

Statistics 112: INTRODUCTORY STATISTICS

January 16, 2019

Course Information

Lecturer: Dr. Xinran Li
Email: lixinran@wharton.upenn.edu

TA: Ruijia Wu
Email: ruijiawu@wharton.upenn.edu

Lectures: TTh 12-1:30pm F60 JMHH

Description: The world and any observation we make of it are full of randomness and uncertainty. These randomness and uncertainty, like a dense fog that obscures traffic signs, hide important patterns and discoveries. Statistics is the art and science of seeing through the fog to uncover the knowledge that lies beneath.

This is a second course on statistics. You are expected to have a basic understanding of *mean, variance, distribution, probability, confidence interval, hypothesis test, and regression* (See me if you don't know what these are).

In the first part of the course, we will study classic hypothesis tests to address questions such as *is there a difference in the effect of different types of drugs* or *is one online advertisement better than another*. We will, critically, learn to quantify uncertainty in our answers. In the second part, we will dig deeper into the problems of regression and classification so that we can describe relationships and make predictions. In the final part, we will learn how to design experiments and how to infer causality.

Office Hours

TBA

Textbook

There is no required textbook. One good reference book is:

- *EZ Statistics*, Downing and Clark (Barron's). A high-level and concise summary of the key ideas in statistics.

Grading

There will be one midterm and a final exam. There will also be biweekly homework assignments. The grades will be:

30% midterm
40% final exam
30% homework

The midterm will be on **March 14st**.

Homework

Homework must be submitted online via Canvas. Scan your written assignment. Submit the scanned file to Canvas. Make sure that all the text and handwriting are clearly legible.

10 points out of 100 points will be deducted from a late homework submission for every day that it is late. Late homework will not be accepted after 4 days. The lowest homework score will be dropped.

Syllabus

Part 0 Probability review

Part 0 Point estimation and confidence interval

Part 1 One and two-sample t -test

Part 1 Analysis of variance (ANOVA)

Part 1 Chi-squared test

Part 2 Simple linear regression

Part 2 Multiple linear regression

Part 2 Logistic regression

Part 3 Causal inference in randomized experiments

(Course Schedule is subject to minor adjustment.)