

Hard Core	STH 551: MULTIVARIATE ANALYSIS	No. of credits: 4
-----------	--------------------------------	-------------------

Course Outcomes:

- CO1: Use multivariate distributions and analysis techniques appropriately
- CO2: Undertake multivariate hypothesis tests, and draw appropriate conclusions.
- CO3: Analyze multivariate data, dimension reduction and the dependence structure of variables to extract the useful information from multivariate dataset
- CO4: Apply suitable tools for classification to formulate and solve real-life problems.
- CO5: Apply suitable tools for exploratory data analysis and clustering techniques to extract the useful information from real-life problems.

Unit I

Multivariate normal distribution – maximum likelihood estimators of the parameters, multiple and partial correlation coefficients, Assessing the assumptions of normality – Q-Q plot, Chi-square plot, transformations to near normality. Inference problems in multivariate normal distribution – Hotelling's T^2 & Mahalanobis D^2 – statistics, likelihood ratio tests – one sample and two sample problems, q- sample problem, test for symmetry.

(12hrs)

Unit II

Confidence regions, simultaneous confidence statements. Independence of subvectors, sphericity test. Test for covariance matrices (Statistics and their distributions- Statements only and applications), MANOVA

Cluster Analysis: distances and similarity measures, hierarchical clustering methods, K – means method.

(10 hrs)

Unit III

Principal Component Analysis (PCA) – definition and properties, graphing the principal components, sample principal components, interpretation of zero, 'small and repeated eigenvalues, component loadings and component correlations, the problem of scaling, tests of hypotheses.

Canonical Correlation Analysis – canonical variates and canonical variables, sample canonical variates, sample canonical correlations, inference problems.

(12hrs)

Unit IV

Classification and Discrimination problems – concepts of separation and classification, Bayes', minmax and Fisher's criteria, classification rules based on Expected Cost of Misclassification (ECM) and Total Probability of Misclassification (TPM), classification with two multivariate normal populations (equal and unequal covariance matrices) evaluating classification rules, classification with several populations, Fisher's linear discriminant function, tests associated with discriminant functions.

(12 hrs)

Unit V

Factor Analysis: orthogonal factor model, factor loadings, estimation of factor loadings, factor scores. Introduction to structural equation modeling (SEM) – nature, purpose, variables, scales, construct validity, mediation effect and goodness of fit measures of SEM model

(10 hrs)

References :

1. T.W.Anderson (1984): An Introduction to Multivariate Analysis, 2nd Ed., John Wiley.
2. Bernard Flury (1997): A First Course in Multivariate Statistics, Springer Texts in Statistics.
3. A.M.Kshirasagar (1972): Multivariate Analysis, Marcel Dekker.
4. K.V.Mardia, J.T. Kent and J.M. Bibby (1979): Multivariate Analysis, Academic Press.
5. C.R. Rao (1973): Linear Statistical Inference and its Applications, Wiley Eastern.
6. Richard Arnold Johnson and Dean W. Wichern (2012) Applied Multivariate Statistical Analysis, Prentice Hall, 6th edition
7. Alvin C. Rencher, William F. Christensen (2012), "Methods of Multivariate Analysis" John Wiley.
8. Rex B kline (2015) : Principles and practices of structural equation modeling , fourth edition ,Guildford press.
9. Geoffrey,M,Maruyama (1998): Basics of structural equation modelling ,Sage Publications ,Inc.
10. A. M. Mathai. Hans, J. Haubold, and Serge B.Provost (2022) Multivariate Statistical Analysis in the Real and Complex Domains. Springer.