

Statistics 431: Statistical Inference

Syllabus, Spring 2015

Classes: Section 001: Tu/Th 10:30 – 11:50 am in G50 JMHH
Section 002: Tu/Th 12:00 – 1:20 pm in G50 JMHH

Instructor: Nathan Stein
Email: natstein@wharton.upenn.edu
Office: 455 JMHH
Office hours: Wed, 10:30 am – noon

Teaching assistant: Linjun Zhang
Email: linjunz@wharton.upenn.edu
Office: 440 JMHH
Office hours: Thurs, 9:30-10:30 am

Course overview

This course offers an advanced undergraduate-level exploration of statistical techniques for data analysis. We will examine a collection of basic concepts and commonly used methods, with an emphasis on understanding when, how, and why to use them. We will also experiment with these ideas on example data sets using the statistical software R.

Topics include (1) collection, summary, and display of data, (2) estimation, hypothesis testing, and confidence statements, and (3) simple and multiple linear regression. If time permits, we will also discuss likelihood-based inference.

Prerequisites

The official prerequisite is Statistics 430. The effective prerequisite is fluency with basic quantitative probabilistic reasoning and analysis (e.g., probability distributions and densities; jointly distributed random variables; conditional probability; independence, correlation, and covariance; normal and binomial distributions; law of large numbers; central limit theorem).

Previous exposure to the statistical computing language R is *not required*, though prior programming experience will be helpful.

Textbook

Statistics and Data Analysis from Elementary to Intermediate, by A. C. Tamhane and D. D. Dunlop, Prentice Hall, 2000.

Homework assignments

- There will be six homework assignments.
- *No late homework will be accepted.* However, the lowest score will not be counted toward your final grade.

Exams

There will be two midterm exams, a final project, and *no final exam*.

- Midterm 1: Thursday, February 12
- Midterm 2: Thursday, March 26
- Both exams will be closed book, but you will be allowed to bring a certain number of pages of notes.

Final project

A written final project will be due at the end of the semester. The goal of the project is to apply the tools you've learned during the semester in the analysis of a real data set. You can work individually or in pairs.

Collaboration policy

Students may help each other in solving the homework problems. However, you are expected to prepare the final write-up individually with acknowledgment of the help received. No collaboration is permitted on the exams.

Grading policy

- Homework assignments: 15% (with the lowest score dropped)
- Midterm 1: 25%
- Midterm 2: 30%
- Final project: 30%

Course website

The course website will use the Canvas platform. Please check the course website for announcements, handouts, sample code, assignments, and other materials, etc.

Statistical computing software

The statistical computing software R will be used in the course. It is free and can be downloaded at

<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 code posted on the course website, and no previous exposure to the software is required.