Syllabus for BIOL/ECOL 563:Statistical Analysis for Ecology and Evolution Spring 2018
Meeting time: TuTh 2:00-3:15
Location: Graham Memorial 0038

Instructor

James Umbanhowar Teaching Assistant Professor of Biology
Office: Genome Sciences Building 4252 and Mitchell 121, Email: jumbanho@unc.edu
Office hours: Scheduled office hours 9:30-11:30 T TH Mitchell 121.

Description

In this course, we will be developing a nuanced view of analyzing ecological and evolutionary data, covering a wide range of statistical techniques. We will focus on drawing connections from basic statistical concepts to analysis of the types of data encountered in experiments and observations in the field of environmental science. By the end of the course you should be prepared to develop, fit and interpret statistical models for a wide variety of data and experimental types. The emphasis on concepts should also empower you to learn more statistical techniques as needed.
Core learning goals for this course are:

- Learn how to connect probabilistic models to data.
  - Estimate model parameters from data.
  - Evaluate how whether the model is a good fit.
  - Make statistical inference based on model fits.
  - Learn how to learn new statistical techniques.
- Learn how to use statistical software package R to analyze data
  - Learn core R functions and libraries.
  - Learn how to visualize data using R.
  - Learn how to use techniques to make your data analysis reproducible.

Prerequisites

You should have taken an introductory statistics course that covered: basic probability theory, common probability distributions (normal, poisson, binomial), hypothesis testing including standard parametric approaches using normal errors such as the t-test and simple linear regression. If you have gotten a bit rusty, we will spend the first week or so reviewing these basic statistical concepts.

Text

We will be using Ecological Statistics: Contemporary theory and application by G.A. Fox, S. Negrete-Yanklevich and V.J. Sosa as our main text. This text is available on UNC computers here: `http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199672547.001.0001/acprof-9780199672547`

Grading

Course grade will be determined on weekly assignments and a take home final exam.
Weekly assignments: 80%
Final exam (take home assignment) 20%

Grading scale:

Percentage	Graduate grade	Undergraduate Grade
93.3	H	A
90	H	A-
86.7	P	B+
83.3	P	B
80	P	B-
75	L	
<75	F	
76.7		C+
73.3		C
70		C-
60		D
<60		F

Assignments

There will be (approximately) weekly assignments for this course. Assignments for the course will consist of a set of questions that will require coding in R and answers to related to that coding. For the first two assignments you will be handing in two files. One a formatted word processed file with answers to the questions and another text (.R file) with all the code that you used to generate that code. If the questions ask for code, copy only the relevant part of the code and similarly for graphs. In the third week of class, we will be learning how to use Rmarkdown to intersperse code and text into one document and you will turn in that markdown file for your assignment.
I have set up Slack channel for discussions of assignments. You will be added to this channel in the first week of the class. If you do not already have the Slack app, you can get it here: `https://slack.com/downloads/` or use the browser based interface.
You are free to work together on your assignments and final, but you are required to turn in your own work. In particular, interpretative questions should be written independently. Direct copying will be considered plagiarism.

Course Policies

- Class attendance is not required, but is highly recommended. If you do miss class, you will probably want to come to office hours to catch up.
- Late assignments will be marked down 5% per day unless previous permission has been granted. Please tell me early if you have conflicts that preclude finishing assignments on time.
- The final exam will be given out in the last week of class. You will have time to finish it during our final exam period

Honor Code

Academic integrity is a core value of the University and the greater academy of ideas. The student led Honor System is the institution that is responsible for monitoring this code. As such this class adheres to that code and any violations of the code will be reported. You can find more information about the honor code here: `http://catalog.unc.edu/policies-procedures/honor-code/`. As in all classes at Carolina, your are expected to adhere to this code.

Course content

This course covers a number of statistical methods that have proven useful in analyzing ecological data, both observational and experimental. The topics this semester will include the following.
Overview of regression
Likelihood theory and its applications in regression
Generalized linear models
Mixed effects models
Bayesian analysis
Spatial analysis
Missing data
Meta analysis

Preliminary course schedule with reading chapters

January 10 Introduction to class and quick probability refresher (Chap 1)
January 15 Introduction to Statistical inference: sampling, confidence intervals and null hypothesis testing (Chap 2)
January 17 Lab: Getting up to speed with R: linear regression
January 22 Statistical inference: Maximum likelihood and information theoretic approaches (Chap 2)
January 24 Lab: Plotting data with ggplot2
January 29 Generalized linear models: theory (Chap 3)
January 31 Lab: Multiple regression

February 5 Generalized linear models: proportional data and logistic regression (Chap 3)
February 7 Lab: Logistic regression
February 12 GLMS: goodness of fit in GLM's and overdispersion (Chap 3)
February 14 Lab: Count data regressions
February 19 Mixed models: theory (Chap 13)
February 21 Lab: fitting linear mixed models
February 26 Generalized linear mixed models: pragmatism (Chap 13)
February 28 Lab: fitting glmms
March 5 Bayesian inference and Markov Chain Monte Carlo (Chap 2)
March 7 GLMM's with Bayesian models
March 12 SPRING BREAK
March 14 SPRING BREAK
March 19 Missing data theory (Chap 4)
March 21 Lab: Accounting for missingness in data
March 26 Generalized additive models: Fitting data with squiggly lines
March 28 Lab: GAMs
April 2 Research synthesis: Meta-analysis (Chap 9)
April 4 Lab: Conducting a meta-analysis
April 9 Analysis of spatial variation (Chap 10)
April 11 Lab: accounting for spatial autocorrelation in data
April 16 Comparative analysis and phylogenetic correlation (Chap 11)
April 18 Lab: conducting an analysis of phylogenetically correlated data
April 23 Mixture models (Chap 12)
April 25 Lab: Fitting mixture models
May 6: Take home final in class session 12-3pm.