Statistical Methods in Ecology and Evolution
ZOL 851
Fall 2015

Dr. Elise Zipkin
Office: 219A NS Building
Website: www.msu.edu/user/ezipkin
Email: ezipkin@msu.edu
Office hours: Tu, Th 12-1pm

Dr. Joshua Nahum
Office: 1440 BPS Building
Website: www.nahum.us
Email: nahumjos@msu.edu
Office hours: M,W 11-noon

Grader: Dr. Sarah Saunders, Email: saund123@msu.edu
Office hours: By appointment only – 305 NS Building

Description and Objectives
This graduate level survey course focuses on the fundamental elements of data analysis in the fields of ecology and evolution. Students will learn how to interpret and model biological data with computationally intensive methods for estimation and inference using the R language. Topics include probability theory, frequentist and Bayesian inference, linear models, generalized linear models, mixed and random effects, hierarchical models, zero-inflation models, power analyses, and model selection.

Class Location: 319 Berkey Hall Meeting time: Tu, Th 10:20-12:10

Recommended Background
STT465 (Statistics for Biologists II) & STT814 or an equivalent course (CSE845). We assume familiarity with probability distributions (PDFs and CDFs), random variables, regression, and ANOVA. While this course will not deal with formal mathematical proofs, we do assume a certain level of comfort with algebraic concepts and a dim recollection of derivatives (we won't be doing any calculus, but we may look at a few derivatives with respect to maximum likelihood). Some basic linear algebra is also useful. We also assume that you have some familiarity with R or another command line statistical software.

If you do not have the background as suggested above, we just want to make sure that you are aware well ahead of time so that you are not shocked in the course. We are happy to suggest some remedial reading, but be prepared for a challenging semester.

Goals of this course:
1. Give you the foundation to learn how to analyze data for your thesis and for the rest of your career.
2. Provide a thorough introduction to model based analysis approaches that allow for easier translation from statistical to biological inference.
3. Introduce R, a powerful and useful programming environment for data analysis and presentation. Develop skills at writing R functions, with the goal of being able to perform advanced, computationally intensive analyses.
4. Provide you with sufficient and appropriate background to be able to teach yourself new methods for data analysis as needed.

Course website
All class materials will be posted to the course's D2L website as soon as possible following the class period. We will also be using Piazza for class discussion. Piazza is highly catered to getting you help fast and efficiently from classmates, and ourselves. Rather than emailing questions to the teaching staff, we encourage you to post your questions on Piazza: https://piazza.com/msu/fall2015/zol851.

Course books
The course will make use of the following textbooks:

Bolker, B.M. (2008). Ecological models and data in R. Princeton University Press.

Clark, J. S. (2007). Models for ecological data: an introduction. Princeton, New Jersey, USA: Princeton university press.

It is not required that you obtain these books but they may be helpful to you. The textbooks are available in the library. Course lectures will be provided.

Grades
Course grades are based on the following components:

Participation (10%): To earn full credit for participation, students are required to attend class (and inform instructors of absences prior to the scheduled class period), contribute to discussions, and actively participate in in-class group activities and coding sessions. Participation points can also be earned by posting questions and answers to the class Piazza site.

Lab assignments (50%): Most weeks we will be building models using R during class. On some occasions exercises will be assigned and will be due one week later. Students may work together on assignments but each student must submit their own homework. The penalty for late homework will be a 10% daily reduction in possible points. However, we strongly encourage students to submit ALL homework assignments (even if they are late). Reviewing concepts outside of class is the best way to learn the material.

Quizzes (40%): Short quizzes (approximately 20 mins) will be administered regularly throughout the semester. The purpose of the quizzes is to encourage students to review course materials outside of class. Quizzes will generally target concepts whereas homeworks will focus on model implementation and inference (primarily through programming in R). The quizzes will be limited to concepts that were covered in recent class periods. The lowest two quiz grades will be dropped. There will be no make-up quizzes.

Additional suggested references

Burnham, K.P., & Anderson, D.R. (2002) Model selection and multi-model inference: a practical information-theoretic approach. Springer.

Crawley, M.J. (2012). *The R book*. John Wiley & Sons.

Gelman, A. & Hill, J. (2007) Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Kéry, M. (2010) Introduction to WinBUGS for ecologists: a Bayesian approach to regression, ANOVA, mixed models, and related analyses. Academic Press.

Link, W.A., & Barker, R.J. (2009) Bayesian inference: with ecological applications. Academic Press.

Logan, M. (2011). *Biostatistical design and analysis using R: a practical guide*. John Wiley & Sons.

Matthiopoulos, J. (2011). *How to be a quantitative ecologist: the'A to R'of green mathematics and statistics*. John Wiley & Sons.

McCarthy, M.A. (2007) Bayesian methods for ecology. Cambridge University Press.

Royle, J.A., & Dorazio, R.M. (2008) Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities. Academic Press.

Quinn, G.P., & Keough, M.J. (2002). *Experimental design and data analysis for biologists*. Cambridge University Press.

Zuur, A., Ieno, E.N., Walker, N., Saveliev, A.A., & Smith, G.M. (2009). *Mixed effects models and extensions in ecology with R*. Springer Science & Business Media.

November 12: Revisiting modes of inference: Bayesian

November 17: Catch up and consulting (Quiz)

Primary Instructor: Dr. Nahum

November 19: Power analysis

November 24: Power analysis

November 26: No class: Thanksgiving

December 1: Resampling methods (Quiz)

December 3: Model selection

December 8: Model selection

December 10: Catch up and consulting (Quiz)