CSM – 68/16
Statistics
Paper – I

Time : 3 hours

Full Marks : 300

The figures in the right-hand margin indicate marks.

Candidates should attempt Q. No. 1 from Section – A and Q. No. 5 from Section – B which are compulsory and any ***three*** *of the remaining questions, selecting at least* ***one*** *from each Section.*

SECTION – A

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

$12 \times 5 = 60$

(a) Let $S = \{1, 2, 3, 4\}$ be the sample space, with probabilities of each point being $\frac{1}{4}$. Consider three events $A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{1, 4\}$. Check the pairwise and mutual independence of the events.

(b) Define the distribution function (df) of the rv X and mention its properties. The pdf of X is $f(x) = \frac{1}{b-a}$, $a < x < b$. Obtain the df of the rv X and sketch the graph.

(c) The joint probability mass function (pmf) of (X, Y) is given in the following table :

Y/X	1	2	3	4
1	1/4	1/8	1/16	1/16
2	1/16	1/16	1/4	1/8

(i) Find the marginal density function of X and Y.

(ii) Compute $P(X = 1 / Y = 3)$ and $P(Y = 2 / X = 2)$

(d) State the pmf of a binomial distribution. Derive its mean and variance. If $E(X) = \frac{5}{3}$ and $V(X) = \frac{10}{9}$, find the parameters of the binomial distribution.

(e) Define multiple regression model with the assumptions. Mention its applications.

(f) What are canonical variates and canonical correlations ? Mention their applications in Multivariate Analysis.

2. (a) A box contains 3 white balls and 4 black balls. Two balls are drawn at random from it, without replacement. Find the probability that (i) the second ball is black, given that the first ball is black and (ii) the second ball has the same color as the first ball.

(b) Define Chebychev's inequality. Mention its applications. Let $X \sim P(\lambda)$.

Verify that (i) $P\left(X \leq \frac{\lambda}{2}\right) \geq \frac{4}{\lambda}$

(ii) $P(X \geq 2\lambda) \leq \frac{1}{\lambda}$

(c) Let $f(x, y) = x + y, 0 < x < 1, 0 < y < 1$, be the joint pdf of (X, Y) :

(i) Check whether X and Y are independent.

(ii) Find the conditional mean of X given $Y = y$.

20×3 = 60

3. (a) Define the moment generating function (mgf) of X and state its properties. Derive the mgf of X with pdf $f(x, \theta) = \theta e^{-\theta x}$, $x > 0$, $\theta > 0$. If $X_1, X_2 \cdots\cdots X_n$ an iid rvs from $f(x, \theta)$, obtain the distribution $Y = X_1 + X_2 + \cdots\cdots X_n$.

(b) (i) Let $\{X_n\}$ be a sequence of rvs with

$$P(X_n = 1) = \frac{1}{n} \text{ and } P(X_n = 0) = 1 - \frac{1}{n}.$$

Examine whether WLLN holds.

(ii) Let $\{X_n\}$ be a sequence of rvs with

$$P(X_n = r) = \binom{n}{r} p^r q^{n-r}, \text{ and}$$

$$Y_n = \frac{X_n - np}{\sqrt{npq}}.$$

Prove that $\{Y_n\}$ converges to $N(0, 1)$ in distribution.

(c) Define convergence in probability and almost sure convergence.

If $X_n \xrightarrow{P} X$ and g is continuous real valued function, then show that $g(X_n) \xrightarrow{P} g(X)$.

20×3 = 60

4. (a) Define Hotelling's T^2-Statistics and Mahalanobis D^2-Statistics and state this relationship. Mention any one application of Hotelling's T^2-Statistics.

(b) State Gauss-Markov theorem. Using this theorem, obtain the best linear unbiased estimator of a linear parametric function in a multiple linear regression model.

(c) What are multiple and partial correlations ? Describe a method to compute them, given the correlation motion. $20 \times 3 = 60$

SECTION – B

5. Attempt any **three** of the following :

(a) (i) Explain the concepts : consistency, sufficiency, unbiasedness and efficiency.

(ii) Based on a random sample of size n from $N(\mu, \sigma^2)$, obtain the sufficient statistics for μ and σ^2. $10+10 = 20$

(b) (i) State Rao-Blackwell and Lehmann-Schiffe theorems. Mention their applications.

(ii) Describe the estimation of parameter in $f(x, \theta)$, by the method of scoring. Illustrate this method for Cauchy distribution.

10+10 = 20

(c) (i) Describe cumulative total method for the selection of n units from N units under PPSWR.

(ii) Distinguish between cluster sampling and two stage sampling. 10+10 = 20

(d) (i) What are the main effects and interaction effects in a 2^3 factorial experiment ? Describe Yates technique to compute the sum of squares in a 2^3 factorial experiment.

(ii) Define a BIBD. Establish any three relationships among the parameters of this design. 12+8 = 20

6. (a) Define exponential family of distributions and derive the sufficient statistic for θ, based on a random sample size n. Hence, obtain the sufficient statistics in (i) $P(\lambda)$ and (ii) $G(\alpha, \beta)$.

12

(b) Explain the maximum likelihood method of estimation. Obtain the MLEs of the parameter in (i) $N(\mu, \sigma^2)$ and (ii) $U(0, \theta)$. 12

(c) State Cramer-Rao inequality along with the regularity conditions. Obtain the CR lower bound for the variance of the unbiased estimator μ in $N(\mu, 1)$. 12

(d) Obtain OC and ASN functions for SPRT for testing $H_0 : p = p_0$, $H_1 : p = p_1$ $(p_1 > p_0)$, where p is a binomial proportion. 12

(e) Explain randomised and non-randomised tests with an example each. 12

7. (a) Describe the following non-parameter tests : 12

(i) Median test

(ii) Mann – Whitney test

(b) Describe the method of moments for estimation of parameters. Obtain MME of the parameter in $G(\alpha, \beta)$. 12

(c) Describe missing plot technique for an LSD and outline the analysis for testing the relevant hypothesis. 12

(d) Define : 12

(i) Loss and risk function

(ii) Prior and posterior distribution

(iii) Bayes and minimax estimators

(e) Based on a random sample of size n from $f(x, \theta) = \theta e^{-\theta x}$, $x > 0$, $\theta > 0$, derive the UMVUEs of $\frac{1}{\theta}$ and $\frac{1}{\theta^2}$. 12

8. (a) Explain the method of estimating the population total Y under Ratio and Regression methods of estimation. Compare the standard errors of these estimators. 12

(b) Describe Horvitz-Thompson estimator of Y under PPSWOR. Derive the variance expression for this estimator. 12

(c) Distinguish between partial and complete confounding. Illustrate the layout in a 2^3 factorial experiment. 12

(d) Discuss briefly the analysis of a 3^2 factorial experiment. 12

(e) Describe Warner's randomised response technique. 12