ERTH 695 BAYESIAN DATA ANALYSIS (R, JAGS & STAN)

The purpose of the course is to help each student develop a Bayesian foundation for analyzing her own data set. The emphasis is on software rather than theory, especially on Monte Carlo software such as JAGS, and Stan. These free, open-source languages enable one to solve difficult data analysis problems without being an expert in either coding or calculus. Knowledge of any particular field of science isn't required for this course – the business of fitting data to models spans all areas of science, and knowing the statistical models used in other fields opens a world of possibilities. As part of the course, students will learn R Markdown, an easy to use facility for literate programming in which the text and code for a research report are in the same document. Weekly coding assignments will be in the form of R Markdown documents.

Instructor: Neil Frazer, neil@hawaii.edu, **Credits:** 3. **Prerequisite:** Instructor approval, but students should already have taken, at a minimum, an introductory course in statistics and an introductory course computer programming. **Meetings**: Two online 80-min meetings/wk. **Hours**: TR, 12:00-1:15 **Grading**: 80% on weekly coding exercises, 10% on class participation, and 10% on progress with student research if applicable. In view of the pandemic, Zoom and email interaction with the instructor is encouraged between classes.

Required texts (free online):
JAGS User Manual,
Stan Modeling Language User's Guide and Reference Manual

Syllabus
• The R language • RMarkdown • Review of distributions (PMFs, PDFs) • Samples as proxies for distributions • Parametric and non-parametric methods • Cox's postulates • Bayes Rule • Conjugate priors • the beta-binomial example • Markov Chain Monte Carlo • Modeling with JAGS • Count data • Metric data with metric predictors • Metric data with nominal predictors • Hierarchical models • Model comparison and selection • Modeling with Stan • Generalized linear models • Logistic regression (dichotomous data with multiple metric predictors) • Multinomial logistic regression (softmax regression, nominal data)

Supplementary reading:
(1) Kruschke (2015) *Doing Bayesian Data Analysis, 2nd edition*.
The level of detail in this fine text can be annoying to sophisticated readers, but it is Bayesian from the get-go, it includes all the code, and the writing is so friendly and typo-free that you cannot help but love the book despite its size. Its many well-thought-out exercises make it a painless way to learn R as well as JAGS and Stan.
(2) Lunn et al. (2013) *The BUGS Book: A Practical Introduction to Bayesian Analysis*, (a marvelously diverse compendium of models with accompanying BUGS code).
(3) Hobbs & Hooten (2015) *Bayesian Models: A statistical Primer for Ecologists*, (no code, but a beautifully well-organized, consistent and comprehensive treatment of ecological models with a foolproof system for constructing the directed acyclic graphs needed for complex problems).

Additional resources:

McElreath (2015) *Statistical Rethinking: A Bayesian course with examples in R and Stan*, (an excellent, very accessible book, right up there with Kruschke (2015), and at about the same level, with many great student exercises and lots of R and Stan code.)

Aster et al. (2013) *Parameter Estimation and Inverse Problems* (2nd ed.) (introductory, accessible, written by geophysicists for geophysicists, Matlab based).

Sivia & Skilling (2006) *Data Analysis: A Bayesian Tutorial* (2nd ed.), (introductory, but very powerful, written for physical scientists; original work by Skilling on nested sampling toward the end; no R, JAGS or BUGS;)

Lancaster (2004) *An Introduction to Modern Bayesian Econometrics*, (beautifully written; S+ and BUGS are treated in appendices with a selection of useful examples; pre-JAGS)

Gelman et al. (2014) *Bayesian Data Analysis (3rd ed.)*, (advanced, very comprehensive; no code, but see Gelman and Hill, next)

Gelman and Hill (2007) *Data Analysis Using Regression and Multilevel/Hierarchical Models* (comprehensive, pre-JAGS, pre-Stan, many representative examples from the social sciences; lots of student exercises, powerful R code, BUGS code and practical advice)

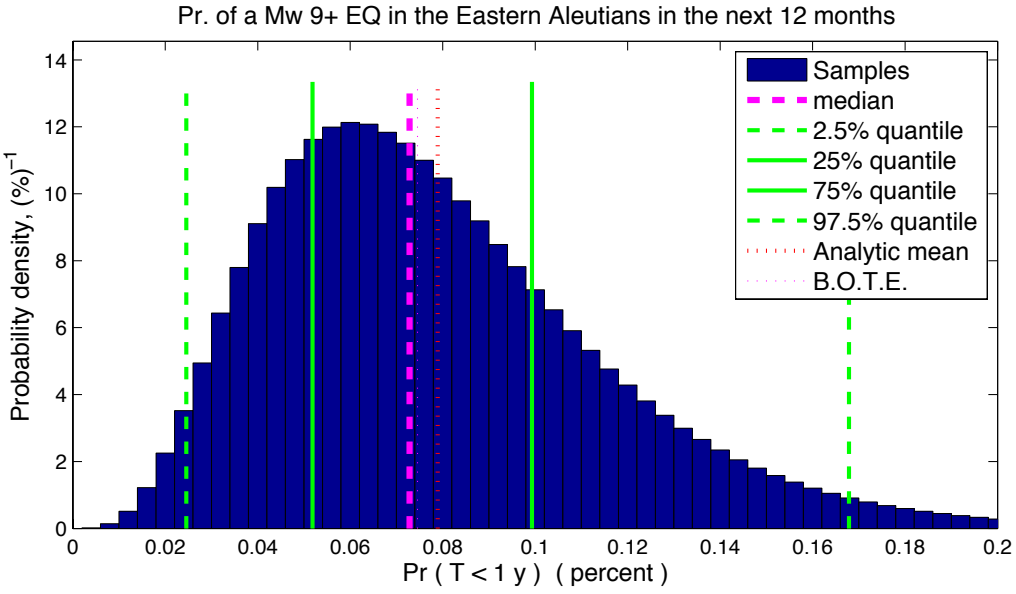
Jackman (2009) *Bayesian Analysis for the Social Sciences* (lots of JAGS code; mainly social science examples but this is also an excellent book for physical scientists because of its exceptionally comprehensive treatment of theory including the multivariate normal)

Cressie & Wikle (2011) *Statistics for Spatio-Temporal Data,* (comprehensive, no codes; difficult, yes, but this is frontier stuff and life is never easy on a frontier)

Fenton & Neill (2013) *Risk Assessment and Decision Analysis with Bayesian Networks*, (very easy and accessible–for example, the Σ and Π notations are explained–but most variables are recoded as factors. The authors' software product, AgenaRisk is available free at http://www.agenarisk.com).

Royle & Dorazio (2008) *Hierarchical Modeling and Inference in Ecology*, (lots of nice models, but the accompanying code-package is reported to have bugs, so *caveat lector*).

Tarantola (1987) *Inverse Problem Theory*, (Published by Elsevier in hard cover, but Albert made sure there was a freely available version on the web long before such things were popular. He was a genius, far ahead of his time. The book may be downloaded from http://www.ipgp.fr/~tarantola/)



Bayes easily addresses difficult questions such as: Should I purchase a PFD?

Class Format:
This is a combined lecture-laboratory. Students are encouraged to actively ask questions in class, to work on weekly assignments in class, and to assist each other in learning, although copying of work is not permitted. In view of the pandemic, Zoom and email interaction with the instructor is encouraged between classes.

Disability Access:

If you have a disability and related access needs the Department will make every effort to assist and support you. For confidential services students are encouraged to contact the Office for Students with Disabilities (known as "Kokua") located on the ground floor (Room 013) of the Queen Lili'uokalani Center for Student Services:

KOKUA Program; 2600 Campus Road; Honolulu, Hawaii 96822 Voice: 956-7511; Email: kokua@hawaii.edu ; URL: http://www.hawaii.edu/kokua

Discrimination:

The University of Hawai'i is committed to providing a learning, working and living environment that promotes personal integrity, civility, and mutual respect and is free of all forms of sex discrimination and gender-based violence, including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence, and stalking. If you or someone you know is experiencing any of these, the University has staff and resources on your campus to support and assist you. Staff can also direct you to resources that are in the community. Here are some of your options:

As members of the University faculty, your instructors are required to immediately report any incident of potential sex discrimination or gender-based violence to the campus Title IX Coordinator. Although the Title IX Coordinator and your instructors cannot guarantee confidentiality, you will still have options about how your case will be handled. Our goal is to make sure you are aware of the range of options available to you and have access to the resources and support you need.

If you wish to remain ANONYMOUS, speak with someone CONFIDENTIALLY, or would like to receive information and support in a CONFIDENTIAL setting, use the **confidential resources available here**: http://www.manoa.hawaii.edu/titleix/resources.html - confidential

If you wish to directly REPORT an incident of sex discrimination or gender-based violence including sexual assault, sexual harassment, gender-based harassment, domestic violence, dating violence or stalking as well as receive information and support, contact: Dee Uwono Title IX Coordinator (808) 956-2299 t9uhm@hawaii.edu.