LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – STATISTICS

THIRD SEMESTER – NOVEMBER 2014

ST 3815 - MULTIVARIATE ANALYSIS

Date : 30/10/2014 Dept. No. Max. : 100 Marks
Time : 09:00-12:00

PART – A

ANSWER *ALL* QUESTIONS (10 x 2 = 20)

1. Let X, Y and Z have trivariate normal distribution with null mean vector and covariance matrix $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$
   Find the distribution of X+Y
2. If $X \sim N_2\left[\begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right]$ then find density function of marginal distribution of X.
3. In a multivariate normal distribution, show that every linear combinations of the component variables of the random vector is normal.
4. Write down the characteristic function of a multivariate normal distribution.
5. Explain the use of partial and multiple correlation co-efficients.
6. Define Hotelling's T^2 – statistic. 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. What are canonical correlation coefficients and canonical variables?
10. Write down any four measures used in cluster analysis.

PART – B

ANSWER *ANY FIVE* QUESTIONS (5 x 8 = 40)

11. Obtain the maximum likelihood estimator of Σ of p – variate normal distribution.
12. Let $Y \sim N_p[0, \Sigma]$. Show that $Y'\Sigma^{-1}Y$ has χ^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. Outline Single linkage and complete linkage clustering procedures with an example.
15. Let $X \sim N_p(\mu, \Sigma)$. If $X^{(1)}$ and $X^{(2)}$ are two sub vectors 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 six variables. Also find the joint distribution of $\bar{x}$ and $\bar{y}$.
    Mean vector: $(\mu, \tau)'$, covariance matrix: $\begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}$
18. Write short notes on repeated measurement design.

PART – C

ANSWER *ANY TWO* QUESTIONS **(2 x 20 = 40)**

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

    b) Consider a multivariate normal distribution of X with

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

    Find i) The conditional distribution of $(X_2, X_4)/(X_1, X_3)$

    ii) $\sigma_{33.42}$ (10+10)

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

    b) What is the difference between classification problem in to two classes and testing problem. (14+6)

21. a) Derive the distribution function of the generalized T^2 – statistic.

    b) Test $\mu = (0 \quad 0)'$ at level 0.05, in a Bivariate normal population with $\sigma_{11} = \sigma_{22} = 5$ and $\sigma_{12} = -2$, using the sample mean vector $\bar{x} = (7 \quad -3)'$ based on sample size 10. (15+5)

22. a) Explain the method of extracting canonical correlations and their variables from a dispersion matrix.

    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)