

STAT 431-920: Statistical Inference

Summer 2014, Syllabus

Instructor: Anru Zhang

E-mail: anrzhang@wharton.upenn.edu

Office: 449 JMHH

Office Hour: M/T/W/R/F after class, 1:00pm - 2:00pm

Meeting Time: M/T/W/R/F, 11:00am - 12:50pm, 07/03/2014 - 08/08/2014
No class on Independence Day, 07/04/2014.

Location: SHDH 213

Course Description: 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) review of necessary probability theory, (2) collection, summary, and display of data, (3) estimation, hypothesis testing, and confidence statements, and (4) simple and multiple linear regression. If time permits, we will also discuss likelihood-based inference.

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

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.

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.

Homework assignments:

- There will be five homework assignments. Each homework is graded out of 100.
- Homework assignments will be posted on the course website.
- Late homework will be penalized 15 points per day with maximum three days. After three days, the homework will not be accepted because I will then post the solution on the course website.
- 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.

Extra Help: Do not hesitate to come to my office during office hours or by appointment to discuss a homework problem or any aspect of the course. If you want to hire a tutor (for a fee), you can find the other PhD students from the UPenn statistics department.

Exams:

- Three in-class quizzes: Fridays, 07/11, 07/25, 08/01.
- Midterm exam: Thursday, 07/17.
- Final exam: Friday, 08/08. The final will be cumulative but focuses on the second half of the course.
- Both exams will be closed book, but you will be allowed to bring a certain number of pages of notes.

Grade Policy:

- Attendance and course participation: 10%
- Homework: 25%
- Quizzes: 15 %
- Mid-term: 20%
- Final: 30%

Important Dates:

Classes Begin	July 3
Quiz 1	July 11
Last Day to Add a Course/Drop a Course with No Financial Obligation	July 11
Mid-term Exam	July 17
Last Day to Drop a course with 50% Financial Obligation	July 18
Last Day to Change Grade Status	July 18
Quiz 2	July 25
Quiz 3	August 1
Last Day to Withdraw from a course	August 1
Final Exam	August 8