ROLLINS SCHOOL OF PUBLIC HEALTH
EMORY

DEPARTMENT: **BIOS**

COURSE NUMBER: **726** **SECTION NUMBER: 1** **SEMESTER: Spring**

CREDIT HOURS: 2

COURSE TITLE: Applied Multivariate Analysis

INSTRUCTOR NAME: Mary E. Kelley, PhD

INSTRUCTOR CONTACT INFORMATION

EMAIL: mekelle@sph.emory.edu

PHONE: (404) 712-0804

SCHOOL ADDRESS OR MAILBOX LOCATION:

OFFICE HOURS: F 2:30 – 4:00

COURSE DESCRIPTION (3-4 Sentences)

The course covers multivariate analysis methods including multivariate regression and ANOVA, data reduction through the use of principal components or factor analysis, multivariate classification and clustering methods. Where applicable, nonparametric and/or categorical adaptations of traditional multivariate normal analyses will also be presented. Prerequisites include BIOS 510, 511 or equivalent

EVALUATION

70% of the grade will be based on homework assignments. Students will also conduct analysis of a research dataset based on one of the presented techniques and write up the conclusions in a scientific paper that will be 30% of the grade.

Datasets will be provided/chosen by the instructor.

LEARNING OBJECTIVES OR COMPETENCIES OF THE COURSE

- to understand the uses of multivariate techniques in the analysis of research data;

LEARNING OBJECTIVES OR COMPETENCIES FOR THE DEPARTMENT OR PROGRAM TO WHICH THE COURSE CONTRIBUTES

The purpose of this course is to give both masters and doctoral level students experience with multivariate techniques in the analysis of research data.

Syllabus: Applied Multivariate Analysis

Required textbooks:

- Johnson RA, Wichern DW. Applied Multivariate Statistical Analysis. Pearson Prentice Hall, New Jersey, 2007. ISBN: 978-0-13-187715-3 (JW)
- Khattree R, Naik DN. Applied Multivariate Statistics with SAS Software, 2nd edition. SAS Institute Inc, Cary NC, 1999. ISBN: 1-58025-357-1 (KN1)
- Khattree R, Naik DN. Multivariate Data Reduction and Discrimination with SAS Software. SAS Institute Inc, Cary NC, 2000. ISBN: 1-58025-696-1 (KN2)

Supplemental readings:

- Bartholomew DJ, Steele F, Moustaki I, Galbraith JI. The Analysis and Interpretation of Multivariate Data for Social Scientists. CRC Press, Boca Raton, 2002.
- Kaufman L, Rousseeuw. Finding Groups in Data: An Introduction to Cluster Analysis. John Wiley & Sons, Hoboken, NJ, 2005

1) Introduction to multivariate data
   a) Data structure and the multivariate problem (JW, chpt 1)
   b) Measures of distance
   c) Matrix Algebra review (length and topics depend on class composition) (JW, chpt 2)
   d) Multivariate estimation (JW, Chpt 4, KN1 chpt 1-2)
2) Multivariate regression and MANOVA (JW 6.3 – 6.10, 7; KN1, chpts 3-4)
3) Repeated measures as a multivariate problem (JW 6.1-6.2, KN1 chpt 5,6 (opt mixed))
4) Data reduction
   a) Principal components analysis (JW Chpt 8; KN2 chpt 2)
   b) Factor analysis
      i) MVN based inference (JW chpt 9; KN2 chpt 4)
      ii) Ordinal and binary FA (supp)
5) Path analysis and structural equation modeling (SEM) (supp)
6) Classification and clustering
   a) Discriminant analysis (JW chpt 11, KN2, chpt 5)
   b) Nonparametric classification (classification and regression trees, recursive partitioning)
   c) MVN-based clustering methods (JW chpt 12; KN2 chpt 6)
      i) hierarchical clustering
      ii) k-means
      iii) multidimensional scaling (MDS)
   d) Categorical clustering methods (supp)