

APMA 1650: Statistical Inference 1

Mo, Tu, We, Th 1-3 pm

Kassar House 105

Instructor: Ross Parker (ross_parker@brown.edu)

TA: James Zhang (james_zhang@brown.edu)

Office Hours and location: TBA

Course Description

In 2008, baseball statistician Nate Silver launched a political blog entitled *FiveThirtyEight.com*. After making predictions for the 2008 Democratic primary elections which were significantly more accurate than those done by professional pollsters, Silver's blog went viral. For the presidential election, Silver correctly predicted the winner of 49 out of 50 states. In addition, he correctly predicted close races in Missouri and North Carolina and forecasted the overall popular vote to within one percentage point. He also correctly predicted the outcome of every senate race.

APMA 1650 is an integrated first course in mathematical statistics oriented towards applications of statistics to the natural and social sciences. It provides a solid mathematical foundation for probability and statistics which is necessary to develop and understand complex statistical models such as the one used by Nate Silver.

The first half of APMA 1650 covers probability, the analysis of random events. We start by defining a probability space and then discussing the axioms necessary to construct a theory of probability which is consistent with our intuition. We will then discuss common probability distributions and how they can be used to model real-world phenomena.

After this, we transition into the realm of statistics. While probability theory allows us to make predictions based on a model, statistics is the art and science of determining which model best fits real-world data. We start with parameter estimation, in which we attempt to predict properties of a population using data from only a small sample. We conclude with hypothesis testing, the use of statistics to determine if there is a significant difference between two data sets.

APMA 1650 is the starting point of paths that lead to other courses at the advanced undergraduate and beginning graduate level. These topics include information theory, operations research, computational statistics, financial models, and computational biology.

Since 2016 is a presidential election year, I will use the election as a unifying theme for the course. I will try to incorporate as much current data as I can to my examples and will try to relate the concepts we learn to polling and electoral predictions. My hope is that this will make probability and statistics both more relevant and more fun. While I cannot promise that you will attain the statistical savvy of Nate Silver after six short weeks, I can promise that you will complete the course with a solid foundation to continue your study of statistics.

Nuts and Bolts

Class format: This is a lecture-based course. We will have two-hour lectures four times per week. There will be a five-minute break at the halfway point. There will also be an optional one-hour recitation session (to be scheduled at a mutually convenient time) during which students will work on problems in groups. Problems in recitation section make great practice for the exams! There will be two midterm exams (July 13 and July 28) and one final exam (August 11). I will hold review sessions prior to each exam.

Homework: There will be eight problem sets total, which is approximately two per week. Problem sets will be due on Mondays and Thursdays (with the exception of the week of July 4; due to the federal holiday, problem sets will be due on Tuesday and Thursday that week). On weeks when there is an exam, there will be no problem set due on Thursday. There are no problem sets due the first week of class. Problem sets will be posted at least one week in advance of the due date. Homeworks may be submitted in class on the due date, or they may be dropped in the APMA 1650 homework box in 182 George St. by 3:45 pm on the due date. (182 George St. closes at 4 pm during the summer, hence the 3:45 deadline for submission). You are encouraged to work together on assignments, but you must write up your own solutions.

Homework policy: Late assignments generally will not be accepted. That being said, I understand that the unexpected does happen. If you have a serious situation in which you will believe you will be unable to complete your assignment on time, contact me directly and we will arrange something.

Textbooks: Lectures will be based on my own notes, which will be posted on this website after each class. Due to the prohibitive cost of textbooks (approximately \$290 for the traditionally required textbook), I cannot with good conscience ask you to purchase a textbook for this class. That being said, a textbook is a valuable reference, both to clarify points made in class and as a source of practice problems. I will place multiple copies of the following textbook on reserve in the Rockefeller Library.

- Wackerly, Mendenhall, and Scheaffer, *Mathematical Statistics with Applications*, 7th Edition. Thomson Brooks/Cole, 2008.

This has been the required textbook in the past for APMA 1650. As it sells for \$275 new on amazon.com, I cannot in good conscience require this textbook. For those who would like a physical reference book, the 6th edition (2001) is not significantly different and can usually be found used on amazon.com for less than \$20.

Prerequisites: One year of university-level calculus. At Brown, this corresponds to MATH 0100, MATH 0170, or MATH 0180. A score of 4 or 5 on the AP Calculus BC exam is also sufficient. Multivariable calculus (MATH 0190, MATH 0200, or MATH 0350 at Brown) will be helpful for one small part of the course, but is not required. I will teach any multivariable techniques we use in detail.

Communication: Email is the best way to reach me. During the week, I will try to respond within 24 hours. Email responses may be slower on the weekends, but I will try to reply by Sunday evening. For complex questions, I may ask you to talk with me after class or come to my office hours.

Grade Distribution:

Problem Sets	30%
Midterm Exam	30%
Final Exam	40%

Academic Honesty Policy

Norms regarding the quality and originality of academic work are often much more stringent and demanding in college than they are in high school. All Brown students are responsible for understanding and following Brown's academic code, which is described below.

Academic achievement is ordinarily evaluated on the basis of work that a student produces independently. Students who submit academic work that uses others' ideas, words, research, or images without proper attribution and documentation are in violation of the academic code. Infringement of the academic code entails penalties ranging from reprimand to suspension, dismissal, or expulsion from the University.

Brown students are expected to tell the truth. Misrepresentations of facts, significant omissions, or falsifications in any connection with the academic process (including change of course permits, the academic transcript, or applications for graduate training or employment) violate the code, and students are penalized accordingly. This policy also applies to Brown alumni, insofar as it relates to Brown transcripts and other records of work at Brown.

Misunderstanding the code is not an excuse for dishonest work. Students who are unsure about any point of Brown's academic code should consult their courses' instructors or an academic dean, who will be happy to explain the policy.

Academic Support

Brown University is committed to full inclusion of all students. Please inform me if you have a disability or other condition that might require accommodations or modification of any of these course procedures. You may speak with me after class or during office hours. For more information contact Student and Employee Accessibility Services at 401-863-9588 or SEAS@brown.edu.