AUD

350031

2011

STATISTICS (Optional)
Paper – I

Standard : Degree | **Total Marks : 200**

Nature : Conventional (Essay) type | **Duration : Three hours**

N.B. : 1) *Answers must be written in* ***English***.

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

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

4) *Though use of log table, Non-Programmable calculator is permitted, but any other Table/Code/Reference book are not permitted.*

5) *Make suitable assumptions, wherever be necessary and state the same clearly.*

6) *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.*

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

8) *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.*

SEAL

Marks

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

(a) (i) Define probability density function of a continuous random variable, with illustration.

(ii) The p.d.f. of a continuous random variable X is given as

$$f(x) = K\,x^2 e^{-x}, \quad 0 \le x < \infty$$
$$\quad = 0 \quad , \quad \text{otherwise}$$

Find the value of K and distribution function (d.f.) of X. 10

(b) Find the maximum likelihood estimator for the parameter λ of a Poisson distribution on the basis of a sample of size n_1. Also find its variance. 10

P.T.O.

Marks

(c) In a study of random sample of 120 students the following results are obtained : $\overline{X}_1 = 68$, $\overline{X}_2 = 70$, $\overline{X}_3 = 74$, $\sigma_1^2 = 100$, $\sigma_2^2 = 25$, $\sigma_3^2 = 81$, $r_{12} = 0.60$, $r_{13} = 0.7$, $r_{23} = 0.65$. Where X_1, X_2, X_3 denote percentage of marks obtained by a student in I, II and III test respectively. Obtain the least square equation of X_3 on X_1 and X_2. Also estimate X_3 when $X_1 = 60$ and $X_3 = 67$. **10**

(d) In simple random sampling without replacement (SRSWOR), show that the sample mean $(\overline{Y}_n)$ is an unbiased estimator of population mean $(\overline{Y}_N)$ i.e. $E(\overline{Y}_n) = \overline{Y}_N$. **10**

(e) Following is a sample drawn from the continuous population in the order in which the observations were made : 79, 13, 138, 129, 59, 76, 75, 53, 122, 98, 38, 37, 84, 110, 76, 58, 98, 70, 24, 52. Test the hypothesis of randomness of the sample. Use 5% level of significance. (Given $C_{0.025} = 6$ and $C^1_{0.025} = 16$ for $n_1 = 10$ $n_2 = 10$). **10**

SECTION – A

2. Answer the following sub-questions :

(a) (i) During a war 1 ship in 10 was sunk on an average in making a certain voyage. What is the probability that exactly three out of a convoy of 6 ships will arrive safety ? **5**

(ii) Examine whether the weak law of large numbers holds for the sequence $\{X_K\}$ of independent random variables defined as follows :

$P[X_K = \pm 2^K] = 2^{-(2K+1)}$

$P[X_K = 0] = 1 - 2^{-2K}$. **5**

(b) X and Y be two random variables with a joint probability distribution given by

X \ Y	0	1	2	3
0	K	2K	3K	4K
1	2K	4K	6K	8K
2	3K	6K	9K	12K

(i) Find the marginal probability distributions of X and Y

(ii) Are X and Y independent ?

(iii) Find $P[X \le 1, Y \le 1]$ and $P[X + Y \le 1]$. **10**

Marks

(c) Define the following :

(i) Moment generating function

(ii) Probability generating function

(iii) Characteristic function.

Also prove that "The moment generating function of sum of a number of independent random variables is equal to the product of their respective moment generating functions". 10

(d) If $X_i = \begin{cases} 1, & \text{with probability } p \\ 0, & \text{with probability } q \end{cases}$

then the distribution of the random variable $S_n = X_1 + X_2 + \ldots + X_n$; where X_i's are independent, is asymptotically normal as $n \to \infty$. 10

3. Answer the following sub-questions :

(a) (I) The defects on a plywood sheet occur at random with an average of one defect per 50 square feet. What is the probability that a 5×8 feet sheet will have (i) no defects and (ii) at least one defect ? 5

(II) Let X_i assume the values i and –i with equal probabilities. Show that the law of large numbers can not be applied to the independent variables $X_1, X_2, \ldots$, i.e. X_i's, $i = 1, 2, 3, \ldots$ 5

(b) A continuous random variable X has p.d.f.

$f(x) = e^{-x}, 0 \leq x < \infty$

$= 0$, otherwise.

find the mean, variance and third central moment of X. Also find coefficient of skewness. 10

(c) The random variable X assume the value 'x' with the probability law :

$P[X = x] = p\, q^{x-1}, x = 1, 2, 3, \ldots\ 0 < p \leq 1, q = 1 - p.$

Find the moment generating function (m.g.f.) of X and hence find mean and the variance. 10

(d) Use Chebyshev's inequality to find the number of times a true coin must be tossed in order that the probability will be at least 0.90 that the proportion of the number of heads will lie between 0.4 and 0.6. 10

Marks

SECTION – B

4. Answer the following sub-questions :

(a) Define, with illustrations the following terms :

(i) Unbiased estimator

(ii) Uniformly Minimum Variance Unbiased Estimator (UMVUE)

(iii) Mean Square Error (MSE)

(iv) Minimum MSE estimators

(v) Sufficient estimator. **10**

(b) (I) Let $X_1, X_2, ..., X_n$ is a random sample from a normal population $N(\mu, 1)$.

Show that $T = \frac{1}{n} \Sigma X_i^2$ is an unbiased of $\mu^2 + 1$. **5**

(II) Define the following :

(i) Simple and composity hypothesis

(ii) Type – I and Type – II errors

(iii) p-value

(iv) Critical region

(v) Power of a test. **5**

(c) A health club advertised a weight reducing program and claimed that on the average participant in program loses weight in 6 months. A person wanted to verify the claim. He selects 10 participants randomly and records their weights before and after the program as given below :

Participant No.	1	2	3	4	5	6	7	8	9	10
Weight before (in lbs)	120	125	115	130	123	119	122	127	128	118
Weight after (in lbs)	111	114	107	120	115	112	112	120	119	112

Do the above data support the claim of the health club ?
[Given $t_{9,0.01} = -2.821$]. **10**

(d) Let X has $N(\mu, \sigma^2)$. Construct the sequential probability ratio test of strength α and β to test $H_0 : \mu = \mu_0$ against $H_1 : \mu = \mu_1$ $(\mu_1 > \mu_0)$. Where σ^2 is known. **10**

Marks

5. Answer the following sub-questions :

(a) State and prove factorization theorem, in case of discrete random variables. 10

(b) (i) A random sample X_1, X_2, X_3, X_4, X_5 of size 5 is drawn from a normal population with unknown mean μ and variance σ^2. Consider the following estimators to estimate μ.

(1) $T_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}$

(2) $T_2 = \frac{X_1 + X_2}{2} + X_3$

Are T_1 and T_2 unbiased for μ ? 5

(ii) If $x \geq 1$ is the critical region for testing $\theta = 2$ against $\theta = 1$, on the basis of a single observation from the population.
$f(x, \theta) = \theta e^{-\theta x}, 0 < x < \infty$.
Obtain the values of type I and type II errors. Also obtain the power of the test. 5

(c) A sample of 1000 students of class X were classified according to their intelligence and economic conditions and the data are given as follows :

		Intelligence			
		Excellent	**Good**	**Mediocre**	**Dull**
Economic condition	**Sound**	40	172	153	35
	Average	33	124	160	33
	Poor	25	82	121	22

Test whether there is any association between intelligence and economic condition, at 5% level of significance.
(Given X^2 at 5% l.o.s. for 6 d.f. = 12.592) 10

(d) Following is a random sample of size 5 from a continuous distribution
$(F_x(x))$: 0.65, 0.35, 0.85, 0.60, 0.90.
Using Kolmogorov-Smirnov test, test whether the sample can be regarded as drawn from a distribution which uniform on (0, 1). (Given $d_{5, 0.05} = 0.565$.) 10

SECTION – C

Marks

6. Answer the following sub-questions :

(a) Determine regression line of price on supply, hence estimate price when supply is 180 units, from the following information. x = supply, y = price in Rs. per unit, n = number of observations = 7.
$\Sigma (x - 150) = 119$, $\Sigma (y - 160) = 84$
$\Sigma (x - 150)^2 = 2835$, $\Sigma (y - 160)^2 = 2387$
$\Sigma (x - 150)(y - 160) = 525$.
Also find correlation coefficient between price and supply. **10**

(b) (i) Find correlation coefficient between x and y given that $n = 25$, $\Sigma x = 75$, $\Sigma y = 100$, $\Sigma x^2 = 250$, $\Sigma y^2 = 500$, $\Sigma xy = 325$. **5**

(ii) Ten competitors in a vocal music test are ranked by the two judges A and B as follows :

Judge A :	1,	6,	5,	10,	3,	2,	4,	9	,7	8
Judge B :	3,	5,	8,	4,	7,	10,	2,	1,	6,	9

Use the method of rank correlation to assess the degree of agreement between their judgement. **5**

(c) If $X_1, X_2, ..., X_k$ have a multinomial distribution with the parameters n and P_i $(i = 1, 2, ..., k)$ with $\Sigma p_i = 1$, obtain the joint probability mass function. Also obtain the moment generating function, hence show that $E(X_i) = np_i$, $V(X_i) = np_i(1-p_i)$ and $\text{cov}(X_i, X_j) = -np_ip_j$, $(i \neq j)$. **10**

(d) A survey of 800 persons on their reading habits revealed the following information, "224 read novels, 301 read detective stories, 150 read short stories, 125 read novels and detective stories, 72 read novels and short stories, 60 read detective stories and short stories and 32 read all the three".
Find the number of persons who read :

(i) novels, detective stories but no short stories,

(ii) only detective stories and

(iii) only short stories. **10**

7. Answer the following sub-questions :

(a) What do you mean by the regression equation of Y on X ? What is its purpose ? Derive the equation to the line of regression of Y on X. **10**

(b) (i) Show that the coefficient of correlation is, equal to geometric mean between the two coefficient of regressions and less than or equal to the arithmetic mean of the coefficients of regression. **5**

(ii) In a trivariate distribution

$r_{12} = 0.7$, $r_{13} = r_{23} = 0.5$.

Find $R_{1.23}$ and $r_{13.2}$. **5**

Marks

(c) For the bivariate normal distribution with p.d.f.

$$f(x,y)=c\exp\left[-\frac{8}{27}\left\{(x-7)^2-2(x-7)(y+5)+4(y+5)^2\right\}\right], -\infty<x,y<\infty,$$

determine the parameters. Also find the value of constant C. 10

(d) Can vaccination be regarded as a preventive measure for small pox from the data given below ?
'Of 1482 persons in a locality exposed to small-pox 368 in all were attacked'.
'Of 1482 persons, 343 had been vaccinated and of these only 35 were attacked' 10

SECTION – D

8. Answer the following sub-questions :

(a) (i) Explain the procedure of stratified random sampling, giving two practical situations. 5

(ii) Write short note on systematic sampling. 5

(b) Explain ratio and regression methods of estimation. Give the estimators of population means and the variances of estimators. Are these estimators of population means are unbiased, in both the methods ? 10

(c) Explain the principles in designing a good questionnaire. 10

(d) Explain meaning, need and importance of agricultural statistics. Also explain how it is collected in India. 10

9. Answer the following sub-questions :

(a) (i) Explain simple random sampling with replacement and without replacement. Also explain the methods of drawing a simple random sample. 5

(ii) Explain the method of drawing a sample using Cluster sampling. 5

(b) Define Systematic Sampling. Obtain the sampling variance of the mean based on systematic sample and compare it with the variance based on simple random sampling without replacement. 10

(c) Find the sample size so that an observed difference of 10% of the mean will be taken as significant at 5% level, the coefficient of variation being 12%. (Given z = 1.96 at 95% level). 10

(d) Explain "Population Statistics". State the importance and sources of population statistics. 10

SEAL