Candidate must not write on this margin.

CSM—64/23

PART—II/PAPER—VI

STATISTICS

PAPER—I

Time : 3 Hours

Full Marks : 250

The question paper contains **18 (Eighteen)** *questions in* **GROUP—A (12)** *and* **GROUP—B (06)** *together.*

GROUP—A

Candidates to attempt **10 (ten)** *questions within word limit of* **250**. *Each question carries* **15** *marks.*

1. Define sequence of sets. State and prove Borel-Cantelli lemma. Apply Borel-Cantelli lemma to prove that if X_1, X_2, ... are i.i.d. random variables with common mean μ and finite fourth moment, then $P\left\{\lim_{n\to\infty} \frac{S_n}{n} = \mu\right\} = 1$, where, $S_n = \sum_{i=1}^{n} X_i$.

2. Define the terms pairwise independence, mutual independence and independence with respect to multiple random variables. Let (X_1, X_2, X_3) be a random variable with joint PMF

$$f(x_1, x_2, x_3) = \begin{cases} \frac{1}{4}, & \text{if } (x_1, x_2, x_3) \in A \\ 0, & \text{otherwise} \end{cases}$$

where $A = \{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$

Are X_1, X_2, X_3 independent? Are X_1, X_2, X_3 pairwise independent? Are $X_1 + X_2$ and X_3 independent?

3. Define expectation of a random variable. Give an example of a random variable whose expectation does not exist. Justify your answer. Let X be a non-negative continuous random variable with distribution function F. Show that

$$E(X) = \int_0^\infty [1 - F(x)]dx$$

in the sense that, if either side exists, so the other two are equal.

4. Describe Negative Binomial Distribution and Chi-square distribution. Let $X : NB(r, p)$. Show that $2pX \xrightarrow{L} Y$ as $p \to 0$, where $Y : \chi^2(2r)$.

5. Let X be a random variable with PMF

$$P\{X = j\} = p_j, \quad j = 0,1,2,...$$

Set $P\{X > j\} = q_j$, $j = 0,1,2,...$ Write $Q(s) = \sum_{j=0}^{\infty} q_j s^j$. Show that

$$Q(s) = \frac{1 - P(s)}{1 - s} \text{ for } |s| < 1,$$

where $P(s)$ is the PGF of X. Find the mean and the variance of X (when they exist) in terms of Q.

6. State WLLN and Lindeberg form of CLT. Let $\{X_n\}$ be a sequence of independent random variables with PMF

$$P\left\{X_n = n^{\lambda}\right\} = P\left\{X_n = -n^{\lambda}\right\} = \frac{1}{2}, \; n = 1,2, \ldots ; \lambda > 0$$

Does WLLN hold? Does Lindeberg condition hold? Comment on the applicability of CLT in this case.

7. Consider the model $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$, $i = 1,\ldots,r$; $j = 1,\ldots,s$

where the errors ε_{ij} are independent and identically distributed with mean zero and variance σ^2. Express this model in the form

$$\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$$

Obtain the rank of matrix X. For a vector of real constants $\underline{\lambda}$, prove that the function $\psi = \underline{\lambda}'\underline{\beta} = \alpha_1$ is not estimable. Show that $\psi = \alpha_1 - \alpha_2$ is estimable.

8. Describe linear model for two-way classified data. Give a real-life situation where it can be applied. Derive least squares estimates of the parameters in the model. Obtain the test for testing overall significance of the model.

9. Define a curvilinear regression model. What is the need of orthogonal polynomials? Consider the following model :

$$y_i = \sum_{j=0}^{k} \alpha_j P_j(x_i) + \varepsilon_i, i = 1,2,\ldots,n$$

where $P_j(x_i)$, is an orthogonal polynomial of order j. Derive least squares estimator for α_j. Explain a test for testing significance of highest order term in the above model. Write expressions for first five orthogonal polynomials.

Candidate must not write on this margin.

10. Define a variance component model with linear structure. Derive MINQUE estimator for the variance component. State the properties of the derived estimator.

Candidate must not write on this margin.

11. Describe the problem of discrimination. Describe Fisher's approach to the problem of discrimination. Derive Fisher's best discriminant function. Find it for discriminating between $N_p\left(\underline{\mu}^{(1)}, \Sigma\right)$ and $N_p\left(\underline{\mu}^{(2)}, \Sigma\right)$ populations.

12. Describe m-factor orthogonal model. Explain principal component method for estimating factor loadings. In factor analysis with variance covariance matrix Σ, does a proper solution to the equations $\Sigma = LL' + \psi$ always exist? Justify your answer.

GROUP—B

Candidates to attempt **05** ***(five)*** *questions within word limit of* **300**.

Each question carries **20** *marks.*

13. Explain the concept of consistency and unbiasedness. State and prove invariance property of consistent estimator. Is an unbiased estimator invariant under linear transformation? Justify your answer. Let $X_1, X_2, \ldots, X_n$ be a random sample from $U(0, \theta)$. Discuss consistency and unbiasedness of the following estimators for θ :

$$T_1 = \frac{2}{n}\sum_{i=1}^{n} X_i;\ T_2 = \max\{X_1, X_2, \ldots, X_n\}$$

14. Explain the concept of prior and posterior distribution. Define conjugate prior. Give a suitable example. State the advantages of using conjugate prior. Define Bayes estimator. Based on a sample of size one, derive Bayes estimator for θ in Binomial (m, θ) assuming uniform prior for θ.

15. Describe Wilcoxon-Mann-Whitney U-test and median test for two-sample problem. Discuss the consistency and asymptotic normality of these tests.

16. Describe stratified sampling. Sate a real-life situation where stratified sampling can be preferred over simple random sampling. Explain proportional and optimal allocation in stratified sampling. Derive the expressions for sample mean under proportional and optimal allocation in stratified sampling. Compare these with simple random sampling with replacement. Comment on the results.

17. Describe sampling and non-sampling errors. Can sampling errors be controlled? Justify your answer. Explain Warner's randomised response technique. State its use.

18. Explain the need for split-plot designs. Give a real-life situation where split-plot designs are useful. Explain the construction and analysis split-plot designs. State the interpretations of results.

★★★

Candidate must not write on this margin.