10000	240	15000	22·0
10 – 25	12000	145	15000	13·0
25 – 60	6000	92	8000	15·0
60 and above	8000	480	9000	47·0

- (b) What are the tests to be satisfied by a good index number? Examine how far they are met using Paasche's index number and Fisher's ideal index number. Using Laspeyres' and Paasche's index numbers, the price index numbers for the year 1970 with 1966 as base year are 121 and 144 respectively. What would be the value of Fisher's index number? **20**
- (c) What is a time series? What do you mean by the analysis of a time series? Describe briefly different components of a time series. Which component of the time series is mainly applicable in the following cases? (i) Fall of death rate due to advances in science (ii) Fire in a factory (iii) Sales of new year greeting cards **15**
8. (a) Prove the following: (i) $p_x = e_x / (1 + e_{x+1})$ (ii) $m_x = \mu_{x+1/2}$ (iii) $\mu_x \approx (8(l_{x-1} - l_{x+1}) - (l_{x-2} - l_{x+2}))/12 l_x$ **20**
- (b) Define generalized least squares estimator. Obtain mean and sampling variance for the estimator. Show that a generalized least square estimator is equivalent to a two-step procedure estimator. **30**
- (c) Describe the test-retest method for estimating the reliability of a test and discuss its merits and demerits. **10**