0,0,1,1,0,1,1,1,0,0,0,0,1,0,1,0,1,1,0,1,0,0,1,0,1,1,1,0,
0,1,1,0,0,0,1,0,1,1,0,0

Set up null and alternative hypotheses. Using a suitable test, can you say that the plant manager's suspicion is true at 5% level of significance?

(6,6)

8. (a) Discuss Mann-Whitney U test for the small, moderate and large samples. What happens when there are ties in large samples?

(b) For a SPRT of strength (α', β'), show the following inequalities :

(i) $\alpha' \leq \frac{\alpha}{1-\beta}$

(ii) $\alpha' + \beta' \leq \alpha + \beta$

(8,4)

[This question paper contains 8 printed pages.]

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

Sr. No. of Question Paper : 2991 **H**

Unique Paper Code : 32371602

Name of the Paper : Multivariate Analysis And Non-parametric Methods

Name of the Course : **B.Sc. (H) Statistics, under CBCS (LOCF)**

Semester : VI

Duration : 3 Hours Maximum Marks : 75

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 **five** more questions selecting **three** questions from **Section A** and **two** questions from **Section B**.

4. Use of a non-programmable scientific calculator is allowed.

P.T.O.

1. Attempt all parts :

(a) Fill in the blanks :

(i) Principal component analysis is a ________ technique.

(ii) If $r_{12} = r_{13} = r_{23} = 0.6$ then the value of $1 - R^2_{1.23} =$ ________ .

(iii) The relation between two values of U statistic is ________ .

(iv) ________ is a non-parametric test for two related samples.

(v) Variance-covariance matrix of $\underset{\sim}{X}$ is ________ and ________ matrix.

(b) Differentiate between nominal and ordinal scales of measurement along with suitable examples.

(c) If $\underset{\sim}{X} \sim N_3(\underset{\sim}{\mu}, \Sigma)$ where

$$\underset{\sim}{\mu} = \begin{pmatrix} 2.5 \\ 3.8 \\ 4.7 \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} 3 & -2 & -1 \\ -2 & 5 & 3 \\ -1 & 3 & 14 \end{pmatrix}$$

then obtain the distribution of $\underset{\sim}{X}^{(1)} = \begin{pmatrix} X_1 \\ X_3 \end{pmatrix}$ and write its pdf.

(i) $P_{\theta_0}(A_0) \geq \frac{1}{B} P_{\theta_1}(A_0)$, where A_0 is the event that ultimately H_0 is accepted.

(ii) $P_{\theta_0}(A_1) \leq \frac{1}{A} P_{\theta_1}(A_1)$, where A_1 is the event that ultimately H_1 is accepted.

(b) Let $X_1, X_2 \dots$ be the iid sequence of random variables from exponential distribution with parameter θ. Obtain the SPRT for testing $H_0: \theta = \theta_0$ against $H_1: \theta = \theta_1 (< \theta_0)$. Also, obtain the O.C. function and ASN function for the same. (6,6)

7. (a) What are the assumptions of non-parametric test? Discuss one sample runs test of randomness.

(b) At a soda factory, the amount of soda put into each 12 ounce bottle by the bottling machine varies slightly. The plant manager suspects that the machine has a non-random pattern of overfilling the bottles (denoted by 1) and underfilling the bottles (denoted by 0). The following result was obtained :

(b) Assume $\underset{\sim}{X} \sim N_4(\underset{\sim}{\mu}, \Sigma)$ where $\underset{\sim}{\mu} = \begin{pmatrix} 1 \\ 2 \\ 4 \\ 6 \end{pmatrix}$ and

$$\Sigma = \begin{pmatrix} 3 & 3 & 1 & 2 \\ 3 & 6 & 1 & -2 \\ 1 & 1 & 6 & 0 \\ 2 & -2 & 0 & 9 \end{pmatrix}.$$

Let $\underset{\sim}{X}^{(1)} = \begin{pmatrix} X_1 \\ X_4 \end{pmatrix}$ and $\underset{\sim}{X}^{(2)} = \begin{pmatrix} X_2 \\ X_3 \end{pmatrix}$.

(i) Obtain the conditional distribution of

$$\left(\underset{\sim}{X}^{(2)} | \underset{\sim}{X}^{(1)} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}\right).$$

(ii) Define $Y = X_1 + X_3 - X_4$. Obtain the distribution of Y. (4,8)

SECTION B

6. (a) Let X_1, X_2 .... be the iid sequence of random variables with pdf/pmf $f_\theta(x)$. Describe Wald's SPRT of strength (α,β) for testing H_0: $\theta = \theta_0$ against H_1: $\theta = \theta_1$. Further, establish the following inequalities :

(d) (i) If (X,Y) ~ BVN(2,3,16,9,0.6), then obtain the distribution of $U = \frac{X}{2} - \frac{Y}{3}$.

(ii) While performing SPRT, obtain the values of A and B when $\alpha = 0.3$, $\beta = 0.25$.

(1×5,3,3,2×2)

SECTION A

2. (a) If $(X,Y) \sim BVN(0, 0, \sigma_1^2, \sigma_2^2, \rho)$ then show that

$$U = \frac{X}{\sigma_1} + \frac{Y}{\sigma_2} \text{ and } V = \frac{X}{\sigma_1} - \frac{Y}{\sigma_2}$$ are independently normally distributed with means 0 and variances $2(1 + \rho)$ and $2(1 - \rho)$ respectively.

(b) If $(X, Y) \sim BVN(0, 0, 1, 1, \rho)$ then obtain the correlation coefficient between X^2 and Y^2.

(8,4)

3. (a) Let $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \Sigma)$. Let $\underset{\sim}{X}$ be partitioned into

$$\underset{\sim}{X}^{(1)} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_q \end{pmatrix} \text{ and } \underset{\sim}{X}^{(2)} = \begin{pmatrix} X_{q+1} \\ X_{q+2} \\ \vdots \\ X_p \end{pmatrix},$$

correspondingly, mean vector and variance-covariance matrix be partitioned as $\mu = \begin{pmatrix} \underset{\sim}{\mu}^{(1)} \\ \underset{\sim}{\mu}^{(2)} \end{pmatrix}$ and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$. Prove that $\underset{\sim}{X}^{(1)}$ and $\underset{\sim}{X}^{(2)}$ are independently normally distributed iff $\Sigma_{12} = \Sigma_{21}' = O$..

(b) Consider the joint pdf of X and Y as

$$f(x,y) = \frac{1}{2}[f_1(x,y) + f_2(x,y)]; \ (x,y) \in R^2$$

where $f_1(x, y)$ is the pdf of BVN(0, 0, 1, 1, ρ) and $f_2(x,y)$ is the pdf of BVN(0, 0, 1, 1, −ρ) then show that the marginal distributions of X and Y are normal. Further, show that X and Y are not independent even when ρ(X,Y) = 0. (7,5)

4. (a) If (X, Y) ~ BVN($\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho$) with the pdf

$$f(x,y) = k \exp\left[-\frac{1}{102}[(y+3)^2 - (2.8)(y+3)(x+1) + 4(x+1)^2]\right];$$

$(x,y) \in R^2$

(i) Obtain the estimates of the parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and ρ.

(ii) Find E(X|Y = 3) and Var(X|Y = 3).

(b) Given two multivariate normal distributions namely $N_p(\underset{\sim}{\mu}^{(1)}, \Sigma)$ and $N_p(\underset{\sim}{\mu}^{(2)}, \Sigma)$ as two populations and a measurement on a p -component vector of variables defined by $\underset{\sim}{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$, explain the discriminant procedure for classifying $\underset{\sim}{x}$ vector into one of the two populations. Further,

(i) Show that $\underset{\sim}{a}'\underset{\sim}{x} = \underset{\sim}{\delta}'\Sigma^{-1}\underset{\sim}{x}$ where

$\underset{\sim}{\delta} = \left(\underset{\sim}{\mu}^{(1)} - \underset{\sim}{\mu}^{(2)}\right)$ and $\underset{\sim}{a} = \Sigma^{-1}\underset{\sim}{\delta}$.

(ii) Establish that $h = \frac{1}{2}\left[\underset{\sim}{a}'\left(\underset{\sim}{\mu}^{(1)} + \underset{\sim}{\mu}^{(2)}\right)\right]$.

(6,6)

5. (a) If r_{12} and r_{13} are given, show that r_{23} must lie in the range

$$r_{12}r_{13} \pm (1 - r_{12}^2 - r_{13}^2 + r_{12}^2 r_{13}^2)^{1/2}.$$

If $r_{12} = 0.75$ and $r_{13} = -0.75$, then obtain the limits of r_{23}.