Program Cover Sheet --- STA306: Applied Multivariate Analysis

I. Basic Course Information

Applied Multivariate Analysis is an upper level statistics option, both for statistics majors and statistics minors. The course meets twice weekly. Students must have completed STA215 prior to enrolling. The course is suitable to students from many disciplines.

II. Learning Goals

The American Statistical Association's guidelines for undergraduate programs in statistical science state that such programs should "emphasize concepts and tools for working with data and provide experience in designing data collection and in analyzing real data that goes beyond the content of a first course in statistical methods." More specifically, they recommend that programs should provide statistical topics that include random sampling, stratification in sample surveys, data exploration in observational studies, and a variety of formal inference procedures. They also recommend that programs should require familiarity with a standard statistical software package.

This course covers aspects of statistical inference in cases where there are multiple variables. In contrast to Regression Analysis, the methods in multivariate analysis focus on areas of unsupervised data exploration and classification. It provides an introduction to methods such as clustering and PCA for exploration and linear discriminant analysis for classification. The rationale underlying each test will be conveyed and statistical software packages SAS or R as decided by the individual instructor.

While topics may expand with the field, on completion students will typically have covered

   (i) Matrices, eigenvalues and eigenvectors

   (ii) Principal components analysis

   (iii) Factor analysis

   (iv) Discriminant analysis (two and three group analyses)

   (v) Cluster analysis, using a variety of metrics and clustering algorithms

   (vi) Multi-dimensional scaling (classical and ordinal)

   (vii) Correspondence analysis

   (viii) Canonical correlation analysis

III. Student Assessment

Students will be assessed through a combination of weekly problem sets, exams during the semester, and a final project or final exam either alone or in combination.

IV. Learning Activities

Learning activities will consist of a combination of lectures, discussions, student presentations, and programming and analysis assignments. The specific choice will

depend upon the individual instructor. Outside of class, students are expected to do a significant amount of individual and group work to achieve the learning goals, including multiple lengthy problem sets that analyze real world data and involve significant use of computer analysis. Students will receive detailed feedback in order to improve their analysis skills and application of analysis methods. Students are encouraged to work together to more fully develop a deep understanding of multivariate statistics. The ability to do many tasks on laptop computers, including working remotely on the ELSA HPC cluster, will permit students to work as a group in the library in addition to in the School of Science computer laboratories.

Approved 2-9-22