2021

STATISTICS — HONOURS

Fifth Paper

(Group - A)

Full Marks : 50

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Notation have their usual significance.

Unit - I

1. Attempt ***any two*** questions from the following : 5×2

(a) Suppose $R = (\rho_{ij})_{i,j=1(1)k}$ is the correlation matrix for a k-dimensional random vector. Show that $|R| \le 1$ and also interpret the case when $|R| = 1$.

(b) If a random vector $X = (X_1, X_2, ..., X_p)' \sim N_p\left(\mu, \Sigma\right)$, evaluate $E\left(e^{\sum_{j=i}^{p} X_j}\right)$.

(c) Let $X = (X_1, X_2, ..., X_p)' \sim N_p\left(\mu_1, I_p\right)$ independently of $Y = (Y_1, Y_2, ..., Y_p)' \sim N_p\left(\mu_2, \Sigma\right)$, $\Sigma > 0$.

Also let $U = \Sigma^{-\frac{1}{2}}\left(X - \mu_1\right)$ and $V = \Sigma^{-\frac{1}{2}}\left(X - \mu_2\right)$, then prove that $\mathrm{Var}(U'V) = p$.

(d) If $(N_1, N_2, ..., N_m) \sim$ Multinomial $(n \mid p_1, p_2, ..., p_m)$, find the condition mean and variance of N_i, given $\sum_{j \ne i}^{m} N_j = c$, for $0 \le c < n$.

2. Attempt ***any one*** question from the following :

(a) (i) Let X, Y and Z be independent random variables having same variance. Define the random variables $W_1 = \left(X - Z\right)/\sqrt{2}, W_2 = \left(X + Y + Z\right)/\sqrt{3}$ and $W_3 = \left(X + 2Y + Z\right)/\sqrt{6}$. Find the multiple correlation coefficient of W_1 on (W_2, W_3).

(ii) Suppose $(X_1, X_2, ..., X_p) \sim$ Multinomial $(n \mid \pi_1, \pi_2, ..., \pi_p)$ and c_i's are non-zero real numbers such that $\sum_{i=1}^{p} c_i \pi_{i=0}$. If $U = \sum_{i=1}^{p} c_i X_i$ and $V = \sum_{i=1}^{p} \left(X_i / c_i\right)$, verify whether two random variables U and V are linearly independent or not. 7+8

(b) (i) For any non null vector $\alpha \in \mathbb{R}'$, establish that $\alpha' X \sim N(\alpha'\mu, \alpha'\Omega\alpha) \Leftrightarrow X \sim N_p(\mu, \Omega)$ where X is a p-dimensional random vector and $\Omega > 0$.

(ii) Let $X_0, X_1, ..., X_n$ be $n + 1$ independent standard normal variables. Define the random variables

$$Y_j = \rho X_0 + \sqrt{(1-\rho^2)}X_j, j = 1,2,...,n.$$

Obtain the distribution of $Y = (Y_1, Y_2,...,Y_n)'$. Also express the partial correlation coefficient $\rho_{12\cdot 34...n}$ in terms of ρ. 7+8

Unit - II

3. Attempt ***any two*** questions from the following : 5×2

(a) Suppose in a group of n people selected at random from a community contains $m(< n)$ men of which m_0 are smokers. Among $(n - m)$ women in the group there are f_0 smokers. Construct Pearsonian Chi-square test to judge whether a man is more prone to smoking habit than a woman.

(b) Let T_n denote the mean of n independent random observations having common pdf $f_\theta(x) = 2(\theta - x)/\theta^2, 0 < x < \theta, \theta > 0.$

Applying delta method, show that $\sqrt{n}(\log T_n - \log(\theta/3)) \xrightarrow{D} N(0,1/2),\ as\ n \to \infty.$

(c) Based on n independent copies of $X \sim N(\mu, \sigma^2)$, obtain the large sample standard error of sample coefficient of variation.

(d) Find the asymptotic distribution of sample median of a random sample of size n drawn from the distribution with pdf $f_\theta(x) = 2\theta^2 / x^3, x > \theta, \theta > 0.$

4. Attempt ***any one*** question from the following :

(a) (i) Explain the term 'convergence in distribution'. If, as $n \to \infty, X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} Y$, prove that $X_n + Y_n \xrightarrow{D} X + Y$.

(ii) State the central limit theorem for a sequence of iid random variables. For a sequence $\{X_k\}_{k\geq 1}$ of independent Poisson (λ) variables prove that, as $n \to \infty$,

$$\sqrt{n}\left(\sqrt{S_n/n} - \lambda\sqrt{n/S_n}\right) \xrightarrow{D} N(0,1)$$

where $S_n = X_1 + X_2 + ... + X_n$. (2+6)+(2+5)

(b) (i) Based on two independent random samples of sizes n_1 and n_2 from $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively, describe a large sample test procedure for testing the null hypothesis $H_0 : \sigma_1 = \sigma_2$. If H_0 is rejected, then determine $100(1 - \alpha)\%$ confidence interval for $|\sigma_1 - \sigma_2|$.

(ii) Suppose $(X_1, X_2, ..., X_n)$ represents a random sample drawn from $N(\mu, \sigma^2)$ population. Derive the asymptotic distribution of sample r-th order central moment. (5+3)+7
