LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – **STATISTICS**

THIRD SEMESTER – **NOVEMBER 2022**

PST 3501 – MULTIVARIATE ANALYSIS

Date: 23-11-2022 Dept. No. Max. : 100 Marks
Time: 09:00 AM - 12:00 NOON

Section - A

Answer ALL the Questions **10 x 2 = 20 Marks**

1. Define Multivariate Normal distribution.
2. Define variance-covariance matrix and correlation matrix.
3. Provide the formula to calculate first order partial correlation coefficient from the values of sample correlation matrix.
4. Define Wilk's Lambda and state its use.
5. What are Common factors and Specific factors in a orthogonal factor model ?
6. How factor naming is done? explain with an example
7. Define discriminant and classification step.
8. State any two association measures and any two distance measures to calculate the similarity between two observations.
9. Explain Ward's linkage method.
10. What is a Dendrogram? State its use.

Section - B

Answer any FIVE Questions **5 x 8 = 40 Marks**

11. Obtain the probability density function of Bivariate normal distribution by substituting p=2 in multivariate normal density
12. Derive the characteristic function of multivariate normal distribution
13. If $X^{(1)} \sim N_{p_1}(\mu^{(1)} \quad, \Sigma_{11})$ *and* $X^{(2)} \sim N_{p_2}(\mu^{(2)} \quad, \Sigma_{22})$

    $X^{(1)} \coprod X^{(2)}$ then $\begin{pmatrix} X^{(1)} \\ X^{(2)} \end{pmatrix} \sim N_{p_{1+p_2}} \left[\begin{pmatrix} \mu^{(1)} \\ \mu^{(2)} \end{pmatrix} \quad, \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} \right]$
14. a) If $X \sim N_p(\mu, \Sigma)$ *then show that* $(X-\mu)'\Sigma^{-1}(X-\mu) \sim \chi^2{}_{(p)}$

    b) Discuss the method to detect outliers in multidimensional data using generalized squared distance.
15. Obtain MANOVA table and Wilk's Lambda for testing equality of three mean vectors based on the data given below

    $sample\ from\ \pi_1: \begin{bmatrix} 9 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 \\ 2 \end{bmatrix}, \begin{bmatrix} 9 \\ 7 \end{bmatrix}$, $sample\ from\ \pi_2: \begin{bmatrix} 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $sample\ from\ \pi_3: \begin{bmatrix} 3 \\ 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 9 \end{bmatrix}, \begin{bmatrix} 2 \\ 7 \end{bmatrix}$
16. Establish Cauchy-Schwarz Inequality for px1 vectors and use the result to establish extended Cauchy-Schwarz Inequality
17. Explain Orthogonal Factor Model and illustrate Varimax rotation with an example
18. a) Explain expected cost of misclassification for classifying two population (3)

    b) Discuss the minimum ECM for two normal population with $\Sigma_1 = \Sigma_2 = \Sigma$ (5)

Section - C

Answer any TWO Questions **2 x 20 = 40 Marks**

19. Determine a) Mean Vector b) Var-Cov matrix c) Correlation matrix and d) Multiple Correlation $R_{1.234}$ based on the data given below **(2 + 10 + 5 + 3)**

X_1 :	91	95	81	83	76	58	89	79	83	74
X_2 :	55	68	72	91	49	78	54	76	69	73
X_3 :	65	67	89	69	78	87	94	79	96	59
X_4 :	72	64	85	65	75	58	48	72	91	82

20. Determine the three principal component equations based on the var-cov matrix given below and also determine the proportion of variance explained by each principal component **(6 +6 + 6 + 2)**

$$\boldsymbol{\Sigma} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

21. Perform Hierarchical Clustering based on a) Single Linkage b) Complete Linkage and c) Average Linkage using the distance matrix given below and obtain the dendrogram based on the three methods **(5 +5 +5 +5)**

$$\boldsymbol{D} = \begin{bmatrix} 0 & & & & \\ 9 & 0 & & & \\ 3 & 7 & 0 & & \\ 6 & 5 & 9 & 0 & \\ 11 & 10 & 2 & 8 & 0 \end{bmatrix}$$

22. **a)** Determine Fisher's sample discriminant function for the three population based on the data given below **(15 Marks)**

$\pi_1 \; (n_1 = 3)$ $\quad$ $\pi_2 \; (n_2 = 3)$ $\quad$ $\pi_3 \; (n_3 = 3)$

$$\boldsymbol{X_1} = \begin{bmatrix} -2 & 5 \\ 0 & 3 \\ -1 & 1 \end{bmatrix} \qquad \boldsymbol{X_2} = \begin{bmatrix} 0 & 6 \\ 2 & 4 \\ 1 & 2 \end{bmatrix} \qquad \boldsymbol{X_3} = \begin{bmatrix} 1 & -2 \\ 0 & 0 \\ -1 & -4 \end{bmatrix}$$

b) Classify a new observation $\boldsymbol{x_0}' = \begin{bmatrix} 1 & 3 \end{bmatrix}$ into one of the three population using Fisher classification procedure **(5 Marks)**

@@@@@@