Do It Yourself Statistics

394U

Syllabus

Instructor: Dr. Larry Cormack

Office: 4.228 SEA

Email: cormack@mail.utexas.edu (this is a great way to get in touch with me)

Phone: 471-5380

Office hours: W 1-2 and by appt.

Meeting time: T.,TH. 11 am - 12:15 pm

Meeting place: SEA 2.116

Unique number: 44020 (psy), 56775 (neu), 58550 (ssc)

The power of modern microcomputers has revolutionized data analysis in many fields and even mathematics itself. Unfortunately, the enormous benefits have been slow to penetrate Neuroscience and Psychology, largely because these fields have clung to traditional hypothesis-testing statistics. Computers have been used only for convenience, to run a standard set of tests more rapidly and easily – tests that were developed when numerical calculation was quite laborious. One unfortunate result of this is that data analysis in the aforementioned fields has not progressed much since the 1920’s (in contrast to other scientific disciplines). Another is that statistical tests are often done capriciously, with little understanding of the underlying theory. This course will provide an overview of doing data analysis and statistics from scratch, that is, from a computer programming environment rather than using commercial statistical software packages. This experience will be valuable for several reasons. First, we will gain a deeper understanding of traditional statistical tests. Perhaps more importantly, we will learn modern computer-based methods of analysis such as Monte Carlo simulation and Bootstrap analysis. Third, we will gain experience in writing our own computer programs for data analysis, thus allowing us to confront situations not anticipated by commercial software packages. Finally, the experience in computer programming will be of use outside the realm of data analysis because, of course, computers are increasingly used in all aspects of experimental science (e.g. automated data collection, experimental control, etcetera).

Projects (in approx. temporal order)

1. Introduction to programming
   a. The MATLAB environment
   b. a simple program
   c. variables and operators
   d. functions
   e. control flow
2. Sampling, Normality, & CLT
3. Trad. Methods, M.C., and the bootstrap
4. Single sample hypothesis testing
5. Two sample hypothesis testing
6. Multi-group hypothesis testing
7. Counting data and chi-squared
8. Contingency tables
9. General curve fitting
10. Final Project *

* The existence and / or complexity of the final projects will depend upon how many students we have.