LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

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

THIRD SEMESTER – **NOVEMBER 2018**

16/17PST3MC01/ 3815 – MULTIVARIATE ANALYSIS

Date: 25-10-2018　　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. In a multivariate normal distribution, show that every linear combination of the component variables of the random vector is normal. Is the converse true?
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. Comment on repeated measurements design.
6. Explain the classification problem into two classes.
7. Let $(X_i, Y_i)'$ i = 1,2,3 be independently distributed each according to $N_2\left\{\begin{pmatrix} \mu \\ \eta \end{pmatrix}, \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix}\right\}$.

   Find the distribution of $(\bar{X}, \bar{Y})'$.
8. Write down any four similarity measures used in cluster analysis.
9. What is the purpose of Multidimensional Scaling?
10. Write a short note on data mining.

SECTION– B

AnsweranyFIVE questions **(5X8=40 Marks)**

11. Obtain the maximum likelihood estimator Σ of p-variate normal distribution.
12. Find the multiple correlation coefficient between X_1 and $X_2, X_3, \ldots, X_p$. Prove that the conditional variance of X_1 given the rest of the variables cannot be greater than unconditional variance of X_1
13. Define Partial correlation coefficient between X_i and X_j . Also prove that

$$\rho_{12.3} = \frac{\rho_{12} - \rho_{13}\rho_{23}}{\sqrt{1-\rho_{13}^2}\sqrt{1-\rho_{23}^2}} .$$

14. Explain the procedure for testing the equality of dispersion matrices of multivariate normal distributions.
15. Obtain the linear function to allocate an object to one of the two given normal populations.
16. Giving suitable examples explain how factor scores are used in data analysis.
17. Obtain the rule to assign an observation of unknown origin to one of two p – variate normal populations having the same dispersion matrix.
18. Describe Profile Analysis with an example.

SECTION– C

AnsweranyTWO questions **(2 X 20 =40 Marks)**

19. a) Define generalized variance.

    b) Show that the sample generalized variance is zero if and only if the rows of the matrix of deviation are linearly dependent.

    c) Find the covariance matrix of the multivariate normal distribution which has the quadratic form $2x_1^{\,2} + x_2^{\,2} + 4x_3^{\,2} - x_1x_2 - 2x_1x_3$. **(3+12+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 $\Pi_1\, and \Pi_2$ 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 π_1 or population π_2 using the decision rule with equal priors and equal costs. **(14+6)**

22. What are canonical correlations and canonical variables? Describe the extraction of canonical correlations and their variables from dispersion matrix. Also show that there will be p canonical correlations if the dispersion matrix is of size p.

(5+8+7)

****