



Course Information

Stat 100: Statistics (3 credits)

Course Description

An introductory statistics course with particular emphasis on understanding statistical concepts and which tools are appropriate for which problems. Includes experimental design (including basics of casual inference from observations), basic probability, descriptive statistics, linear regression, sampling and statistical inference, and hypothesis testing. Includes easy-to-use data analysis program (<http://dataprogram.stat.illinois.edu/>)

For more details, see <http://stat100website.web.illinois.edu/>

Prerequisite:

High School Algebra

Course Objectives

Large stores of information have become readily available on crucial issues, e.g. Covid-19. You can choose either to ignore the information or to make sense of it, which means learning statistics. Statistics is a collection of real tools. The key is to understand which one to use when and why, not just to memorize some complicated rules. We teach

1. How to use a conceptual, intuitive approach to understand a set of complex statistical methods. We build a unifying framework for general predictive models.
2. How to determine whether predictors are also causes. Would changing X change Y? We examine real data sets to sort out such causal effects from confounders.
3. How to use statistical software both to help us understand what the statistical methods are doing and to do the calculations for us.

Course Outline

Study Design - observational studies vs. randomized experiments, why randomized controls are key, identifying possible confounders in observational studies.

Descriptive Statistics - mean, median, SD, histograms, box plots, normal curve, etc.

Simple Linear Regression – correlation coefficient, regression equation, etc.

Probability - multiplication rule, addition rule, conditional probability, Bayes rule.

Diagnostic Tests - Type I and II errors, sensitivity and specificity.

Statistics for Random Variables - expected value and standard error of chance processes, probability histograms and convergence to normal curve. Focus is on developing simple chance models (box models- drawing numbers at random from a box) that more complicated sampling processes can be translated into.



Sampling and Statistical Inference - using sample means and percents to estimate population means and proportions, and attaching margins of errors to our estimates by computing confidence intervals. Why randomized sampling is key.

Significance Tests - one sample and two sample Z-tests and t-tests and chi-square tests for goodness of fit and independence. Focus is on understanding how these tests depend on chance models.

Format

- This course is entirely online. All homework, surveys and bonus work are submitted and graded on Lon-Capa.
- Text: A free text/workbook you'll fill in while watching the lecture videos will be provided.
- This course has fixed due dates and deadlines for homework and exams. You'll turn in homework twice. All homework is submitted on Lon-Capa and you'll get immediate feedback on each problem as soon as you submit it. You also get multiple tries on all the problems. No late HW will be accepted but your lowest 3 assignments will be dropped.
- This course has 2 midterms and a Final that you'll take on Lon-Capa. All exams are open book.