LINEAR MODELS

STAT 4430/8436

Course Description:

This is an introduction to linear statistical models which will include: simple linear regression models, multiple linear regression models, ANOVA models including one way ANOVA, randomized block design, and other designs. Also, logistic regression models, Poisson regression models, bootstrapping/resampling models, survival analysis. Some necessary linear algebra and mathematical statistics ideas will be covered in the course also. If time allows, some mixed models and/or survival models. Much use of computer software will be made. **3 credits**

Prerequisites:

Undergraduate and graduate: Math 4750/8756 with a C- or better or Stat 3800/8805 with a C- or better or permission of instructor. Permission based upon students' having taken a basic statistics course with a C- or better and having at least a rudimentary knowledge of calculus.

Overview of content and purpose of the course:

The purpose of the course is to provide the student with the background necessary to apply linear statistical models to make predictions and inferences about the relationships between a set of variables. Unlike other linear models courses offered in non-mathematics departments, this course will discuss some of the theoretical aspects of statistical inference from linear models and will be approximately 20% theory and 80% applications (data analysis). The theoretical part of the course will require some mathematics from linear algebra and mathematical statistics concepts such as likelihood functions and sufficient statistics and will allow the student to understand where the formulas are coming from and how they work. However, the greatest emphasis is on data analysis and applications.

Heavy use of the computer packages will be made for the applications part of the course**.** This course is required for students interested in data science.

Anticipated audience/demand:

There has been increasing interest in statistics among our math majors over the past five or ten years. We would like to expand our offerings in statistics as well as give students interested in data science a firm background in linear models, a very important part of data science. The audience would be math/computer science/economics/sciences/business majors. More and more employers are looking for students who can analyze mountains of data, so the demand is high.

Major topics:

1) Brief Review of Mathematical Statistics Concepts

2) Some Linear Algebra Topics

3) Likelihood Functions and Sufficiency

4) Simple Linear Regression

5) Multiple Linear Regression

6) Model Selection

7) Correlation and Partial Correlation

8) One Way ANOVA

9) Randomized Block Design

10) Multiple Comparison Procedures

11) Other ANOVA Models

12) Logistic Regression

13) Resampling/Bootstrapping

14) Poisson Regression

15) Other Linear Models Including Survival Analysis Models, Mixed Effects Models

Methods:

The primary method of instruction is the lecture format with ample opportunity for student questions and participation.

Student role:

Class attendance is expected. Some material discussed will not be in the textbook. Students will be expected to do homework assignments which require both 'pencil and paper' solutions as well as use of computer software.

Textbook:

Kutner, M.H., Nachstein, C.J., and Neter, J., (2004). Applied Linear Statistical Models, Fourth Edition, McGraw Hill.

January 2016