2010
STATISTICS - I (Optional)

100043

Standard : Degree **Total Marks : 200**

Nature : Conventional **Duration : 3 Hours**

Note :

(i) *Answers must be written in* **English** *only.*

(ii) *Question* **No. 1** *is* **Compulsory.** *Of the remaining questions, attempt* **any four** *selecting one question from* **each section.**

(iii) *Figures to the* **RIGHT** *indicate marks of the respective question.*

(iv) *Use of log table, Non-Programmable calculator is permitted, but any other Table / Code / Reference book are not permitted.*

(v) *Make suitable assumptions, wherever be necessary and state the same.*

(vi) *Number of optional questions upto the prescribed number in the order in which they have been solved will only be assessed. Excess answers will not be assessed.*

(vii) *Credit will be given for orderly, concise and effective writing.*

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

1. Answer *any four* of the following :

(a) Three fair coins are tossed 3000 times. Find the frequencies of the distribution of heads and tails and tabulate the result. Also calculate mean and s.d. of the distribution. **10**

(b) Find the maximum likelihood estimator (MLE) for the parameter λ of a poisson distribution from n sample values. Also find its variance. **10**

(c) Determine the regression equation of : **10**

1) X_1 on X_2 and X_3

2) X_2 on X_1 and X_3.

$r_{12} = 0.8,\ r_{13} = 0.6,\ r_{23} = 0.5$

$\sigma_1 = 10,\quad \sigma_2 = 8,\quad \sigma_3 = 5$

P.T.O.

Marks

(d) The mean value of random sample of 60 items was found to be 145 with standard deviation of 40. Find the 95% confidence limits for the population mean. What size of sample is required to estimate the population mean with error of 5 units with 95% confidence using the sample mean. **10**

(Given $Z_T = 1.96$ at 95% confidence level).

(e) Using the sign test to see there is a difference between the number of days until collection of an account receivable before and after a new collection policy. Use the 0.05 significance level. **10**

Before	30	24	32	38	35	40	42	45	43	38	37
After	42	18	12	16	33	37	46	49	46	32	32

SECTION – A

2. Answer the following sub-questions :

(a) (i) Verify that the function given below is a distribution function. Also find $P(X \leq 4)$. **5**

$$F(X) = \begin{cases} 0 & , x < 0 \\ 1 - e^{-x/4} & , x \geq 0 \end{cases}$$

(ii) Using Chebychev's inequality determine how many times a fair coin must be tossed in order that the probability will be atleast 0.90 that the ratio of the observed number of heads to the number of tosses will be between 0.4 and 0.6. **5**

(b) Find the probability distribution function of number of heads obtained when a fair coin is tossed 4 times. Also find expectation and variance. **10**

(c) Find the moment generating function of the random variable X whose probability mass function is **10**

X	-2	3	1
$P(X=x)$	$1/3$	$1/2$	$1/6$

Also find first two moments about the origin.

(d) State and prove Chebychev's inequality. **10**

Marks

3. Answer the following sub-questions :

(a) (i) Define : 5

(a) Independent events.

(b) Sample space.

(c) Event.

(d) Complementary event.

(e) Mutually exclusive events.

(ii) A random variable X has the probability function 5

X	−2	−1	0	1	2	3
$P(X=x)$	0.1	K	0.2	2K	0.3	K

Find

(a) K.

(b) $P(X \leq 1)$.

(c) Obtain the distribution function of X.

(b) A discrete random variable X has the probability density function 10

X	−2	−1	0	1	2	3
$P(X=x)$	0.2	K	0.1	2K	0.1	2K

Find the mean and variance.

(c) Find the moment generating function of a random variable X if the r^{th} moment about the origin is given by $\mu_r' = r!$ 10

(d) State and prove central limit theorem for identically distributed variables. 10

Marks

7. Answer the following sub-questions :

(a) Calculate correlation coefficient : **10**

X	5	10	15	20	25
Y	2	4	10	8	6

(b) (i) Explain in terms of correlation coefficient. **5**
$r=0$, $r=+1$, $r=-1$, $r<0$, $r>0$.

(ii) Calculate multiple correlation coefficient $R_{1.23}$, if $r_{12}=0.8$, $r_{13}=0.5$, $r_{23}=0.3$. **5**

(c) Define Bivariate normal distribution. State and explain the properties of it. **10**

(d) Of 1000 people consulted, 811 liked chocolates, 752 liked toffees, and 418 liked sweets, 570 liked chocolates and toffees, 356 liked chocolates and sweets, 348 liked toffees and sweet, 297 liked all three. Is this information correct ? **10**

SECTION – D

8. Answer the following sub-questions :

(a) (i) Write short note on stratified sampling. **5**

(ii) Write short note on multistage sampling. **5**

(b) Write short note on cluster sampling. State its merits and demerits. **10**

(c) Describe briefly the methods of collecting primary data. **10**

(d) Describe the methods of collection of official statistics and their limitations. **10**

9. Answer the following sub-questions :

(a) (i) What do you mean by SRSWR and SRSWOR ? **5**

(ii) What do you mean by systematic sampling ? **5**

(b) Explain the principal steps in sample survey. **10**

(c) What are non-sampling errors ? What are the sources of non-sampling errors ? **10**

(d) Write note on the following : **10**

(i) N.S.S.O.

(ii) I.S.I.

-o O o-

Marks

SECTION – B

4. Answer the following sub-questions :

(a) State and prove Rao-Blackwell theorem. **10**

(b) (i) If $S^2 = \frac{1}{n-1} \Sigma (X_i - \overline{X})^2$ then S^2 is an unbiased estimator of σ^2. **5**

(ii) Prove that in sampling from a N (μ, σ^2) population sample mean is a consistent estimator of μ. **5**

(c) The means of two large samples of sizes 1000 and 2000 are 67.5 and 68 respectively. Test the equality of means of the two populations each with standard deviation 2.5 (Given Z_T at 5% for 2998 degrees of freedom = 3) **10**

(d) The following is an arrangement of 20 Boys (B) and 15 Girls (G) lined up to purchase tickets for a picture show. **10**

B, GG, BBB, G, BB, G, B, G, BB, GGG, BBB, G, BB, GGG, BBBBBB, GGG.

Test for randomness at the 5% level of significance.

(Given $Z_T = 1.96$ at 5% level of significance)

5. Answer the following sub-questions :

(a) Define Minimum Variance Unbiased Estimator (MVUE) and solve the problem. **10**

Let Y_1 and Y_2 be two unbiased estimators of θ such that Var $(Y_1) = 2$ Var (Y_2) obtain constants K_1 and K_2 such that $K_1Y_1 + K_2Y_2$ is MVUE of θ.

(b) (i) Let joint density of X and Y be **5**

$$f(x, y) = \frac{2}{\theta^2} e^{-\frac{(x+y)}{\theta}}, \theta < x < y < \infty$$

Find E (Y) and Variance of Y.

(ii) Define : **5**

(1) Type I and Type II errors.

(2) Null and Alternate Hypothesis.

P.T.O.

Marks

(c) A manufacturing company sold their items in three package sizes Large (Economy), Normal, Small (Handy). Experience has shown that they sell in the ratio 3 : 5 : 2. Sales returns from a particular region shows that 40 dozen, 45 dozen, and 15 dozen respectively of different package sizes have been sold. Is this pattern of sales significantly different from the previously established one ? **10**

(Given χ^2_T at 5% level for 2 degrees of freedom = 5.99)

(d) The following data relate to the daily production of cement (In m. tonnes) a large plant for 30 days. **10**

11.5, 10.0, 11.1, 10.0, 12.3, 11.1, 10.2, 9.6, 8.7, 9.3, 9.3, 10.7, 11.3, 10.4, 11.4, 12.3, 11.4, 10.2, 11.6, 9.5, 10.8, 11.9, 12.4, 9.6, 10.5, 11.6, 8.3, 9.3, 10.4, 11.5

Use sign test to test the null hypothesis that the plant's average daily production of cements is 11.2 m. tonnes at 5% level of significance.

(Given $Z_T = -1.645$ at 5% level of significance)

SECTION – C

6. Answer the following sub-questions :

(a) Find the regression equation of Y on X and X on Y. **10**

X	1	2	3	4	5
Y	1	4	5	0	3

(b) (i) Why are there in general two regression lines ? When do they coincide ? **5**

(ii) Given $r_{12} = 0.5$, $r_{13} = 0.4$, $r_{23} = 0.1$. Find $r_{12 \cdot 3}$ **5**

(c) Define multinomial distribution and solve the problem. **10**

A community consists of 50 percent Hindus, 40 percent Muslims and 10 percent Sikhs. In a sample of 6 individuals selected at random, find the probability that 3 are Hindus, 2 are Muslims and 1 is a Sikh.

(d) From the following data. Find out the frequencies. **10**

($A \beta$), (αB), ($\alpha \beta$), (α), (β).

(AB) = 100, (A) = 300, N = 1000, (B) = 600.

P.T.O.