[This question paper contains 4 printed pages.]

Your Roll No...............

Sr. No. of Question Paper : 5776 **K**

Unique Paper Code : 2372014701

Name of the Paper : Multivariate Analysis

Name of the Course : **B.Sc. (Hons.) Statistics under NEP-UGCF**

Semester : VII

Duration : 3 Hours Maximum Marks : 90

Instructions for Candidates

1. Write your Roll No. on the top immediately on receipt of this question paper.
2. Question **1** is compulsory.
3. Attempt **six** questions in all.
4. Use of a non-programmable calculator is allowed.

1. Attempt any three parts:

   (i) Find Coefficient of correlation between X and Y for jointly distributed normal distribution

$$f(x,y) = \left(2\pi\sqrt{3}\right)^{-1} \exp\left[-\frac{\{(2x-y)^2+2xy\}}{6}\right]; -\infty < x, y, \infty.$$

   (ii) For the given the covariance matrix $\begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & 3 \\ 2 & -3 & 25 \end{bmatrix}$, obtain the standard deviation matrix and the correlation matrix.

(ii) For the given the covariance matrix $\begin{bmatrix} 4 & 1 & 2 \\ 1 & 9 & 3 \\ 2 & -3 & 25 \end{bmatrix}$, obtain the standard deviation matrix and the correlation matrix.

(iii) Data matrices from two populations π_1 and π_2 along with the sample covariance matrices are given by $X_1 = \begin{pmatrix} 3 & 2 \\ 1 & 2 \\ 2 & 5 \end{pmatrix}$, $S_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ and $X_2 = \begin{pmatrix} -2 & 5 \\ 0 & 3 \\ -1 & 1 \end{pmatrix}$, $S_2 = \begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$, respectively. Classify an observation $\underset{\sim}{x}_0' = (2.5 \quad 3.5)$ to π_1 or π_2, using sample discriminant function.

(iv) From the data relating to the yield of dry bark (X_1), height (X_2) and girth (X_3) for 18 cinchona plants, the following correlation coefficients were obtained:

$$r_{12} = 0.77\,, r_{13} = 0.72\,, r_{23} = 0.52\,.$$

Find the partial correlation coefficient $r_{12.3}$ and the multiple correlation coefficient $R_{1.23}$.

(5×3)

2. (i) If $(X, Y) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ then show that linear combination of X and Y is a normal variate. Prove that converse also holds good.

(ii) If $(X_i, Y_i) \sim BVN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, $i = 1,2, \ldots, n$ is a random sample of size n, then find the joint p.d.f. of $\bar{X}$ and $\bar{Y}$.

(8, 7)

3. (i) If X and Y are standard normal variates with coefficient of correlation ρ, then show that $(X + Y)$ and $(X - Y)$ are independently distributed.

(ii) Suppose the random vector $\boldsymbol{X}$ is distributed as $N_3(\boldsymbol{\mu}, \Sigma)$. Find the distribution of $\boldsymbol{AX}$, where $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$.

(8, 7)

4. (i) Derive the moment generating function of $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$.

(ii) Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ be a random sample of size n from p-variate normal distribution with mean $\underset{\sim}{\mu}$ and covariance matrix Σ. Then show that

(a) $\bar{\underset{\sim}{X}} \sim N_p\left(\underset{\sim}{\mu}, \frac{\Sigma}{n}\right)$

(b) $A = \sum_{\alpha=1}^{n}(\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})(\underset{\sim}{X}_\alpha - \bar{\underset{\sim}{X}})'$ is distributed as Wishart distribution with $(n-1)$ d.f.

(c) $\bar{\underset{\sim}{X}}$ and A are independently distributed. (6, 9)

5. (i) If $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$. Let $\underset{\sim}{X}$, $\underset{\sim}{\mu}$ and Σ be partitioned as follows:

$$\underset{\sim}{X} = \begin{pmatrix} \underset{\sim}{X}^{(1)}_{k\times 1} \\ \underset{\sim}{X}^{(2)}_{p-k\times 1} \end{pmatrix}, \ \underset{\sim}{\mu} == \begin{pmatrix} \underset{\sim}{\mu}^{(1)}_{k\times 1} \\ \underset{\sim}{\mu}^{(2)}_{p-k\times 1} \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

If $\Sigma_{12} = O = \Sigma_{21}$, then prove that $\underset{\sim}{X}^{(1)}_{k\times 1}$ and $\underset{\sim}{X}^{(2)}_{p-k\times 1}$ are independently distributed as multivariate normal with mean vector $\underset{\sim}{\mu}^{(1)}_{k\times 1}$ and $\underset{\sim}{\mu}^{(2)}_{p-k\times 1}$ and dispersion matrix Σ_{11} and Σ_{22}, respectively.

(ii) Let X_1, X_2 and X_3 are three variables measured from their respective means. Show that, with usual notations, the equation of the plane of regression X_1 on (X_2, X_3) is given by

$$\frac{X_1}{\sigma_1}\omega_{11} + \frac{X_2}{\sigma_2}\omega_{12} + \frac{X_3}{\sigma_3}\omega_{13} = 0,$$

where ω_{ij} is the cofactor of ith row and j th column of the determinant

$$\omega = \begin{vmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{vmatrix}.$$

(8, 7)

6. (i) Derive the discriminating procedure in the case of two p-variate normal population with different mean vectors but same dispersion matrix.

(ii) Explain briefly the concept of Hierarchical clustering, Agglomerative hierarchical method and divisive hierarchical method. Illustrate any one of the linkage methods of agglomerative hierarchical procedure using the distance matrix

$$D = \begin{array}{c c} & \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} & \begin{bmatrix} 0 & & & & \\ 9 & 0 & & & \\ 3 & 7 & 0 & & \\ 6 & 4 & 9 & 0 & \\ 11 & 10 & 2 & 8 & 0 \end{bmatrix} \end{array}$$

and draw dendrogram for distances between five objects.

(7, 8)

7. (i) If $\Sigma = \begin{bmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix}$ with eigenvalue-eigenvector pairs as:

$\lambda_1 = 5.83, e_1' = (0.383, -0.924, 0);$

$\lambda_2 = 2.00, e_2' = (0,0,1)$; and

$\lambda_3 = 0.17, e_3' = (0.924, 0.383, 0).$

Obtain the Principal Components. Also show that they are uncorrelated and have variances equal to the eigen value of Σ.

(ii) Describe an orthogonal factor model stating clearly the underlying assumption. Prove that the covariance structure of such a model is expressed by $\Sigma = LL' + \Psi$, defining clearly the ith communality and the specific variance.

(7, 8)