ISyE 7401  Advanced Statistical Modeling  Spring 2015

Instructor:  Nicoleta Serban (nserban@isye.gatech.edu) 438 Groseclose Building
Office Hours:  Tuesday and Thursday 2-3pm. Please come prepared.
Class Meets:  TTH 12:05 - 1:25pm Groseclose 402.

Web Address:  T-square
Most of our class material will be submitted via t-square including
• Course syllabus
• Lecture and computer notes
• Homework assignments and solutions
• Your course grades
• Questions and replies for homework assignments
• Important announcements

Textbooks:  Optional:

Objective:  This course will cover the fundamentals of advanced statistical modeling. Topics include: point estimation, likelihood principle, interval estimation, hypothesis testing, linear regression, generalized linear regression, nonparametric inference and bootstrap. This course is excellent preparation for advanced work in statistics and machine learning.

Class Notes:  Class notes will be posted on the web regularly. The notes are not meant to be a substitute for the optional material and hence are generally quite terse. Read both the notes and additional material on each topic.

Honor Code:  For any questions involving Academic Honor Code issues, please consult me, the class teaching assistant, or www.honor.gatech.edu. In this course, you are allowed to work together in the homework, as long as you write up and turn in your own solutions including the computer codes.
Midterm: There will be one midterm exam with problems reviewing the material (lectures and assignments) provided in this course throughout the first part of the semester. The midterm exam is close notes (including homework) and books but a two (one-sided) pages with formulas will be allowed. The intent of this midterm is to put the basis of theoretical aspects of statistical inference used in the rest of this class. All students will take the midterm in the same day at the same time. **No make-ups.**

**Midterm Date:** February 5th in class

Homework: Assignments will include both theoretical and computer problems; the latter problems will ask you carry out analysis of data sets using computer software. Keep in mind that you should not hand in raw computer output. Conclusions and interpretation of results are more important than good printouts. These assignments are intended to help you prepare for the midterm exam and final project. You are allowed (and encouraged) to work together with other students on homework, as long as you write up and turn in your own solutions. You are also allowed (and encouraged) to ask me questions, although you should try to think about the problems before asking. **Late Homework will not be accepted.**

Project: This project is a requirement you must fulfill in order to pass this course. The general goal of the project is to provide you with experience in applying advanced statistical methods to real data. For this project, you must find a data set on your own and write a final report on the analysis of the selected data. The data cannot be a data set found in a textbook or been analyzed in detail and results published. In grading, I will primarily look for a sensible approach to the problem, and clearly-made connections between your analyses and the substantive questions. You can use any computing equipment and any computing resources in the school, any written source material you can find, in or out of the school. However, replicating results which have been already published without referencing to the source of publication is subject to plagiarism. Plagiarizing is defined by Websters as “to steal and pass off (the ideas or words of another) as one’s own : use (another’s production) without crediting the source.” Be sure to document carefully your project work.

The project reports will be graded by the instructor and students in this class. Each student will be assigned to grade projects of other students. The report grade will be a weighted average between students’ grade and the instructor grade. Please refer to grading guidelines document.

**Deadline to submit an abstract of the project:** March 10th, 2015.

**Deadline to submit the project work (report):** April 27th, 2015.
Grading: The course grade is based on Class Attendance (5%), Homework (25%), Midterm (30%), and Final Exam (40%).

Course Topics

- Convergence
- Estimation Methods and Likelihood Principle
- Hypothesis Testing & Confidence intervals
- Multiple Linear Regression
- Generalized Linear Models
- Model Selection
- Nonparametric Inference and Bootstrap
- Nonparametric Regression