LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

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

THIRD SEMESTER – NOVEMBER 2017

16PST3MC01/ST3815 - MULTIVARIATE ANALYSIS

Date: 01-11-2017 Dept. No. Max. : 100 Marks
Time: 09:00-12:00

SECTION – A

Answer ALL the questions (10 x 2 = 20 Marks)

1. Let X,Y and Z have trivariate normal distribution with null mean vector and Covariance matrix $\begin{bmatrix} 2 & 5 & 0 \\ 5 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$, find the distribution of X+Y.
2. If $X'=(X_1, X_2) \sim N_2\left\{\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ then write the characteristic function of the marginal distribution of X_1.
3. Explain use of the partial and multiple correlation coefficients.
4. Give an example in the bivariate situation that, the marginal distributions are normal but the bivariate distribution is not.
5. Describe a) Common factor b) Communality.
6. Find the maximum likelihood estimates of the 2X1 mean vector « and 2X2 covariance matrix Σ based on random sample $X' = \begin{pmatrix} 6 & 8 & 10 & 8 \\ 12 & 8 & 14 & 14 \end{pmatrix}$ from bivariate population.
7. Let $(X_i, Y_i)'$ i = 1,2,3 be independently distributed each according to $N_2\left\{\begin{pmatrix} « \\ ¦ \end{pmatrix}, \begin{pmatrix} ±_{xx} & ±_{xy} \\ ±_{yx} & ±_{yy} \end{pmatrix}\right\}$. Find the distribution of $(\bar{X}, \bar{Y})'$.
8. Briefly explain K – means method in clustering.
9. What is the purpose of Multidimensional Scaling?
10. Write a short note on data mining.

SECTION– B

Answer anyFIVE questions (5X8=40 Marks)

11. Obtain the maximum likelihood estimator Σ of p-variate normal distribution.
12. Derive the characteristic function of multivariate normal distribution.

13. Explain the procedure for testing the equality of dispersion matrices of multivariate normal distributions.
14. Obtain the linear function to allocate an object to one of the two given normal populations.
15. Let $X \sim N_p(«, \Sigma)$. If $X^{(1)}$ and $X^{(2)}$ are two sub vectors of X then obtain the conditional distribution of $X^{(1)}$ given $X^{(2)}$.
16. Giving suitable examples explain how factor scores are used in data analysis.
17. Explain the method of extracting canonical correlations and their variables from a dispersion matrix.
18. Outline single linkage and complete linkage procedures with an example.

SECTION– C

Answer anyTWO questions **(2 X 20 =40 Marks)**

19. a) Derive the distribution function of the generalized T^2 – Statistic.

    b) Test $« = (0 \quad 0)'$ at level 0.05, in a bivariate normal population with $±_{11} = ±_{22} = 10$ and $±_{12} = -4$, using the sample mean vector $\bar{x} = (7 \quad -3)'$ based on sample size 20. (15+5)

20. a) What are the principal components? Outline the procedure to extract principal components from a given dispersion matrix.

    b) What is the difference between classification problem into two classes and testing problem? (15+5)

21. Consider the two data sets from populations Π_1 *and* Π_2 respectively,

$$X_1 = \begin{pmatrix} 3 & 7 \\ 2 & 4 \\ 4 & 7 \end{pmatrix} \quad and \quad X_2 = \begin{pmatrix} 6 & 9 \\ 5 & 7 \\ 4 & 8 \end{pmatrix}$$

for which $\bar{x}_1 = (3 \ \ 6)'$, $\bar{x}_2 = (5 \ \ 8)'$ *and* $S_{pooled} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

a) Calculate the linear discriminant function.

b) Classify the observation $x_0' = (2 \ 7)$ to population $®_1$ or population $®_2$ using the decision rule with equal priors and equal costs. (14+6)

22. Write short notes on:-

    a) Repeated Measurements Design

    b) Correspondence Analysis (10+10)

$$$$$$$$$