Course title: Multivariate Data Analysis

Course code: NRE 112	**No. of credits:** 3	**L-T-P distribution:** 28-14-0	**Learning hours:** 42

Pre-requisite course code and title (if any): NRE 115 Environmental Statistics

Faculty: Prof. Prateek Sharma	**Department:** Department of Natural Resources

Course coordinator (s): Prof. Prateek Sharma	**Course instructor (s):** Prof. Prateek Sharma

Contact details:

Course type	**Compulsory**	~~**Core**~~	**Elective**	
Course offered in	**Semester 1**	**Semester 2**	~~**Semester 3**~~	**Other**

Course Description

Large amount of data is collected on many different variables across disciplines in order to understand the underlying process(es). The multivariate analysis of data deals with examining interrelationship between three or more equally important variables or explaining of variation in, usually one (or more than one) dependent variable(s) on the basis of two or more independent (explaining) variables. With the availability of inexpensive, fast and efficient computing resources and statistical packages there has been a growth in the application of these techniques. This course introduces the student to various multivariate data analysis tools. The focus is on cross-disciplinary application of these techniques.

Course objectives

1. Introduce the language of multivariate data analysis
2. Understand the characteristics of multivariate quantitative research, including strengths and weaknesses
3. Understand the principles and characteristics of the multivariate data analysis techniques

Course content

SNo	Topic	L	T	P
1.	**Introduction** Basic multivariate statistics–mean, variance, ovariance, correlation, linear combination of variables, geometric concepts, distances	4	2	
2.	**Interdependence methods** Principal component analysis, factor analysis, cluster analysis, correspondence analysis, multidimensional scaling	12	6	
3.	**Dependence methods** Multiple regression models, logistic regression canonical correlation, discriminant analysis.	12	6	
	Total	**28**	**14**	

Evaluation criteria

- 2 minor test: 15% each
- Assignment + viva voce: 20
- Major test: 50

Learning outcomes

1. distinguish between dependence and interdependence methods in multivariate data

analysis
2. identify the most appropriate statistical techniques for a multivariate dataset
3. carry out and apply commonly used multivariate data analysis techniques, and interpret results
4. use statistical software packages for the analysis of multivariate data

Pedagogical approach

Materials

Required text

1. Afifi A., May S. and Clark V.A. (2012) *Practical Multivariate Analysis*, CRC Press, Taylor &Francis, Boca Raton.
2. Johnson R.A. and Wichern D.W. (2002) *Applied Multivariate Statistical Analysis*, Prentice Hall of India Pvt Ltd., New Delhi.
3. Sharma S. (1996) *Applied Multivariate Techniques*, John Wiley and Sons, New York.

Suggested readings

1. Alt M. (1990) *Exploring Hyperspace–A Non-mathematical Explanation of Multivariate Analysis,* McGraw-Hill Book Company, New York
2. Everitt B.S. and Dunn G. (2001), *Applied Multivariate Data Analysis*, Arnold, London.
3. Haan C.T. (1977)*Statistical Methods in Hydrology*, The Iowa State University Press/Ames.
4. Harris R.J. (1985) *A Primer in Multivariate Statistics*, Academic Press, New York.
5. Manly B.F.J. (1994) *Multivariate Statistical Methods–A Primer*, Chapman and Hall, London.

Case studies

Websites

Journals

1. Applied Statistics
2. Biometrics
3. Biometrika
4. Environmental and Ecological Statistics
5. Environmetrics
6. Journal of the American Statistical Association
7. Psychometrika
8. Statistical Science
9. Technometrics
10. The American Statistician

Additional information (if any)

Student responsibilities

Attendance, feedback, discipline, guest facultyetc