



Lecture 0

Course overview

WILD6900 (Spring 2020)

LOGISTICS

Lecture: Tuesday and Thursday 1:30-2:45 ENLAB 248

Credits: 3

INSTRUCTOR

Dr. Clark Rushing

clark.rushing@usu.edu

Office: NR 146

Office hours: Monday and Wednesday 1:00-2:30 (or by appointment)

COURSE FORMAT

Refer to [syllabus](#)

WHY BAYESIAN?

Why Bayesian?

Philosophical advantages

- 1) Probabilistic treatment of all unknown quantities
- 2) Coherent framework for incorporating prior knowledge into analysis
- 3) Proper accounting of uncertainty
- 4) Ease of estimating latent variables (and uncertainty)

Why Bayesian?

Practical advantages

- 1) Ability to develop custom/complex models to suite your needs
- 2) Many statistical concepts (e.g., random effects) make more sense (no blackbox)
- 3) Expanded "toolkit" for quantitative analysis
- 4) Ability to keep up with the literature
- 5) Ability to review manuscripts/proposals that use Bayesian methods

Why Bayesian

Disadvantages

- 1) Computationally intensive
- 2) Few (no?) "canned" software

Bayesian methods in ecology

The use of Bayesian methods is growing rapidly in ecology, largely due to:

- powerful computers and the development of accessible software
- large amount of code online and in books that be used as the starting point for analyses

These developments have led to many people adopting Bayesian methods even if they don't fully understand what they're doing. They also led to some ecologists becoming full-on Bayesians and then a predictable backlash from those that view these methods as needlessly complex and overly trendy (i.e., statistical machismo)

The methods you'll learn about in this class, just like all statistical methods, are *tools* to help you answer questions

Your job as a researcher is to choose the tools that best suite your question and your data. The goal of this course is to **expand your toolbox**

Why are you interested in learning
Bayesian methods?

REPRODUCIBLE RESEARCH

Reproducible research

What makes research "reproducible"?

Many definitions (Goodman et al. 2016) but for the purposes of this class, we will define it as:


The ability of independent researchers to reproduce scientific results using the original data and methods (adapted from Markwick et al. 2018)

- Reproducibility is a gradient from unreproducible (no data, no methods, etc.) to completely reproducible (the ability to start with raw data and reproduce all results, figures, tables, etc) ¹.
- At a minimum, reproducibility means someone could re-run your analyses and come up with exactly the same answer

Why make your research reproducible?

Making your work reproducible involves extra work. Why bother?

- 1) To help advance science - if your work can't be reproduced, it's not science
- 2) To meet requirements of journals/granting orgs
- 3) To make it easier to share your work with collaborators
- 4) To make it easier to revise your analysis later



You always have at least one collaborator on every project - you future self. And your past self doesn't respond to email

Course philosophy

- Stats courses are the primary source of experience for learning to work with data
- But in-class examples often ignore the fact that **most** time will be spent wrangling raw data to get it ready for analysis
- Because data wrangling skills aren't taught, students (often) develop bad habits
- In the "lab" activities in this course, we will emphasize the tasks of cleaning and processing raw data to prepare it for analysis

Additional resources

- Cooper, N. & Hsing, P. (2017) *A guide to reproducible code* British Ecological Society