Advanced Biological Statistics Bi 610 Fall 2017

Course Description:

This course aims to provide students with an understanding of the core concepts and approaches for the analysis of biological data, particularly large data sets. This is meant as a first, foundational course for graduate students. It is advanced in that we will move through the material quickly with the goal of covering all major topics in univariate data analysis, forming a foundation for subsequent learning. Students will learn to use the powerful statistical programming language R.

Instructors of Record Information:

Dr. Bill Cresko 312 Pacific Hall wcresko@uoregon.edu Office hours TBA

Course Information:

Schedule: Tue/Thur 10:00 AM - 11:50 AM, 40 Science Library
Website: Bi 610 Canvas site (https://canvas.uoregon.edu/)

Textbooks:

Wickham, H. & G. Grolemund. 2016. R for Data Science.
Logan, M. 2010. Biostatistical Design and Analysis Using R. Wiley-Blackwell.
Quinn, G. & M. Keough. 2002. Experimental Design and Data Analysis for Biologists. Cambridge Univ. Press.
* Wickham, H. 2009. GGPlot2: Elegant Graphics for Data Analysis. Springer. (*Recommended*)

Software:

Latest version of **R** from the R Project (www.r-project.org)
An R scripting environment such as **RStudio** (www.r-project.org) & **TextWrangler** (free version of **BBEdit**)

Prerequisites:

None, but students should be comfortable with algebra and the basics of calculus.

Grading:

Homework assignments	35%
Independent analyses	35%
Final Project	30%

The final course letter grade will be determined based on a curve of total performance in the course

Homework assignments:

Two problem sets will be assigned to be turned in the following week (see course schedule). The intent of each homework is to expose students to running specific analyses in R. You will complete the analyses and submit both your R code and resultant output via the Canvas course website. More details regarding each home work will be given at the appropriate time.

Independent analyses:

There will be two small projects during the course of the term. The projects will consist of analyzing data sets provided by the instructor. For each project you will produce a write-up of the analyses similar to the results section from a journal article. More details regarding the projects will be given at the appropriate time.

Final Project:

Each student will be required to use the concepts and approaches throughout the term to analyze a rich data set provided by the instructor. The final project will be due during finals week, and should comprise a write-up similar to the independent analyses, as well as a set of appropriately formatted and finished tables and figures of the quality expected for a conference presentation or manuscript. More details regarding the final project will be given at the appropriate time.

Students with Disabilities:

The UO is working to create inclusive learning environments. Please notify me if there are aspects of the instruction or design of this course that result in disability related barriers to your participation. I would like to meet with you in privacy the first week of the term to be sure you are appropriately accommodated.

Academic Honesty:

All persons involved in academic dishonesty will be disciplined in accordance with University regulations and procedures.

Advanced Biological Statistics Bi 610 Fall 2017

Lecture Schedule

Week	***Date***	***Subject***	***Readings***	***Assignments***
1	Sep 26th Sep 28th	Tidy data and data wrangling R programming and reproducible data analysis Exploratory Data Analysis (EDA)	W & G: Chps 1-16 Logan: Chps 1-3 Q & K: Chps 1	
2	Oct 3rd Oct 5th	Grammar of Graphics Probability and sampling distributions Estimation & confidence intervals	W & G: Chps 1-16 Logan: Chps 4-6 Q & K: Chps 2-4	Homework 1
3	Oct 10th Oct 12th	Collaborative analysis: Git, GitHub & R markdown Hypothesis testing and significance Modeling basics and ingredients	W & G: Chps 17-29 Logan: Chps 4-6 Q & K: Chps 2-4	
4	Oct 17th Oct 19th	Correlation and covariance Ordinary Linear Models (OLMs) Simple linear regression	Logan: Chps 7-9 Q & K: Chps 5-6	Project 1
5	Oct 24th Oct 26th	Multiple linear regression Model selection Principal Component Analysis (PCA)	Q & K: Chps 15-17	
6	Oct 31st Nov 2nd	Clinical trials and experimental design Revisiting type I & II errors Power and False Discovery Rate (FDR)	Logan: Chps 10 Q & K: Chps 7	Homework 2
7	Nov 7th Nov 9th	Analysis of Variance (ANOVA) Single factor ANOVA Planned and post hoc comparisons	Logan: Chps 10 Q & K: Chps 8	
8	Nov 14th Nov 16th	Multiple factor ANOVA Factorial & Nested ANOVA	Logan: Chps 11-12 Q & K: Chps 9 & 12	Project 2
9	Nov 21st T-giving	Random effects Hierarchical and mixed models ANCOVA	Logan: Chps 11-12 Q & K: Chps 9 & 12	
10	Nov 28th Nov 30th	Frequency analyses Contingency tests Generalized Linear Models (GLMs)	Logan: Chps 16-17 Q & K: Chps 13-14	
Finals Week Week of Dec 5th		**FINAL PROJECT**		***Final Project Due***