LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

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

THIRD SEMESTER – NOVEMBER 2015

ST 3815 - MULTIVARIATE ANALYSIS

Date : 03/11/2015 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. Write the statistic used to test the hypothesis is $H : °_{12.3} = 0$ in a bivariate normal distribution.
3. Mention any two properties of multivariate normal distribution.
4. Write down the characteristic function of a multivariate normal distribution.
5. Explain use of the partial and multiple correlation coefficients.
6. Define Hotelling's T^2 – Statistics. How is it related to Mahalanobis' D^2?
7. Give an example in the bivariate situation that, the marginal distributions are normal but the bivariate distribution is not.
8. Outline the use of Discriminant analysis.
9. Find the maximum likelihood estimates of the 2*1 mean vector « and 2*2 covariance matrix Σ based on random sample $X' = \begin{pmatrix} 6 & 8 & 10 & 8 \\ 12 & 8 & 14 & 14 \end{pmatrix}$ from bivariate population.
10. Write down any four similarity measures used in cluster analysis.

PART– B

Answer anyFIVE questions **(5X8=40 marks)**

11. Obtain the maximum likelihood estimator Σ of p-variate normal distribution.
12. Let $Y \sim N_p(0, \Sigma)$. Show that $Y'\Sigma^{-1}Y$ has ¡ ^{2} distribution.
13. Obtain the rule to assign an observation of unknown origin to one of two p-variate normal populations having the same dispersion matrix.
14. Show that the sample generalized variance is zero if and only if the rows of \ the matrix of deviation are linearly dependent.
15. Let $X \sim N_p($«$, \Sigma)$. If $X^{(1)}$ and $X^{(2)}$ are two subvectors of X, obtain the conditional distribution of $X^{(1)}$ given $X^{(2)}$.

16. Giving suitable examples explain how factor scores are used in data analysis.

17. Let $(X_i, Y_i)'$ i = 1,2,3 be independently distributed each according to bivariate normal with mean vector and covariance matrix as given below. Find the joint distribution of the six variables. Also find the joint distribution of $\overline{X}$ *and* $\overline{Y}$.

Mean Vector: $(«, ²)'$, covariance matrix: $\begin{pmatrix} ±_{xx} & ±_{xy} \\ ±_{yx} & ±_{yy} \end{pmatrix}$.

18. Prove that the extraction of principal components from a dispersion matrix is the study of characteristic roots and vectors of the same matrix.

PART– C

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

19. a) If $X \sim N_p(«, \Sigma)$ then prove that $Z = DX \sim N_p(D«, D\Sigma D')$ where D is qxp matrix rank q≤p.

b) Consider a multivariate normal distribution of X with

$$« = \begin{pmatrix} 8 \\ -2 \\ 0 \\ 3 \end{pmatrix} \quad , \quad \Sigma = \begin{pmatrix} 7 & 5 & 1 & 4 \\ 5 & 4 & 8 & -6 \\ 1 & 8 & 3 & 7 \\ 4 & -6 & 7 & 2 \end{pmatrix}.$$

Find i) the conditional distribution of $(X_2, X_4) \mid (X_1, X_3)$.

ii) $±_{33.42}$ (10+10)

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

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

21. 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 $±_{11} = ±_{22} = 5$ and $±_{12} = -2$, using the sample mean vector $\bar{x} = (7 \quad -3)'$ based on sample size 10. (15+5)

22. a) Outline single linkage and complete linkage clustering procedures with an example.

b) Prove that under some assumptions (to be stated), variance covariance matrix can be written as $\Sigma = LL' + \Psi$ in the factor analysis model. Also discuss the effect of an orthogonal transformation. (8+12)

****