 LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

**M.Sc.** DEGREE EXAMINATION - **STATISTICS**

THIRD SEMESTER – NOVEMBER 2012

# ST 3811 - MULTIVARIATE ANALYSIS

Date : 01/11/2012 Dept. No. Max. : 100 Marks

Time : 9:00 - 12:00

SECTION – A

Answer ALL the following questions: (10 x 2 = 20 marks)

1. Distinguish between ‘Data Exploration’ and ‘Confirmatory Analysis’.
2. Define Variance-Covariance matrix and Generalized Variance of a random vector.
3. Give the motivation for ‘Statistical Distance’ and express its form.
4. Define ‘partial correlation coefficient’ and give the expression for it in a trivariate normal distribution.
5. Explain the objectives of Principal Component Analysis.
6. Explain the ‘Varimax’ criterion for factor rotation.
7. Distinguish between ‘Agglomerative’ and ‘Divisive’ methods in clustering.
8. Present the decomposition of the total sum of squares and cross products in one way MANOVA.
9. State the test for significance of correlation coefficient in a bivariate normal population.
10. Define Wishart distribution.

SECTION – B

Answer any FIVE Questions (5 x 8 = 40 marks)

1. Find the mean vector and var-cov matrix of **X** = (X, Y) whose p.d.f. is

*f* (*x , y*) = 

1. Describe p-p plot and q-q plot and state how the multivariate normality assumption is verified.
2. Let X =  ~ Np (μ,Σ) and μ and Σ be correspondingly partitioned as  and , Xi be the ith component of X(1) and let **β** = σ(i)  where σ(i) is the ith row of . Derive an expression for multiple correlation coefficient between Xi and X(2).
3. Derive the moment generating function of multivariate normal distribution.
4. Derive the T2 test for hypothesis concerning the mean vector of a multivariate normal population using the likelihood ratio criterion.
5. Define Principal Components and extract the same for a given random vector, stating the lemma on maximization of quadratic forms (without proof).
6. Present Fisher’s method of discriminating two populations and derive the linear discriminant function. Explain the classification rule based on it.
7. Describe Hierarchical clustering of objects and its algorithm giving figurative depiction of three linkage methods.

SECTION – C

Answer any TWO Questions: (2 x 20 = 40 marks)

1. (a) Explain the ‘lowess’ curve enhancement of a scatter plot.

(b) Consider the partition in Q. No. (13). If **X(1)** and **X(2)** are uncorrelated, then show that their distributions are multivariate normal of appropriate dimensions. Also, establish the same result even when **X(1)** and **X(2)** are correlated. (7+13)

1. (a) Present the ‘Orthogonal Factor Model’ and develop the ideas of ‘communality’ and ‘specific variance’.

(b) Explain the ‘Principal Factor’ Method of Factor Analysis. Bring out the approach to ‘Reduction of factors’ and ‘Decision on number of factors’. (10+10)

1. (a) Carry out the ‘single linkage’ process for clustering six items whose distance matrix is given below (Dendrogram not required):



(b) Derive the expression for ‘Expected Cost of Misclassification’ in the case of two

populations and obtain the Minimum ECM Rule. Discuss the special cases of equal prior

probabilities and equal misclassification costs. (10+10)

1. (a) Derive the MLEs of the parameters of multivariate normal distribution.

(b) Establish the independence of the sample mean vector and sample var-cov matrix of a sample from multivariate normal distribution. (10+10)

\*\*\*\*\*\*