BBSP 610 - Introductory Statistics for Laboratory Scientists
Fall 2016

Couse Time/Location: Koury 3615, 11:00 am-12:15 pm TTh

Instructor: Eric Bair
Instructor E-mail: ebair@email.unc.edu
Office: Koury 4503
Office Phone: 919-537-3276
Office Hours: Mon./Tues. 4pm-5pm (or by appointment)

Teaching Assistant: Nuvan Rathnayaka (nuvanrath@gmail.com)

Textbook (optional): Samuels and Witmer, *Statistics for the Life Sciences*
Textbook (optional): Gonick and Smith, *The Cartoon Guide to Statistics*

Course Description

BBSP 610 introduces the basic concepts and methods of statistics with emphasis on applications in the experimental biological sciences. Emphasis is on mastery of basic statistical skills and familiarity with situations in which advanced analytical skills may be needed. The primary focus of the course is on applications in basic science research, and students with primary interests in prospective epidemiological studies strongly consider BIOS 600 instead. Course objectives include learning to use statistical reasoning to formulate scientific questions in quantitative terms, learning to design and interpret graphical and tabular displays of statistical information, using basic probability models to describe trends and random variation in laboratory data, and using basic statistical models, including tests and confidence intervals, to draw inferences from data. Topics include experimental design, basic summary statistics, graphical methods for visualizing data, probability, confidence intervals, hypothesis testing, and regression. The course introduces and employs the freely available statistical software, R, to explore and analyze data, although support will also be provided for students who wish to use Prism or Excel. Emphasis will be placed on mastery of basic statistical analysis skills, familiarity with situations in which advanced analytic skills may be needed, the ability to critically review statistical analysis presented in relevant manuscripts, and the ability to clearly describe the results of statistical calculations when writing papers and grant applications.

Course Objectives

The objectives of this course are to provide graduate students in biomedical research programs familiarity with basic experimental design and elementary statistical methods. By the end of the course, students should understand the principles of experimental design, be familiar with basic statistical methods (and how they are implemented in computer software), and know which methods are appropriate in a given circumstance.

Prerequisites

Students should have a basic understanding of algebra and arithmetic. If you are enrolled in a graduate program in a science-related field, you should be fine. Meet with the instructor if you are concerned about your preparation for this course. No previous background in probability or statistics is required, nor is experience with statistical computing.

Course Format

Prior to each class, a set of lecture notes and recorded lecture videos will be posted online. Students should be familiar with the material presented in these lectures (either by reading the slides or watching the videos) before each class. Each Thursday class period will involve discussions about the assigned reading as well as selected homework problems. Each Tuesday class session will begin with a short quiz based on a homework problem. The TA's will then provide a brief review of the lecture material and examples with the remaining class time used to answer students' questions. Class attendance is mandatory unless authorized by the instructor.

Course Assignments
There will be short written assignments due approximately every two weeks. Students will also be assigned papers to read and discuss during class. You may turn in your homework by either e-mailing it to me or uploading it to Sakai. **PLEASE INCLUDE YOUR NAME IN THE FILE NAME.**

Please try to turn in your homework by the assigned due date. I understand that you have other classes and lab work that are competing for your time, so I try not to be an ogre about due dates, but I also want to get the homework graded and returned as quickly as possible. You may turn your homework in up to one week late without penalty. However, try to avoid this unless it is really necessary. Homework that is turned in more than a week late will be penalized unless you get permission from the instructor to turn it in later.

Final Project
The final project will involve using the statistical methods discussed in the class to analyze a real-world data set. This data set may be data collected during your lab/rotation project, data you download from the Internet, or any other suitable data set. The final project should include a written description of your analysis that includes: 1) a description of the data and how it was collected, 2) the hypotheses that you wish to test using the data, 3) the steps to taken to recognize/remove outliers or other problematic data, 4) a description of the statistical methods used to analyze the data and why these methods are appropriate, and 5) the conclusions of your analysis. The writing style should be suitable for a scientific journal or a grant application and should be 1-2 pages in length. In addition to the written description of the analysis, you will give a brief presentation (5-10 minutes) describing your project during the last week of class. The final project may be done individually or in groups of 2-3 people.

Assessments
Grading will be based on performance on the written assignments, final project, and participation in the class discussions. As long as students make a reasonable effort to complete all of the assignments in the course, they should expect a grade of no lower than "P." Outstanding work will earn an "H" grade.

Tentative Course Schedule

Weeks 1-3: Experimental Design, Basic Summary Statistics, Graphical Methods, and Introduction to R (Aug. 30-Sep. 15)

- Statistical Reasoning in the Basic Biological Sciences
- Experimental Design - Controlled Experiments
- Experimental Design - Observational Studies
- Blocking and Sampling Methods
- Introduction to R/Sampling in R

- Basic Summary Statistics (Mean, Median, etc.)
- Graphical Methods for Summarizing Data
- Outliers

Weeks 4-6: Probability, Confidence Intervals, Hypothesis Testing (Sept. 20-Oct. 6)

- Probability
- Random Variables and the Normal Distribution
- Sampling Distributions and the Central Limit Theorem
- Confidence Intervals for the Mean
- Hypothesis Testing
- Two-Sample T-Tests
- Confidence Intervals and Hypothesis Tests for Proportions
- Hypothesis Tests for Categorical Data

Weeks 7-8: Introduction to Regression (Oct. 11-Oct. 20)

- Correlation
- Introduction to Regression
- Inference for Regression
- Multiple Regression
- Introduction to ANOVA
- Logistic Regression

Weeks 9: Review and Case Study (Oct. 25-Oct. 27)

- Review
- Case Study

Weeks 10: Final Presentations (Nov. 1-Nov. 3)

Mandatory Honor Code Blurb

Cheating is bad. If you cheat, you have to go to honor court and lots of bad things can happen to you. Thus, cheating would be a very silly thing to do because 1) I already told you that you won't get a grade lower than “P” if you do all the work, and 2) you are in graduate school, so nobody cares what your grades are anyway. So please don't cheat. :)

It is perfectly acceptable to work together on your homework assignments. (In fact, I encourage it.) However, each person must turn in her or his assignment separately.