Statistics 4420
Spring 2026

Course Description: Introduction to Bayesian Data Analysis

The course will introduce data analysis from the Bayesian perspective to undergraduate students. We will cover important concepts in Bayesian probability modeling as well as estimation using both optimization and simulation-based strategies. We will start from first principles, building up both the mathematical framework and R implementation for our models along the way. Key topics covered in the course include hierarchical models, mixture models, hidden Markov models and Markov Chain Monte Carlo.

Prerequisites:

1. A course in probability (Statistics 430 or equivalent)
2. A course in statistical inference (Stat 431 or equivalent)
3. Basic knowledge of linear algebra (e.g. matrix multiplication and inversion, etc.)
4. Experience with the statistical software R (at the level of Stat 405 or Stat 470)

Professor:
Dr. Shane Jensen ARB 415
stjensen@wharton.upenn.edu 215-573-2211

Lectures: Tue Thu 1:45-3:15pm (SHDH 213)

Office Hours: Wed 4:00-5:00pm (ARB 415)

Recommended Textbook:
Bayesian Data Analysis (3nd Edition) by Gelman, et.al.

Required Software:
The R statistical package is needed and can be downloaded at `www.r-project.org`

Course Website:
All course materials will be available on Canvas

Evaluation

Your course grade will be calculated as:
- 50% homeworks
- 25% midterm exam
- 25% final exam

Midterm exam will be a **24 hr take home,** starting at **2pm on Thu March 5**

Final exam will be a **24 hr take home,** starting at **2pm on Wed May 6**

No late homeworks will be accepted, for any reason whatsoever

No make-up midterm will be given

Generative AI Course Policy

The use of Generative AI tools (such as ChatGPT) is permitted, provided you cite its use explicitly and comprehensively. However, this course will focus on building models from first principles while generating both the mathematical framework and R code base for our analyses along the way. So while generative AI tools may be helpful for validation or code checking, it is required that you produce your own **replicable** code base for each course assignment and exam in this course. Also note that all AI tools are prone to errors, including falsifying sources. The university's policy on plagiarism still applies to uncited or improperly cited work, whether from an AI or another human being.

Course Topics

1. Introduction to Bayesian Inference (Ch.1)
2. Simple Parametric Models (Ch. 2, 3)
3. Regression Models from the Bayesian Perspective (Ch. 14,15)
4. Hierarchical and Mixture Models (Ch. 5)
5. Optimization Algorithms for Model Estimation (Ch. 13)
6. Monte Carlo Simulation for Model Estimation (Ch. 10,12,13)
7. Posterior Predictive Model Checking (Ch. 6,7)
8. Bayesian Binary Regression
9. Nonparametric and Semiparametric Bayesian models (Ch. 23)
10. Hidden Markov Models
11. Bayesian Tree Models