Candidate must not write on this margin.

CSM—64/22

STATISTICS

PAPER—I

Time : 3 Hours

Full Marks : 250

The figures in the right-hand margin indicate marks.

Candidates should attempt ***any* 10 *(ten)*** *questions of* **GROUP—A** *with word limit of 250 words and should attempt* ***any* 5 *(five)*** *questions from* **GROUP—B** *with word limit of 300 words.*

GROUP—A

Attempt *any* **10 (ten)** questions from the following :

1. Consider the following bivariate PMF :

$$P[X = x, Y = y] = \frac{(x+y+k-1)!}{x!y!(k-1)!} p_1^x p_2^y (1 - p_1 - p_2)^k$$

where $x, y = 0, 1, \dots ; k \geq 1$ is an integer, $0 < p_1 < 1, 0 < p_2 < 1$ and $p_1 + p_2 < 1$.

(a) Find the marginal PMFs of X and Y.

(b) Find the conditional PMF of X given $Y = y$. 15

2. Let (X, Y) have the PDF $f(x, y) = \begin{cases} xe^{-x(1+y)} & , \quad x > 0, y > 0, \\ 0 & , \quad \text{o.w.} \end{cases}$

(a) Find $E(X|Y)$.

(b) Find $E(XY)$ by conditioning on X. 15

3. Define convergence in probability and almost sure convergence. Let $\{X_n\}$ be a sequence of random variables and X be a random variable.

(a) Show that $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X$.

(b) Is converse of statement in *(a)* true? Justify your answer. 15

4. *(a)* Define Probability Generating Function (PGF) of a random variable X. Describe computation of $E(X)$ and Var (X) using PGF.

(b) Find the PGF of random variable X whose PMF is

$$P[X = x] = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})} & , \quad x = 0, 1, 2 \ldots; \lambda > 0 \\ 0 & , \quad \text{o.w} \end{cases}$$

(c) Let X be an integer valued random variable with PGF $P(s)$. Let α and β be non-negative integers and write $Y = \alpha + \beta X$. Find the PGF of Y. 15

5. Describe exponential family of distributions. Check whether following family of distributions are exponential families :

(a) $f(x; \theta) = \dfrac{1}{2} e^{-|x-\theta|}, -\infty < x < \infty; -\infty < \theta < \infty$

(b) Binomial(n, θ), $0 < \theta < 1$ 15

6. *(a)* Define characteristic function and state its inversion property.

(b) Find the density function corresponding to the characteristic function $\phi(t) = \dfrac{1}{1+t^2}$.

(c) Let X be the random variable with the characteristic function $\phi(t) = \dfrac{1}{1+t^2}$. Using $\phi(t)$, find $E(X)$ and Var (X). 15

7. *(a)* Define a linear statistical model.

(b) Derive Least Squares Estimator (LSE) of the coefficients in the linear statistical model.

(c) Stating necessary conditions, establish that the LSE of the coefficients in the linear model is Best Linear Unbiased Estimator (BLUE). 15

8. *(a)* For a multiple linear regression model, derive a test of goodness-of-fit. State the associated ANOVA table.

(b) Explain the construction and application of Normal Probability Plot in a multiple linear regression model. 15

Candidate must not write on this margin.

9. Let X_1, X_2 and X_3 be three random variables.

(a) Define multiple correlation and partial correlation.

(b) Express multiple correlation coefficient $R_{1 \cdot 23}$ in terms of total and partial correlation coefficients. Justify your claim. 15

10. Let $\underline{X} = \left(\underline{X}^{(1)}, \underline{X}^{(2)}\right)' \sim N_p\left(\underline{\mu}, \Sigma\right)$.

(a) Derive the marginal distribution of $\underline{X}^{(2)}$.

(b) Derive the conditional distribution of $(\underline{X}^{(2)} \mid \underline{X}^{(1)})$. 15

11. (a) Define Mahalanobis D^2 and Hotelling's T^2 statistics.

(b) Establish relationship between Mahalanobis D^2 and Hotelling's T^2 statistics.

(c) Show that Hotelling's T^2 statistics is invariant under non-singular linear transformation.

(d) Explain the use of Hotelling's T^2 statistics in Profile Analysis. 15

12. (a) What is dimension reduction?

(b) Derive expressions for principal components and variance explained by them.

(c) For the Dispersion Matrix $\sum = \begin{pmatrix} 1 & 0 \cdot 5 \\ 0 \cdot 5 & 1 \end{pmatrix}$, perform PCA and interpret the results. 15

GROUP—B

Attempt *any* **5 (Five)** questions from the following :

13. Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\theta, 1)$ and let $\psi(\theta) = \theta^2$.

(a) Obtain Cramer-Rao lower bound for the variance of an unbiased estimator of $\psi(\theta)$.

(b) Obtain UMVUE of $\psi(\theta)$. Does it attain the Cramer-Rao lower bound? Justify. 20

Candidate must not write on this margin.

14. Let $X_1, X_2, ..., X_n$ be a random sample from the PDF

$$f(x;\theta)=\frac{1}{\beta}e^{-(x-\alpha)/\beta},\ \alpha<x<\infty; -\infty<\alpha<\infty \text{ and } \beta>0$$

Find the Maximum Likelihood Estimator of

(a) α and β;

(b) $P[X_1 \geq 1]$. 20

15. (a) Define Most Powerful (MP) Test.

(b) State and prove Neyman-Pearson Lemma.

(c) Let $X_1, X_2, ..., X_n$ be a random sample from the PDF

$$f(x;\theta)=\frac{1}{2}e^{-\frac{|x|}{\theta}}, -\infty<x<\infty; \theta>0$$

Find a size α MP test for testing $H_0:\theta=\theta_0$ vs. $H_1:\theta=\theta_1(>\theta_0)$. 20

16. (a) Define Balanced Incomplete Block Design (BIBD). Give a real life situation, where BIBD can be applied.

(b) Explain the construction of BIBD with a suitable choice of parameters.

(c) Explain the statistical analysis of BIBD. State the interpretations of results. 20

17. For a Probability Proportional to Size with Replacement (PPSWR), let y_i be the value of the unit drawn in the i^{th} draw and p_i be the corresponding selection probability, $i = 1, 2, ..., n$.

(a) Show that an unbiased estimator for the population total is

$$\widehat{Y}_{PPS}=\frac{1}{2}\Sigma\frac{y_i}{p_i}$$

Find expression for $\text{Var}(\widehat{Y}_{PPS})$.

(b) Show that an unbiased estimator of the gain due to PPSWR sampling as compared to SRSWR is $\frac{1}{n^2}\sum_{i=1}^{n}\left[N-\frac{1}{p_i}\right]\frac{y_i^2}{p_i}$, where N is population size. 20

18. Define UMPU test. Let $X_1, X_2, ..., X_n$ be a random sample from $N(\theta, 1)$.

(a) Find UMPU test of size α for testing $H_0:\theta=\theta_0$ vs. $H_1:\theta\neq\theta_0$.

(b) Find UMAU confidence interval θ at confidence level $1-\alpha$. 20

★★★