

Statistics 431: Statistical Inference

Syllabus, Spring 2014

Classes: Section 001: Tu/Th 10:30 – 11:50 am in F55 JMHH
Section 002: Tu/Th 12:00 – 1:20 pm in F55 JMHH

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

Teaching assistant: Anru Zhang
Email: anrzhang@wharton.upenn.edu
Office: 449 JMHH
Office hours: Thurs, 1:30 – 3:00 pm

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 and how to use them, and why. Students will also experiment with these ideas on data examples 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.

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.

Students from other schools should first apply for a class account at <http://accounts.wharton.upenn.edu>

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*.

Homework assignments

- There will be six homework assignments.
- Homework assignments will be posted on the course website, and after the due dates, solutions will be posted.
- *No late homework will be accepted*. However, the lowest score will not be counted toward your final grade.
- Students may help each other in solving the problems, but are expected to prepare the final write-up individually with acknowledgment of the help received.

Exams

- Midterm exam: Wednesday, March 19, 6:30-8:30 pm
- Final exam: Monday, May 5, 6-8 pm
- Both exams will be closed book, but you will be allowed to bring a certain number of pages of notes.

Grading policy

- Homework assignments: 20% (with the lowest score dropped)
- Midterm exam: 35%
- Final exam: 45%