Introduction to Bayesian modeling with Stan

Spring 2017

Instructor:	Dr. Zachary Marion	**Time:**	MW 4:40–6ish PM
Email:	zmarion@utk.edu	**Place:**	488 Dabney
Office:	435 Hesler		

Office Hours: After class or by appointment.

Main References: This is a list of suggested books that will be useful throughout this course:

- Richard McElreath, *Statistical Rethinking*, Chapman & Hall, 2016.
- John Kruschke, *Doing Bayesian Data Analysis, 2nd Edition*, Academic Press, 2014.
- Andrew Gelman, John Carlin, Hal Stern, David Dunson, Aki Vehtari, & Donald Rubin, *Bayesian Data Analysis, 3rd Edition*, Chapman & Hall, 2014.

GitHub Repository: `https://github.com/zmarion1/BayesianGymnasium2017.git`

Objectives: Bayesian data analysis is an old field of statistics, but recent computational advances—as well as the use of Markov-Chain Monte Carlo (MCMC) samplers—have dramatically increased the utility of the method. This class is aimed at introducing Bayesian modeling to graduate students who are already familiar with traditional (frequentist) methods of analysis. The goal is to take the mystery out of hierarchial Bayesian modeling through practical applications and examples rather than diving into the theory and math (though some is required). To this end, the class takes a computational approach using the Stan programming language in R. Parameter estimation and biological interpretation is stressed above all. We will cover simple Bernoulli and binomial models, what hierarchy means, linear regression, Bayesian model selection, generalized linear multilevel modeling, and posterior simulation.

Prerequisites: The class is primarily geared towards graduate students who have had one or two semesters of statistics (e.g., EEB 560 Biometry) and are familiar with (but probably rusty on) linear algebra and calculus. A working knowledge of functional computer programming in R is **strongly** recommended. Students should be comfortable with writing for-loops and multidimensional indexing in R.

Course Outline:

- Introduction: What is Bayes? How is it different than likelihood?
- Simple Bernoulli and binomial models
- Hierarchical models and why they're awesome!
- Simple and multiple linear regression
- Maximum entropy and model selection using loo and WAIC
- Generalized linear modeling
- Hierarchical modeling: varying intercepts & varying slopes
- Advanced topics: Missing data, mixture modeling, autocorrelation, and tricks

Software: We will be using the statistical programming language R. R is free (https://www.r-project.org) and is fast becoming established as the language of choice for statistics. We will also be using the Stan programming language (mc-stan.org) via the RStan package.