

Welcome to NRES 746

Advanced Analysis Methods in Natural Resources and Environmental Science

Fall 2019

Instructor

Kevin Shoemaker

Office: FA 220E

Phone: (775)682-7449

Email: kshoemaker@cabnr.unr.edu

Office hours: Wednesdays at 1pm in FA 220e

Course Meeting Times

Lecture & Discussion: M, W at noon in KRC 105 (50 mins)

Lab: Tuesday at 3pm in LME 315 (3 hours)

Course Website

<http://naes.unr.edu/shoemaker/teaching/NRES-746/>

Course Objectives

Modern computers have reduced or eliminated many of the barriers to advanced data analysis, and as a result computational algorithms now often supersede elegant and simple mathematical formulae for complex data analysis. Armed with basic concepts of probability and statistics, and with some facility with computer programming, ecologists and natural resource professionals can get more out of their data than ever before. **In this course, we embrace the primacy of the algorithm.**

By the end of this course, students should have the ability to (1) develop computational routines to simulate data generation under alternative mechanisms, (2) fit these computational models to data using maximum likelihood and Bayesian inference, (3) validate these computational models, and (4) understand where and when to use a wide variety of additional advanced data analysis methods. The goal is for students to emerge from this course as creative data analysts with the tools and intuition needed to draw inferences from a wide variety of data types.

The course motto: *Be Dangerous!* What does that mean?? It is *safer* to use standard analytical tools (e.g., in a software like SAS or SPSS) because these methods have been validated and tested in many ways over the years. When we build our own algorithms, we can be entering uncharted territory. And exploring these territories can be dangerous... And daunting... And exciting!!

The focus of this course is on using computational algorithms to infer ecological processes and relationships from pattern in *observational studies*; *we will not directly address experimental methods or design in this course*. However, the data analysis methods covered will be of broad utility for a wide variety of disciplines and data types, including analysis of experimental data. The general focus will be on model-based inference, including regression-based approaches, hierarchical/mixed models, and multi-model inference. Additional student-led modules will cover other advanced analysis topics such as classification and regression trees, structural equation modeling, survival analysis, and species distributions models. *This is not a "statistics" course per se*; we will focus on the implementation and leave the nitty-gritty stats questions to statisticians.

Each student will be responsible for leading discussions and demonstrations on a data analysis method of their choice (working in groups). The laboratory portion of the class will provide students the opportunity to try out some of the data analysis methods. Structured labs with example data sets will be interspersed with open lab periods where students work in small groups on a research project involving analysis of a real-world data set.

Student Learning Objectives

Students will be able to: 1. Identify and contrast the major classes of statistical models used by ecologists (e.g., Bayesian vs frequentist, likelihood-based, machine learning) and explain how and why ecologists use these models. 2. Apply analysis tools such as logistic regression, non-linear regression, mixed-effects models, and machine-learning algorithms (e.g., Random Forest) on diverse data sets representative of those commonly considered in ecology. 3. Learn to explore data sets quantitatively and graphically and to prepare data appropriately for analysis. 4. Perform statistical analysis, data visualization, simulation modeling, model validation and programming with the statistical computing language R. 5. Critically evaluate the strength of inferences drawn from statistical models by understanding and testing major assumptions and using tools such as cross-validation. 6. Communicate statistical and computational concepts by leading lectures and discussion on advanced topics in data analysis.

Prerequisites

Curious scientific mind, broad research interests, comfort with (or at least, lack of fear regarding) equations and computed programming. Students are expected to already have a solid foundation in standard statistical concepts and methods, obtained through other coursework. If this is not the case, they should be prepared to work harder to develop the necessary prerequisite knowledge.

Textbooks and Readings

We will use the book, *Ecological Models and Data in R*, by Ben Bolker, as a general class reference. However, additional readings will be assigned, and will be available on the course website.

Additional readings will be assigned as indicated in the course schedule (which is ever-evolving!).

Grading

Course component	Weight
Student-led topics	20%
Participation	20%
Laboratories	20%
Research Project, written	30%
Research Project, presentation	10%

Course components

Student-led presentations: Each student will work with a small group to lead a lecture/demonstration that introduces an advanced data analysis method using a worked example (clear, concise, informative tutorial), and provides examples of real-world applications from the published literature. Presenters are encouraged to work with the instructor (and other faculty, graduate students!) to better understand their data, methods, papers and topics.

Class Participation: Students are expected to actively participate in the classroom education process. Don't be afraid to ask questions- fear of embarrassment can be a major impediment to learning. So consider this a safe space for making mistakes- *be dangerous AND fearless!*

Laboratory Reports: Students will submit (1) R functions that accomplish specific tasks (which will be graded using an automated algorithm- in R of course!), and (2) a brief written report (in Word or Google Docs, submitted via WebCampus) succinctly answering any questions, and stating any questions or points of confusion. While students are encouraged to work on the labs in small groups, all lab submissions must be made individually.

Group Projects: Students will work on projects in groups of 2 - 4. Projects will require analysis of previously published or publicly available data sets that are NOT intended to be part of a student's planned thesis or dissertation chapters. The instructor can assist with identifying suitable data sets. Although a primary goal is to enhance knowledge and facility with the data analysis methods, an important secondary goal could be to develop a collaborative manuscript for publication! Therefore, careful thought should go into choice of a data set and relevant scientific questions to guide the analysis.

The group project will take the form of a manuscript for submission as a research paper (with fully worked methods and results and brief intro and discussion: see below). This will be submitted to the instructor as a complete draft by Dec 6, 2018, and (after review and comment by the instructor) as a final version by December 18.

Group projects: expectations

Students are expected to perform (and report on) a data analysis using *state-of-the-art analytical methods*. The write-up will loosely take the format of a scientific paper to be

submitted to a professional journal. However, because of the nature of this course, the most important pieces of the write-up are the **methods** and **results** sections (with several publication quality figures, of course!). Nonetheless, I expect at least a few paragraphs introducing the topic and why it's important (introduction section), and a few paragraphs discussing the implications of the results (discussion section). The methods and results section can (and in many cases should) be longer than you typically see in a scientific paper- don't feel constrained by space for these sections! Not that you need to be wordy, I just want to make sure you have the space to clearly explain the analyses you performed and why you made the choices you did!

Here is a more detailed description of expectations for the final group project:

Introduction: Provide enough description so that the reader understands why the research is important and (if appropriate) what research question(s) and/or hypotheses are being addressed. (ca. 3 paragraphs)

Methods:

- Provide just enough details about the data collection to give the reader the context necessary to understand the data.
- Provide plenty of detail about the analytical approach- enough detail to replicate the analysis. Justify all decisions that were made and (where appropriate) discuss why you did not use alternative approaches.
- Discuss key analytical assumptions and how they were tested. How did you assess model adequacy? Did you attempt any other approaches?

NOTE: This section can be longer than the Methods section of a standard manuscript.

Results: Present all relevant results completely and concisely. Wherever possible, results should be presented via figures and tables. *There is a limit of 5 figures and 3 tables, so choose carefully which figures and tables to present. Figures should be *publication quality*.*

Discussion: Write at least three paragraphs that put the results in a larger context (returning to the key research questions) and discuss areas of uncertainty. Potential topics are possible violations of assumptions, and future work that your analysis suggests would be profitable.

Supplement Provide all code used to run the analyses presented in the paper as an R script or GitHub link.

Course Schedule

NOTE: the course schedule is not set in stone, so please check back frequently!

Note: this schedule is subject to change. Please check for updates frequently!

Week	Lecture (Mon)	Lecture (Wed)	Lab (Tues)	Final project	Readings
Aug. 26	Course Introduction	Writing algorithms	Programming algorithms in R	Start getting into groups, identify dataset	Clark Ch. 1; Touchon and McCoy 2016
Sept. 2	NA (labor day)	Working with probabilities	Programming algorithms in R (continued)		Bolker ch. 4; Zurell et al. 2010;
Sept. 9	The Virtual Ecologist	The Virtual Ecologist	Final project #1	Finalize groups and project ideas	Bolker Ch. 1, Ch 5.; Zuur et al. 2010 (optional)
Sept. 16	Likelihood	Likelihood	"Virtual Ecologist" lab	Work on one-page project description	Bolker Ch. 6; Hobbs and Hilborn 2006 (optional)
Sept. 23	Likelihood	Optimization	"Virtual Ecologist" lab (and project work)	DUE: one-page project descriptions	Bolker Ch. 7
Sept. 30	Optimization	Bayesian inference	Final project #2	Review proposals with instructor	Bolker Ch. 6 and 7 (Bayesian section); Ellison 2004
Oct. 7	Bayesian inference	Markov Chain Monte Carlo (MCMC)	Maximum likelihood	Start running analyses and generating figures	Bolker Ch. 7 and 8
Oct. 14	Markov Chain Monte Carlo (MCMC)	Model selection and multi-model inference	Maximum likelihood		Anderson et al. 2000; Anderson et al. 2001; Wintle et al. 2003
Oct. 21	Model selection and	Model validation and	Final project #3		TBD

Week	Lecture (Mon)	Lecture (Wed)	Lab (Tues)	Final project	Readings
	multi-model inference	performance evaluation			
Oct. 28	Model validation and performance evaluation	Machine-learning?	Bayesian model fitting in JAGS		TBD
Nov. 4	TBD	TBD	Bayesian model fitting in JAGS		TBD
Nov. 11	NA (veteran's day)	TBD	Final project #4		TBD
Nov. 18	TBD	TBD	Optional: Model selection and performance evaluation		TBD
Nov. 25	TBD	TBD	Final project #5		TBD
Dec. 2	TBD	TBD	Complete "mini-lab" assignments	Final project complete draft due this week	TBD
Dec. 9	Class wrap-up	NA (prep day)	Final project presentations		
Dec. 18	NA (classes over)	NA	NA	Final projects due	

Make-up policy and late work:

If you miss a class meeting or lab period, it is your responsibility to talk to one of your classmates about what you missed. If you miss a lab meeting, you are still responsible for completing the lab activities and write-up on your own time. You do not need to let me know in advance that you are going to miss class or lab.

Students with Disabilities

Any student with a disability needing academic adjustments or accommodations is requested to speak with the Disability Resource Center (Thompson Building, Suite 101) as soon as possible to arrange for appropriate accommodations.

Statement on Academic Dishonesty

Cheating, plagiarism or otherwise obtaining grades under false pretenses constitute academic dishonesty according to the code of this university. Plagiarism is using the ideas or words of another person without giving credit to the original source; this includes copying another student in class. Always cite the source of your information. This includes copying or paraphrasing from a book, journal, or unpublished material without giving credit to the author(s), and submitting a term paper that was used in another course. Academic dishonesty will not be tolerated and penalties can include filing a final grade of "F"; reducing the student's final course grade one or two full grade points; awarding a failing mark on the coursework in question; or requiring the student to retake or resubmit the coursework. For more details, see the University of Nevada, Reno General Catalog.

This is a safe space

The University of Nevada, Reno is committed to providing a safe learning and work environment for all. If you believe you have experienced discrimination, sexual harassment, sexual assault, domestic/dating violence, or stalking, whether on or off campus, or need information related to immigration concerns, please contact the University's Equal Opportunity & Title IX Office at 775-784-1547. Resources and interim measures are available to assist you. For more information, please visit: <http://www.unr.edu/equal-opportunity-title-ix>

Statement on Audio and Video Recording

Surreptitious or covert video-taping of class or unauthorized audio recording of class is prohibited by law and by Board of Regents policy. This class may be videotaped or audio recorded only with the written permission of the instructor. In order to accommodate students with disabilities, some students may have been given permission to record class lectures and discussions. Therefore, students should understand that their comments during class may be recorded.