LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

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

THIRD SEMESTER – NOVEMBER 2010

# ST 3811 - MULTIVARIATE ANALYSIS

Date : 29-10-10 Dept. No. Max. : 100 Marks

Time : 9:00 - 12:00

**SECTION – A**

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

1. State and briefly explain any one broad objective for which multivariate analysis techniques are used.
2. Define Variance-Covariance matrix of a random vector.
3. Explain ‘Bubble Plot’ for three-dimensional data display.
4. Define ‘Partial Correlation Coefficient’.
5. Explain the dimensionality reduction role of Principal components.
6. Distinguish between ‘Clustering’ and ‘Classification’.
7. State any two similarity measures for pairs of items when variables are binary.
8. State any one of the commonly employed scalings on the coefficient vector of the Fisher’s linear discriminant function and the interpretative use of the scaling.
9. State the Hotelling’s T2 statistic for testing of hypothesis about the mean vector of a multivariate normal population.
10. State the test for significance of the correlation coefficient in a bivariate normal population.

SECTION – B

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

1. If is the sample mean vector and Sn is the sample var-cov matrix based on a random sample of size ‘n’ from a p-variate distribution with mean vector μ and Var-cov matrix Σ, show that Var-Cov() = Σ /n and Sn is a biased estimator of Σ.
2. Explain the construction of ‘probability plots’ in general and the investigation of multivariate normality assumption in particular.
3. If X ~ Np(μ, Σ ) and C is a non-singular matrix of order p x p, show that CX ~ Np ( Cμ, CΣCT ). Hence, deduce the distribution of DX where D is a q x p matrix with rank q (≤ p).
4. Under usual notations of partitions of a random vector and its var-cov matrix, with  = , show that ρ( Xi ,X(2) ) ≥ ρ( Xi , X(2) ) for every vector.
5. Develop the MANOVA for comparing mean vectors of a number of normal populations and explain the test procedure for the same.
6. Establish the relationship of principal components to the eigen values and eigen vectors of the var-cov matrix of the underlying random vector.
7. State the three linkage methods for hierarchical clustering. Present a figurative display of the measure of between-cluster distances in each method.
8. Derive the ‘Minimum ECM Rule’ for classification involving two multivariate normal populations with a common var-cov matrix.

(P.T.O)

SECTION – C

Answer any TWO questions (2 x 20 = 40 marks)

1. (a) Describe the enhancement of scatter plots with ‘*lowess’* curves.

(b) Develop the multivariate normal density function. (8 +12)

1. (a) Derive the MLE’s of the parameters μ, Σ of multivariate normal distribution.

(b) Show that these MLEs are independently distributed. (12 + 8)

1. Present the motivation, definition and derivation of Fisher’s (multiple) discriminant functions.
2. (a) Present the ‘Orthogonal Factor Model’ in detail and develop the notions of ‘Communality’ and ‘Specific Variance’.

(b) Bring out the ambiguity in a factor model, the rationale for ‘factor rotation’ and explain the varimax criterion for factor rotation. (12 + 8)

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