OCEANS 140: Statistical Modeling
Syllabus, Winter 2025

Course description

Intermediate course focused on statistical modeling in a Bayesian framework, with applications in the biological and environmental sciences. Topics will include probability, causal inference, and generalized linear models. Classroom activities will be a mix of lecture, discussion, and problem sets. We will take a hands-on, computational approach (R, Stan) to gain intuition so that students can design their own inferential models. Outside of class, students will watch pre-recorded lectures and complete readings. Prerequisites: Introductory statistics course, such as STATS 60, STATS 110, STATS 141 and some experience with scientific computing in R, such as STATS 32.

Learning objectives

This course is a practical introduction to statistical modeling in a Bayesian framework. By the end of the course, you will be able to:

- Understand the basics of probability necessary for statistical inference
- Describe the components of Bayes theorem and their relevance for model specification
- Develop and diagnose Bayesian statistical models for prediction and inference
- Confidently learn new approaches to data analysis as you progress in your research

Instructor

Dr. Robin Elahi (elahi@stanford.edu)

Times and location

10:30am-12:00pm; T, Th in Agassiz 12

Office hours

11:00am-12:00pm; W in Agassiz 18 (or email for appointment)

Course structure

We will meet for 1.5 hours, twice a week. Prior to class, students will watch a video lecture, complete a reading, or complete a problem set. In class, we will review the assignment together (led by a student), and then launch into the day's activity. Please bring your laptop to all classroom sessions.

Course website

Homework will be submitted through: https://canvas.stanford.edu/

Course texts
We will use the following text, available online:

Statistical Rethinking, 2nd edition (McElreath 2020):
http://xcelab.net/rm/statistical-rethinking/

Book can be accessed here via Stanford's library:
https://searchworks.stanford.edu/view/13631911

McElreath's lectures and slides from his most recent course here:
https://github.com/rmcelreath/stat_rethinking_2024

Evaluation and grading
Here is a general breakdown of graded tasks:

- Homework (70%)
- Discussion lead (10%)
- Participation (10%)
- Adhering to code of conduct (10%)

Homework will be completed prior to our class meeting on Tuesday. Answers to problems should include the code necessary to recreate the answer, where relevant. Homework will be graded for completion only – you do not have to get the correct answers, but you will demonstrate that you made an honest effort (e.g., with an explanation of your answer; code). I encourage you to take advantage of the solutions provided by the author (or other answers online) – but only *after* you submit your own answers. You will learn much more if you make mistakes and then fix them. You are welcome to work together in groups, but please submit your own work. Upload your homework through Canvas. The preferred format is PDF; but you can also submit a .Rmd or .Qmd file.

Once during the quarter, you will lead a discussion of the video and/or homework. This may involve a discussion, code review, or some blackboarding. Please see me ahead of your discussion lead date to briefly review the plan.

It is important that you watch the videos, work through the code, and complete the homework in a timely manner because we are building a community of statistical learners. You are expected to be an active colleague during our class meetings. I anticipate, many questions, failures, and growth spurts as we move through this difficult but intellectually satisfying material together. You should expect to allocate 6-8 hours per week outside of class to dig into the material.

Before class
Before the first class please download and install the latest versions of R (https://www.r-project.org/) and RStudio (https://www.rstudio.com/products/rstudio/).

To fit our models, we will be using the 'rethinking' package which relies on Stan, as well as several accessory packages. For installation, please refer to the following website (note that you will need the full 'rethinking' package with 'cmdstan'.

- https://github.com/rmcelreath/rethinking

If you have never used R, it is imperative that you get up to speed with the basics of R programming before the first day of class. I recommend working through this set of Data Carpentry lessons. It will take about two full days to complete, if you are starting from scratch:

- https://datacarpentry.org/R-ecology-lesson/index.html

Resources for learning R abound on the web, but here are some other starting points:

- https://education.rstudio.com/learn/beginner/
- https://ourcodingclub.github.io/course.html

Tips for success
- Do the assignment before class
- Be an active participant in class
- Ask questions – of me and your classmates. We will all benefit!
- Do not procrastinate. Don't let a week pass with unanswered questions as it will make the following week's material more difficult to follow

Code of conduct
I will maintain an inclusive learning environment that is productive and fun. Here are some simple rules for conduct:
- Use welcoming and inclusive language
- Be respectful of different viewpoints and experiences
- Assume best intentions
- Gracefully accept constructive criticism
- Focus on what is best for the community
- Show courtesy and respect towards other community members
- Make space, take space

Additional resources
Bolker, B.M. 2008. *Ecological Models and Data in R*. Princeton University Press.
Devlin, D., Guo, J., Kunin, D., and Xiang, D. 2018. *Seeing Theory*.
https://seeing-theory.brown.edu/index.html

Edge, M.D. 2019. *Statistical Thinking from Scratch: A Primer for Scientists*. Oxford University Press.
Hobbs, N.T. and Hooten, M.B. 2015. Bayesian Models: *A Statistical Primer for Ecologists*. Princeton University Press.
Johnson, A.A., Ott, M.Q., and Dogucu, M. 2022. *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. CRC Press. https://www.bayesrulesbook.com/index.html
Kruschke, J. 2014. *Doing Bayesian Data Analysis: A tutorial with R, JAGS, and Stan*.

Schedule

Week	Date	Topic	Pre-class	SR2	Due
1	Tue., Jan. 7	Intro to course; intro stats and coding review	Science Before Statistics	1	
1	Thu., Jan. 9	Components of a Bayesian model	Garden of Forking Data (to 0:45)	2	
2	Tue., Jan. 14	Sampling the imaginary	Garden of Forking Data (0:45-1:16)	3	HW1
2	Thu., Jan. 16	Intro to linear models	Geocentric Models	4.1-4.4	
3	Tue., Jan. 21	Categorical effects, polynomial regression, and splines	Categories and Curves (to 1:14)	4.5-4.6; 5.3	HW2
3	Thu., Jan. 23	Multiple regression and intro to causal inference	Elemental Confounds	5.1, SR 6	
4	Tue., Jan. 28	(More) causal inference	Good and Bad Controls	6	HW3
4	Thu., Jan. 30	Overfitting, regularization, model comparison	Overfitting	7.1, 7.3-7.6	
5	Tue., Feb. 4	Markov chain Monte Carlo	MCMC	8.1, 9	HW4
5	Thu., Feb. 6	Generalized linear models; Binomial regression	Modeling Events (to 1:16)	10.2, 11.1	
6	Tue., Feb. 11	Confounds; Poisson regression	Counts and Confounds	11.2	HW5
6	Thu., Feb. 13	Ordered categorical outcomes	Ordered Categories	12	
7	Tue., Feb. 18	Multilevel models	Multilevel Models	13	HW6
7	Thu., Feb. 20	Multilevel models	Multilevel Adventures	13	
8	Tue., Feb. 25	Multilevel models	Correlated Features	14	HW7
8	Thu., Feb. 27	Continuous categories	Gaussian Processes	14.5	
9	Tue., Mar. 4	Measurement error	Measurement	15.1	HW8
9	Thu., Mar. 6	Missing data	Missing Data	15.2	
10	Tue., Mar. 11	Scientific models	Generalized Linear Madness	16	HW9
10	Thu., Mar. 13	Course wrap-up	Horoscopes	17	

Plagiarism, dishonesty, and academic misconduct
It is expected that Stanford's Honor Code will be followed in all matters relating to this course. You are encouraged to meet and exchange ideas with your classmates while studying and working on homework assignments, but you are individually responsible for your own work and for understanding the material. Passing anyone else's scholarly work, which can include: written material, computer code, exam answers, graphics or other images, and even ideas as your own, without proper attribution, is considered academic misconduct. Plagiarism, cheating, and other misconduct, including bullying, discrimination, and harassment, are serious violations of the University's *Fundamental Standard* and *Honor Code*:
https://communitystandards.stanford.edu/policies-and-guidance

Affordability of course materials
Stanford University and its instructors are committed to ensuring that all courses are financially accessible to all students. If you are an undergraduate who needs assistance with the cost of course textbooks, supplies, materials and/or fees, you are welcome to approach us directly. If would prefer not to approach us directly, please note that you can ask the Diversity & First-Gen Office for assistance by completing their questionnaire on course textbooks & supplies: http://tinyurl.com/jpqbarn or by contacting Joseph Brown, the Associate Director of the Diversity and First-Gen Office (jlbrown@stanford.edu; Old Union Room 207). Dr. Brown is available to connect you with resources and support while ensuring your privacy.

Students with documented disabilities
Students who may need an academic accommodation based on the impact of a disability must initiate the request with the Office of Accessible Education (OAE). Professional staff will evaluate the request with required documentation, recommend reasonable accommodations, and prepare an Accommodation Letter for faculty. Unless the student has a temporary disability, Accommodation letters are issued for the entire academic year. Students should contact the OAE as soon as possible since timely notice is needed to coordinate accommodations. The OAE is located at 563 Salvatierra Walk (phone: 723-1066, URL: https://oae.stanford.edu/).