Statistics 471: Intermediate Statistics

Instructor   Nancy R. Zhang
                        nzh@wharton.upenn.edu
                        Office hours are Wednesdays, 1:45-2:45 PM at JMHH 467.

Graduate Student

Helper: Peichao Peng (ppeichao@wharton.upenn.edu)

Canvas   Most of the materials including announcements, reading material, homeworks,
                        solutions, etc. will be available on our Canvas site.

Course overview

This is a second course in Statistics aimed to equip students with tools to analyze diverse real-world data and to justify their use through theory.   The first half of the course covers advanced regression-based concepts such as weighted least squares, ecological regression, response surface analysis, logistic and Poisson regression, and model building. The second half of the course introduces students to a mixed set of tools such as random effects ANOVA, principal components analysis, survival analysis, Bayesian statistics, bootstrap, and cross-validation. The emphasis will be on modeling and data analysis for real-world problems.

Prerequisites

The prerequisite of the course is fluency with basic probabilistic reasoning (e.g., probability distributions and densities; joint distributions; conditional probability, independence, correlation), familiarity with statistical estimation, hypothesis testing, and basic regression modeling (at the level of STAT 102 or 431). It would be helpful to have previous exposure to linear algebra, but it is not required.

Textbook

There is no textbook for this course. Reading material and lecture notes will be posted on to the course website.

Statistical computing software

The statistical computing software R will be used in the course. It is free, and can be downloaded at the R-project website:

```
http://www.r-project.org/
```

The above website also contains a list of manuals for using the software. Basic usage of R will be illustrated in class and through sample codes posted on the course website. Again, no previous exposure to the software is required.

Grading policy

- Assignments (10%): There will be 5-6 exercise sets designed to reinforce lecture material, assigned approximately bi-weekly.
- Project (30%): There will be a project for the course. You can form teams of up to 3 people. You will be graded on a final report and in-class presentations.
- Midterm (30%): Exam will be in class on ***Wednesday, Feb 27***.
- Take-home Final (30%): The Take-home Final will be due on ***Wednesday, April 30 at 5PM.***