ONS

2008 100102

STATISTICS (Optional)

सांख्यिकी (वैकल्पिक)

Time : 3 hours **Maximum Marks : 200**

Note :

(i) *In all attempt __Five__ Questions.*

(ii) *Question No. **1** is **Compulsory**.*

(iii) *Of the remaining Questions, Attempt __Any four__ by selecting __one Question__ from each section.*

(iv) *Numbers of optional questions upto the prescribed number in the order in which questions have been solved, will only be assessed and excess answers of the question/s will not be assessed.*

(v) *Statistical and logarithmic tables will be supplied on request.*

(vi) *Use of your own simple electronic calculator is allowed.*

(vii) *Candidate should not write roll number, any names (including his/her own), signature, address or any indication of his/her identity anywhere inside the answer book otherwise he/she will be penalised.*

1. Answer ***any four*** of the following :

(a) $X_1, X_2, \ldots, X_{11}$ are independent random variables (r.v.s.) such that X_i has the exponential distribution with mean $(1/i)$ $i = 1, 2, \ldots, 11$ respectively. Obtain the distribution of **10**

(i) $X = \sum_{i=1}^{10} i . X_i$ (ii) $Y = \sum_{i=2}^{11} i . X_i$

Also find the coefficient of correlation between X and Y.

(b) In the analysis of variance (ANOVA) of one-way classified data, prove that : **10**

Total sum of squares = within group sum of squares + Between group sum of squares.

Obtain the computational formula for each of the above sum of squares.

(c) $X_1, X_2, \ldots, X_n$ is a random sample of size n from an uniform distribution defined over the interval (a, b). Obtain the moment estimators of a and b. **10**

(d) Five artists have submitted bids to complete 4 sets of plates and the costs are given in the following table : **10**

Artists	Cost per plate (in Rs.)			
	P_1	P_2	P_3	P_4
A	245	295	210	360
B	220	300	155	300
C	250	220	190	320
D	240	250	180	390
E	235	245	200	370

One set of plates is to be commissioned per artist so that the total cost is minimum. Which artist will not receive a commission ? What is the minimum total cost ?

(e) If a random variable X has beta distribution of first kind with parameters $\frac{m}{2}$ and $\frac{n}{2}$.

Find the distribution of $Y = \frac{nX}{m(1-X)}$. Hence or otherwise state the mean and variance of Y. **10**

SECTION - A

2. Answer the following sub-questions :

(a) A two-dimensional random variable (X, Y) has the following joint probability density function (p.d.f.). **15**

$$f(x, y) = x + y \quad \text{if } 0 \le x \le 1, 0 \le y \le 1.$$
$$= 0 \quad \text{elsewhere.}$$

Find (i) $P(X \le Y)$

(ii) $P\left(X^2 + Y^2 \le \frac{1}{4}\right)$

(iii) $P\left(Y \ge \frac{1}{2}\right)$

(b) A random variable (r.v.) X has a normal distribution with parameters μ and σ^2. Obtain the expressions for the odd ordered and even ordered central moments of X. Hence obtain the values of β_1 and β_2. Comment on the nature of the distribution. **15**

(c) If X_1, X_2 and X_3 are the independent observations from an univariate normal population with mean zero and unit variance, state giving reasons the sampling distribution of $U = \frac{\sqrt{2}\,X_3}{\sqrt{X_1^2 + X_2^2}}$. **10**

By using an appropriate table find

(i) C such that $P[U \leq C] = 0.995$ and

(ii) $P[|U| \leq 4.3]$.

3. Answer the following sub-questions :

(a) A two-dimensional discrete random variable (X, Y) has the joint probability distribution as follows : **15**

X \ Y	1	2	3	4	Total
0	1/24	1/12	1/12	1/24	1/4
1	1/12	1/6	1/6	1/12	1/2
2	1/24	1/12	1/12	1/24	1/4
Total	1/6	1/3	1/3	1/6	1

Obtain (i) marginal distribution of X,

(ii) marginal distribution of Y,

(iii) conditional distribution of X given $Y = 3$,

(iv) $E[X|Y=3]$ and $V[X|Y=3]$.

Are X and Y independent r.v.s. ?

(b) The joint probability mass function (p.m.f.) of X_1, X_2, X_3 is : **15**

$$P(x_1, x_2, x_3) = \frac{n!\ p_1^{x_1}.p_2^{x_2}.p_3^{x_3}}{x_1!\,x_2!\,x_3!}$$

where $x_i = 0, 1, \ldots, n$

For $i = 1, 2, 3$;

$0 < p_i < 1$ for $i = 1, 2, 3$;

$x_1 + x_2 + x_3 = n$, $p_1 + p_2 + p_3 = 1$.

Derive the joint moment generating function of X_1 X_2 X_3. Hence or otherwise determine :

(i) The value of coefficient of correlation between X_1 and X_2.

(ii) The value of coefficient of partial correlation between X_1 and X_2 with respect to X_3.

(c) Show that a chi-square distribution having n degrees of freedom is positively skewed. **10**

SECTION - B

4. Answer the following sub-questions :

(a) With two strata, a sampler would like to have $n_1 = n_2$ for administrative convenience instead of using the values that are obtained by Neyman's allocation. **15**

If $V(\bar{y}st)$ and $V(\bar{y}st)opt$ denote the variances given by equal and Neyman's allocation respectively, show that ignoring finite population correction,

$$\frac{V(\bar{y}st) - V(\bar{y}st)opt}{V(\bar{y}st)opt} = \frac{(r-1)^2}{(r+1)^2}$$

where $r = \frac{n_1}{n_2}$ as given by Neyman's allocation.

(b) Describe the technique of analysis of covariance (ANCOVA) in Randomised Block Design (R.B.D.). **10**

(c) Let n triplets of observations be available on X_1, X_2 and X_3. Obtain the equation of the regression plane of X_3 on X_1 and X_2 by using the method of least squares. **15**

5. Answer the following sub-questions :

(a) State the estimators of the population mean in ratio and regression methods. Are these estimators unbiased ? Comment on the relative efficiency of these estimators with respect to that in simple random sampling without replacement (SRSWOR). Justify your comment. **15**

(b) The following table gives the plan of an experiment involving three factors N, P, K in blocks of size 4 plots each : **10**

Replication	Blocks	Factorial Effects
1	I	(1) pk nk np
	II	n p k npk
2	III	(1) k npk np
	IV	p nk pk n
3	V	p npk (1) nk
	VI	n k pk np

(i) Identify the type of the experiment.

(ii) Construct the ANOVA table for the experiment.

(c) Find the distribution function and the p.d.f. of the rth order statistic $X_{(r)}$ when $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$ are the order statistics of a random sample of size n from a continuous distribution with p.d.f. $f(x)$. **15**

SECTION - C

6. Answer the following sub-questions :

(a) $X_1, X_2, \ldots, X_n$ is a random sample (r.s.) from a geometric distribution with following p.m.f. : **15**

$$p(x, \theta) = \theta(1-\theta)^{x-1} \text{ for } 0 < \theta < 1 \text{ ; } x = 1, 2, 3, \ldots$$
$$= 0 \qquad \text{elsewhere.}$$

Show that the sample mean $\overline{X}$ is an unbiased estimator of $\frac{1}{\theta}$. Determine variance of $\overline{X}$ and show that $\overline{X}$ attains the lower bound given by Crammer - Rao inequality.

(b) Let $X_1, X_2, \ldots, X_n$ be a r.s. from a gamma distribution with parameters (α, λ). Find the asymptotic variance of the maximum likelihood estimator (m.l.e.) of α when λ is known. **10**

(c) $X_1, X_2, \ldots, X_{10}$ is a r.s. from an exponential distribution with mean θ. Find the most powerful test of size $\alpha = 0.05$ to test the hypothesis $H_0 : \theta = 1$ against the alternative $H_1 : \theta = 4$. **15**

7. Answer the following sub-questions :

(a) X is a r.v. having uniform distribution on the interval $(0, \theta)$. Let $X_1, X_2, \ldots, X_n$ be a r.s. from this distribution. Obtain a sufficient statistic for θ. Is it unbiased for θ ? If not, obtain an unbiased estimator of θ from it. Further, if $T = 2 \sum_{i=1}^{n} X i / n$, evaluate the relative efficiency of T with respect to the unbiased estimator of θ obtained above. **15**

(b) A r.s. of 100 boys and 100 girls from a certain college were asked whether they attend any coaching classes. It was observed that 36 boys and 64 girls attend the coaching classes. Find the 95% confidence interval for the difference between the proportion of boys and girls who attend coaching classes. **10**

(c) Describe Mann - Whitney test. **15**

SECTION - D

8. Answer the following sub-questions :

(a) Solve the following linear programming problem (L.P.P.) by Big M method : **10**

Maximize $Z = 3\ X_1 + 2\ X_2$

Subject to

$2\ X_1 + X_2 \leq 2$

$3\ X_1 + 4X_2 \geqslant 12$

$X_1, X_2 \geqslant 0$

(b) Explain the following methods of estimating the trend values in a time series : **15**

(i) Method of moving averages.

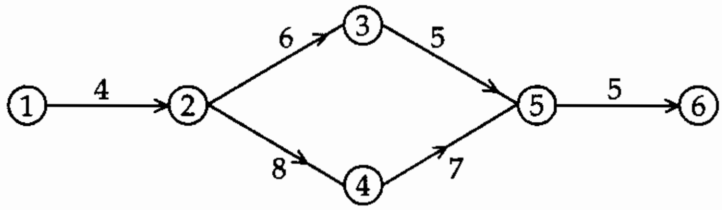
(ii) Curve fitting by least squares method.

(iii) Graphical method.

(c) Suppose a large number of lots, each of size 2000 and 0.5% defective are submitted for inspection. Given three single sampling plans (i) $n = 40$, $c = 0$, (ii) $n = 80$, $c = 1$ and (iii) $n = 120$, $c = 2$; state giving reasons which sampling plan involves the minimum total inspection, considering both sampling inspection and screening of the rejected lots. **15**

9. Answer the following sub-questions :

(a) Calculate the slack times for the activities in the following network. Hence or otherwise find the critical path in this network. **10**



(b) Let L (p) and P (q) represent Laspeyre's index number for prices and Paasche's index number for quantities. Show that $L(p) \times P(q) = V_{01}$ where V_{01} is the value index. Hence or otherwise show that **15**

$$\frac{L(p)}{L(q)} = \frac{P(p)}{P(q)}$$

where L (q) and P (p) represent Laspeyre's index number for quantities and Paasche's index number for prices respectively.

Also show that if (i) $A(p) = \frac{L(p) + P(p)}{2}$ then A (p) > Fisher's ideal index number for prices.

(ii) $H(p) = \frac{2\,L(p)\,P(p)}{L(p) + P(p)}$ then

Fisher's ideal index number for prices > H (p).

(c) Explain the following procedures of setting limits on a control chart for fraction defective in a production process when the sample size is not fixed and standard of the process is not specified : **15**

(i) Method of variable control limits.

(ii) Method of stabilised control limits.

(iii) Method based on average sample size.

-oOo-