Syllabus
EES 4891/5891: Bayesian Statistical Methods

Jonathan Magnolia Gilligan
Vanderbilt University

Spring 2026

Contents

1 Nuts and Bolts

1.1 Class Meetings

MW 8:40–9:55 Classroom SC5737 (the Jewell Room)

1.2 Professor

Jonathan Magnolia Gilligan
Professor of Earth & Environmental Sciences
Professor of Civil & Environmental Engineering
Professor of Climate & Environmental Studies
Office: Stevenson 5735 (Stevenson #5, 7th floor),
`jonathan.gilligan@vanderbilt.edu`
`www.jonathangilligan.org`
Office Hours: Tuesday 1:30–2:30, Wednesday 1:30–3:00, or by appointment.

1.3 Email

If you want to communicate with me be sure to begin the subject line of your email with “EES 5891” This helps assure that I will see your message quickly and respond to it.

I have set my email reader to flag all messages like this as important, so I will read them first. This also assures that I do not mistake your email for spam. I typically receive over 100 emails per day, so if you do not follow these instructions, I may not notice your email.

1.4 Course web site

In addition to Brightspace, I have set up a companion web site for this course at `https://ees5891.jgilligan.org`, where I post the reading and homework assignments, my slides from class, and other useful material. That web site will be the central place to keep up with material for the course during the semester. This web site will direct you to Brightspace if there is anything you need to find there.

2 Course Description

2.1 Concise Description

The class will begin with an introduction to Bayesian statistics and then focus on practical application of regression methods to data. We will use R together with the Stan software package (`https://mc-stan.org`) for Hamiltonian Monte Carlo methods. The course will combine practical applications of Bayesian methods to real (often messy) data with more philosophical discussions of Bayesian approaches to statistics and how to interpret results of statistical analyses. We will focus on regression methods, including hierarchical or multilevel regression modeling methods, which can be very powerful when you have data that has a nested structure (e.g., cities and counties within states or species within genera). Students will do projects applying Bayesian methods to their own data sets.

2.2 Prerequisites

You should be comfortable with differential and integral calculus and have some previous experience with standard statistics.

This course will be very mathematical and will make extensive use of the R software system, but I do not assume that you already know R or advanced mathematics beyond calculus.

2.3 Narrative Description

Bayesian statistics is a branch of statistics that has been around for almost 300 years, but for most of that time, it was very difficult to apply to practical problems because the mathematical equations were too difficult to solve. In the last 30 years, as computers have become much faster and more powerful, new computational methods have emerged that make Bayesian statistics practical for research and applications.

Bayesian analysis is widely used across a wide variety of research as well as practical applications. It is used to analyze results from high-energy particle physics experiments to discover new subatomic particles. There are many other applications in a wide variety of domains. It's used by geologists to improve estimates of mineral distributions and radon hazards. It's used by biologists to identify and categorize variations in the genomes of humans and other species. It's used extensively in medicine to analyze the results of clinical trials, to determine the pharmacokinetics of drug metabolism, and to assess the predictive value of tests for diseases such as cancer or COVID infection. It's used in political science and sociology to improve the accuracy of public opinion surveys and to understand patterns of voting. It's widely used in marketing to identify consumer preferences and improve the effectiveness of advertising. If you use Google, Amazon, Netflix, Stitchfix, or practically any large online platform for shopping or entertainment, advanced Bayesian methods form the basis of their recommendations. Bayesian analysis has also been applied effectively to law and criminology to assess the value of evidence in proving guilt or innocence. It has been applied to public health to estimate the prevalence of diseases and to make more effective treatment decisions when medical tests are uncertain. It is widely used in meteorology to make weather forecasts and in climate science to combine data from many different sources and come up with quantitative predictions and detailed understanding of their associated uncertainties. Bayesian methods are also widely used in computational applications, such as image analysis and reconstruction, computational text analysis, and natural language processing. One of the earliest practical applications of

Bayesian textual analysis, in 1964, identified the anonymous authors of the Federalist Papers. More recent applications of Bayesian textual analysis are used to separate desired email from spam.

Bayesian statistical methods are valuable because they provide a systematic way to combine what you already know about a problem with new data from experiments or observations, and the results of Bayesian analyses are more straightforward to interpret than conventional statistics.

This course will provide a general introduction to Bayesian statistics and will combine practical instruction in how to do Bayesian data analysis and philosophical discussions about how to think about the assumptions that go into a Bayesian analysis and how to interpret the results that it produces.

You do not need to have any prior knowledge of computer programming, but I do expect that you are familiar with basic statistics and calculus (both derivatives and integrals).

3 Goals for the Course

By the end of the semester, you will:

- Understand Bayes's theorem and how to apply it.
- Understand problems with the traditional statistical emphasis on null-hypothesis significance testing (NHST), why Bayesian approaches to NHST don't solve these problems, and how Bayeian statistics offers superior alternatives to NHST.
- Understand how think about statistical models, how to choose an appropriate model for your problems, and understand the tradeoffs between different kinds of models.
- Be able to design and conduct a comprehensive Bayesian analysis of data from start to finish.
- Understand how to choose appropriate priors for your Bayesian analyses and how to test whether your choice of priors is sound.
- Understand how to set up and perform Bayesian regression analysis, assess the validity of the analysis, and interpret the results.
- Understand why Markov Chain Monte Carlo (MCMC) sampling is used in Bayesian analysis, what the limits of MCMC are, and how to test your MCMC analyses for validity.
- Understand and be able to perform analyses using more complex statistical models, such as interaction models, generalized linear models, models of discrete (categorical and count) data.
- Understand what multilevel or hierarchical models are, when to use them, and how to interpret the results of a multilevel analysis.

4 Structure of the Course:

I divide the semester into three parts:

1. **Introduction to Bayes's Theorem and its Applications:** The first part of the course introduces the basic concepts of Bayesian statistics, using simplified approximations to calculate difficult equations. This section will focus on linear regression methods.

2. **Monte Carlo Methods:** Next, we study Monte Carlo methods, which help us solve more difficult problems that our earlier approximations are not powerful enough for.

3. **Advanced Applications of Bayesian Methods:** This section will introduce statistical models of discrete data (counts, categories, etc.), and generalized linear models. It will conclude with multilevel statistical models, which can be very powerful methods for working with large and complex data sets.

4.1 Reading Material

There is one required textbook and two optional books. Supplementary reading on the Internet or in handouts will also be assigned during the term and posted on Brightspace.

Required Reading Materials

- Richard McElreath, *Statistical Rethinking: A Bayesian Course with Examples in R and Stan,* 2nd ed. (CRC Press, 2020; ISBN 978-0-367-13991-9). This will be the principal textbook for the semester. Be sure you get the second edition because it is significantly different from the first.

  There is a PDF version of *Statistical Rethinking* on the web that many people download. It is substantially similar to the printed version that we are using for our official textbook, but the page numbers are different.

  To avoid confusion, I have posted PDF versions of the homework exercises from the printed book, which you can download when you do the homework.

  There is a companion web site to *Statistical Rethinking* at `github.com/rmcelreath/stat_rethinking_2024`, which has links a number of resources, and the author has posted videos of his lectures on YouTube.

  McElreath uses basic R, which is fine, but many people have learned to use a more modern dialect of R called the "`tidyverse`," which is described at length in our companion book, *R for Data Science*.

  For people who are familiar with R and like to work in the `tidyverse` dialect, there is a free companion e-book on the web at `bookdown.org/content/4857/`, that has translated almost all the R code in the book into the `tidyverse` dialect of R.

Optional Reading Materials

There are two optional textbooks that you don't need to buy, but which may be very useful as companions to the assigned textbook.

- John Kruschke, *Doing Bayesian Data Analysis,* 2nd ed. (Academic Press, 2015; ISBN 978-0-12-405888-0). This is an excellent introduction to Bayesian data analysis for beginners. It is gentler than *Statistical Rethinking*, and would be better suited for an undergraduate course, but I decided not to use it as the main textbook for this class because it focuses more on the statistical methods and does not give as much application of them to real scientific problems.

  The author writes very clearly and this book may be helpful if you find some of the material in *Statistical Rethinking* confusing. I have asked the Science and Engineering Library to put a copy on reserve so you will be able to access it without buying a copy, and the library also has access to online editions of the book.

- Hadley Wickham and Garrett Grolemund, *R for Data Science* (O'Reilly, 2017; ISBN 978-1-491-91039-9). This book is the best practical introduction I have found for getting started in R and getting things done in data analysis. The author is the chief data scientist at Posit (the makers of RStudio) and wrote a huge number of widely used free packages to extend and enrich R.

  This book uses the `tidyverse` dialect of R, which Hadley Wickham developed, and it follows his philosophy of how to organize data sensibly for analyzing and presenting it.

  This book is freely available online at `https://r4ds.had.co.nz/`, so you won't need to buy it.

Overview of Reading Materials

I have posted detailed reading assignments that give specific pages to read for each class and notes on important things you should understand. **I expect you to complete the reading before you come to class on the day for which the reading is assigned**, so you can participate in discussions of the assigned material and ask questions if there are things you don't understand.

4.2 Graded Work

Basis for Grading

Class participation	5%
Homework	45%
Project	50%

Homework

Unless the assignment specifies differently, homework must be turned in on the day it is due (before midnight at the end of the day).

I will create Brightspace assignments for the homework and you can turn the assignments in there. You can turn in your assignments as Word or PDF documents, or if you write them by hand, you can take photos of the pages and upload those (in JPG or PNG format, please).

There is a PDF version of *Statistical Rethinking* on the web, which many people download. It is substantially similar to the printed version that we are using for our official textbook, but the page numbers are different.

To avoid confusion, I have posted PDF versions of the homework exercises from the printed book, which you can download when you do the homework.

Projects

In the second half of the semester, you will do a research project, in which you will choose a data set that's interesting to you and apply Bayesian methods to analyze it. You will present the results of your project in class during the last week of the semester and turn in a written report about your project.

Tests and Examinations

There will not be any tests or examinations in this course. Your grade will be based on class participation, homework, modeling projects, and in-class presentations.

5 Honor Code:

This course, like all courses at Vanderbilt, is conducted under the Honor Code.

Studying: As you study for this class, I encourage you to to seek help from me or from other classmates or friends.

Homework: I encourage working together. I also encourage you to talk with other classmates, as well as friends and acquaintances outside of class. You may discuss assignments, compare notes on how you are working a problem, and you may look at your classmates' work on homework assignments. But you must work through the problems yourself in the work you turn in: **Even if you have discussed the solution with others you must work through the steps yourself and express the answers in your own words. You may not simply copy someone else's answer.**

Research project: The research project will be conducted under the same ethical principals that apply to publishing papers in scientific journals. The work must be your own, but you may consult any other resources. If anyone else makes a substantial contribution, you must list them and their contributions in an Acknowledgements section.

If you ever have questions about how the Honor Code applies to your work in this course, please ask me. **Uncertainty about the Honor Code does not excuse a violation.**

6 Final Note:

I have made every effort to plan a busy, exciting, and instructive semester. I may find during the term that I need to revise the syllabus to give more time to some subjects or to pass more quickly over others rather than covering them in depth. Thus, while I will attempt to follow this syllabus as closely as I can, you should realize that it is subject to change during the semester.

7 Meet Your Professor

Jonathan Magnolia Gilligan has worked in many areas of science and public policy. Their past research includes work on laser physics, quantum optics, laser surgery, electrical properties of the heart, using modified spy planes to study the ozone layer in the stratosphere, and connections between religion and care for the environment. They are a Professor of Earth & Environmental Sciences, Professor of Civil & Environmental Engineering, and Professor of Climate & Environmental Studies. They serve on the steering committee for the Vanderbilt Center for Sustainability, Energy, and Climate (VSEC), the leadership team for the Center for Climate Leadership and AI-driven Integrity in Mitigation (CLAIM), the technical advisory board of the Southwestern Urban Corridor Integrated Field Laboratory (SW-IFL), which is helping cities in the Southwestern U.S. plan for climate change, and the steering committee for the Community Surface Dynamics Modeling System (CSDMS), a large interdisciplinary modeling center that studies environmental changes to landscapes and coasts.

Professor Gilligan's current research investigates the role of private-sector organizations as well as individual and household behavior in cutting greenhouse gas emissions; environmental justice aspects of access to safe and affordable drinking water in the US; vulnerability and resilience to environmental stress in Bangladesh; and developing new directions for climate policy in the US.

In 2017, Professor Gilligan and Professor Michael Vandenbergh shared the Morrison Prize for the highest-impact paper on sustainability law and policy published in the previous year. Professors Gilligan and Vandenbergh have developed this work into a book, *Beyond Politics: The Private Governance Approach to Climate Change* (Cambridge University Press, 2017), which the Environmental Law Institute named one of the most important books in environmental law and policy of the last 50 years. In 2023, Gilligan received Vanderbilt's Alexander Heard Distinguished Service Professor Award in honor of their distinctive contributions to understanding problems of contemporary society.

In addition to their academic work, Professor Gilligan dabbles in writing for the theater. Their stage adaptation of Nathaniel Hawthorne's *The Scarlet Letter*, co-written with their mother Carol Gilligan, has been staged at The Culture Project in New York City, starring Marisa Tomei, Ron Cephas Jones, and Bobby Cannavale, and was later performed at Prime Stage Theatre, Pittsburgh and in a touring production by The National Players. Prof. Gilligan and Carol Gilligan also wrote the libretto for an opera, *Pearl*, in collaboration composer Amy Scurria, and producer/conductor Sara Jobin, which was performed at Shakespeare & Company in Lenox MA, starring Maureen O'Flynn, John Bellemer, Marnie Breckenridge, John Cheek, and Michael Corvino, and in Shanghai China, starring Li Xin, Wang Yang, John Bellemer, and Lin Shu.

Schedule of Classes (Subject to Change)

IMPORTANT NOTE: This schedule gives a rough indication of the reading for each day. See the detailed daily assignments on the course web site at `https://ees5891.jgilligan.org`.

Date	Topic	Reading
Mon., Jan. 5	Introduction	No reading
Wed., Jan. 7	Rethinking statistics	*McElreath* Ch. 1–2 ("The Golem of Prague" and "Small Worlds and Large Worlds")
Mon., Jan. 12	Sampling	*McElreath* Ch. 3 ("Sampling the Imaginary")
Wed., Jan. 14	R Review	No reading
Mon., Jan. 19	**MARTIN LUTHER KING, JR. DAY**	No class
Wed., Jan. 21	Geocentric Models	*McElreath* Ch. 4 ("Geocentric Models")
Mon., Jan. 26	Many variables (part 1)	*McElreath* Ch. 5 ("The Many Variables and The Superfluous Waffles"), section 5.1 pp. 123–144
Wed., Jan. 28	Many variables (part 2)	*McElreath* Ch. 5 ("The Many Variables and The Superfluous Waffles"), sections 5.2-5.4 pp. 144–158
Mon., Feb. 2	Designing statistical models	*McElreath* Ch. 6 ("The Haunted DAG & The Causal Terror")
Wed., Feb. 4	Regularization	*McElreath* Ch. 7 ("Ulysses' Compass")
Mon., Feb. 9	Interactions	*McElreath* Ch. 8 ("Conditional Manatees")
Wed., Feb. 11	Monte Carlo sampling	*McElreath* Ch. 9 ("Markov Chain Monte Carlo") sections 9.1–9.3
Mon., Feb. 16	Monte Carlo Software	*McElreath* Ch. 9 ("Markov Chain Monte Carlo") sections 9.4–9.5
Wed., Feb. 18	Generalized linear models	*McElreath* Ch. 10 ("Big Entropy and the Generalized Linear Model") sections 10.1–10.2.1

Date	Topic	Reading
Mon., Feb. 23	Generalized linear models	*McElreath* Ch. 10 ("Big Entropy and the Generalized Linear Model") sections 10.2–10.4
Wed., Feb. 25	Discussion of student projects	No reading
Mon., Mar. 2	Discrete statistical models	*McElreath* Ch. 11 ("God Spiked the Integers")
Wed., Mar. 4	Mixture models	*McElreath* Ch. 12 ("Monsters and Mixtures")
Mon., Mar. 9 Wed., Mar. 11	Spring Break	No class
Mon., Mar. 16	Multilevel models	*McElreath* Ch. 13 ("Models with Memory")
Wed., Mar. 18	Multilevel models, part 2	*McElreath* Ch. 13 ("Models with Memory")
Mon., Mar. 23	More multilevel models	*McElreath* Ch. 14 ("Adventures in Covariance")
Wed., Mar. 25	Messy data	*McElreath* Ch. 15 ("Missing Data and Other Opportunities")
Mon., Mar. 30	Review of Bayesian Regression	No reading
Wed., Apr. 1	Discussion of projects	No reading
Mon., Apr. 6	Generalized Linear Madness	*McElreath* Ch. 16 ("Generalized Linear Madness") sections 16.1–16.2
Wed., Apr. 8	Differential Equations	*McElreath* Ch. 16 ("Generalized Linear Madness") sections 16.3—16.5
Mon., Apr. 13	Diagnosing model output	Reading TBA
Wed., Apr. 15	STAN Hamiltonian Monte Carlo sampler	Reading TBA
Mon., Apr. 20	Project presentations	No reading