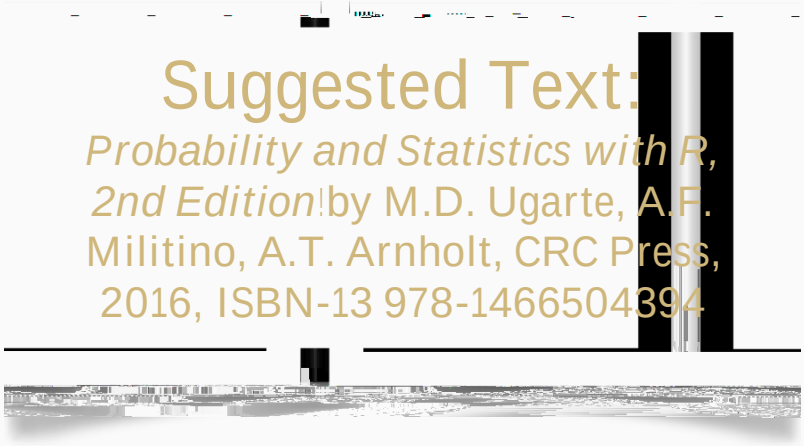
Course Description

Probability and statistics are essential tools for engineers, data scientists, and many other professionals. *Statistics* can be understood as the study and application of (1) methods for reliably gathering and presenting information (*descriptive* statistics), and (2) methods for drawing conclusions about the world from limited information (*inferential* statistics). In this course, we will be concerned with both (1) and (2). Especially for (2), we will need to learn how to analyze chance events. As such, before studying statistical inference, we will study some basic probability theory. Our ultimate goal is to use tools from mathematics to analyze sample data and try to make predictions or draw conclusions about law-like relationships that hold in more general populations. My aim in this course is to help students become proficient in the

4. estimate population parameters of interest by calculating point and interval estimates from a sample/data;
5. perform statistical hypothesis tests;
6. compute Bayesian posterior distributions and conduct Bayesian data analyses in R.
7. compute bootstrap confidence intervals in R.
8. construct basic data visualizations in R and organize analysis, findings, and recommendations into reports clean, easy to interpret reports.



Course Webpage

Course materials, such as this syllabus, the course schedule, homework assignments, and general updates will be uploaded to our Canvas page. Please check our Canvas page frequently!

Assignments

A Note about Jupyter and R

On many assignments in this course, we will use the R programming language; we will run R within the web application Jupyter. Both Jupyter and R are great (free!) tools for data analysis/science, and I think they will be beneficial to you beyond this course. We will spend a few class periods downloading these programs and becoming comfortable using them. Learning new languages and applications can be difficult and frustrating, but also rewarding. I'm here to help you on this journey! Taking the first in-class Jupyter assignment seriously will make all future assignments much easier.

Homework (30%)

Homework will be due (roughly) once every week (with a few weeks off here and there). Due dates and times will appear on Canvas assignments. Late homework will not be accepted or graded, except in extraordinary circumstances. Homework assignments will have a theoretical section and a computational section. You are asked to electronically submit a single Jupyter file to Canvas; this file will contain your answers to both sections. Ideally, your answers to theoretical section will be typeset in Markdown cells (below each question) using LaTeX. You may also embed a properly scanned pdf of handwritten answers to theoretical questions into the Jupyter file. The computational portion of the homework should include all relevant R code and output, and a write-up and interpretation of your results. You can collaborate with your classmates on the homework assignments, but you must write up the results independently of each other.

Exams (15% each)

There will be two midterm exams and one final exam. The date of these exams are on our course schedule (on Canvas).

Final Project (15%)