Biostatistics 682: Applied Bayesian Inference
Lecture 1: Introduction

Jian Kang

Department of Biostatistics
University of Michigan, Ann Arbor

Instructor

- Jian Kang, PhD, Professor of Biostatistics
- Research Interests: Bayesian methods and theory; Machine Learning; Artificial Intelligence; Imaging statistics; Large-scale complex data analysis.
- Email: jiankang@umich.edu
- Office: M4537, SPH II (1415 Washington Heights)
- Office Hour: Monday 4:30PM – 5:30PM ET (Aug 28 - Dec 8) or by appointment
- Class Time: Monday & Wednesday 1:00PM (Sharp) - 2:20PM ET
- Class Room: 1655 SPH I

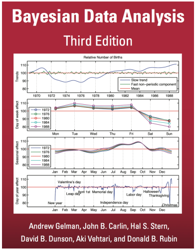
Objectives

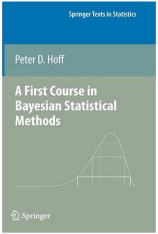
- This course introduces foundational and modern Bayesian inference with a focus on applications in artificial intelligence (AI), biomedical research, and data science.
- Statistical programming language for Bayesian computation
  - R, `https://cran.r-project.org`
  - JAGS (Just Another Gibbs Sampler), `http://mcmc-jags.sourceforge.net`
  - Python, `https://www.python.org`
  - PyMC (a probabilistic programming library for Python), `https://www.pymc.io/welcome.html`

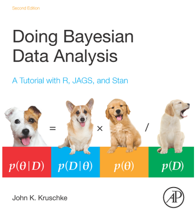
Prerequisites

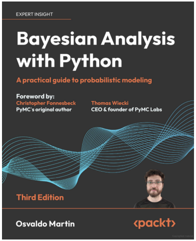
- Linear algebra (matrix theory).
- Multivariate calculus.
- Basic statistics including probability distribution.
- Linear model and hypothesis testing.
- Biostatistics 601 and Biostatistics 650 or equivalent is required prior to or in parallel to taking Biostatistics 682.
- Previous experience in programming in R or Python

Reference Books (Not Required)









- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. *Bayesian data analysis (Third Edition)*, CRC press, 2014
- Hoff, Peter D. *A first course in Bayesian statistical methods.* Springer Science & Business Media, 2009.
- Kruschke, John. *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan.* Academic Press, 2014.
- Martin, O. *Bayesian Analysis with Python: A practical guide to probabilistic modeling*. Packt Publishing Ltd., 2024

Grading (100 points)

- Homework: 50 points
  - Five homework assignments in total
  - Due in two weeks.
  - Plagiarism will NOT be tolerated and late homework will NOT be graded
- Midterm Exams: 25 points
  - In-class, Written exam, Open-book, No Computer
  - Questions for Bayesian theory, methods and computation for data analysis
- Course project: 25 points
  - Develop Bayesian methods for statistical analysis for a given dataset.
  - Write a report to describe the methods and summarize the analysis results
  - Submit the software package or source code for the analysis

Grading Bonus Points (up to 10 points)

- Pop-up quiz in class: 1 point for each question

Final Grades

Letter Grades	Points	Expected Proportions
A+	[95, 110]	$\approx 10\%$
A	[90, 94]	$\geq 25\%$
A−	[85, 90]	$\geq 35\%$
B+	[80, 84]	$\leq 15\%$
B	[70, 79]	$\leq 10\%$
B− and below	[60, 69]	$\leq 5\%$

Gen-AI Use

Generative AI, such as U-M GPT and similar technologies are rapidly becoming part of our professional lives. As such, I expect that you will incorporate these technologies into your work in this class as appropriate and will treat the work you produce as demonstration of your abilities to engage with these new tools. We do ask that you cite the technologies used as part of your submission so that we're all engaging in a dialogue around the role and efficacy of these tools.

Overview of Topics

- Foundations of Bayesian Inference
- Bayesian Computation
- Bayesian Models for Classical AI
- Bayesian Deep Learning
- Bayesian Generative Models
- Applications in AI and Health

Foundations of Bayesian Inference

- Probability as belief: prior, likelihood, posterior
- Conjugate priors and analytical examples
- Bayesian decision theory

Bayesian Computation

- Monte Carlo integration, importance sampling
- MCMC: Gibbs sampling, Metropolis–Hastings
- Variational inference and the ELBO
- Probabilistic programming: JAGS, PyMC, Pyro

Bayesian Models for Classical AI

- Bayesian linear and logistic regressions
- Gaussian processes (regression/classification)
- Naive Bayes, latent Dirichlet allocation
- Mixture models and hidden Markov models
- Posterior predictive checks and model comparison

Bayesian Deep Learning

- Bayesian neural networks
- Stochastic Variational Inference (SVI) for large datasets

Bayesian Generative Models

- Variational autoencoders (VAEs)
- Latent variable models
- Diffusion Models

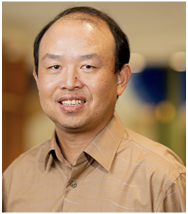
Applications in AI and Health

- Bayesian imaging data analysis
- Bayesian adaptive clinical trials
- Bayesian methods in genomics
- Bayesian analysis of mobile health data

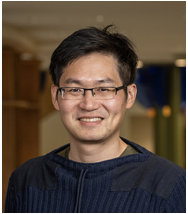
Guest lectures

Fan Bu (Nov 19)

Michele Peruzzi (Nov 24)



William Wen (Dec 1)



Zhenke Wu (Dec 3)

Important Dates

- August 27: **Homework 1 Release**
- September 1: **No class for Labor Day**
- September 15: **Drop / Add deadline for full term classes**
- October 13: **No class for Fall Break**
- October 15: **Midterm Exam**
- November 3: **Homework 5 Release**
- November 17: **Course project release**
- November 19: **Guest Lecture by Fan Bu**
- November 24: **Guest Lecture by Michele Peruzzi**
- November 26: **No class for Thanksgiving**
- December 1: **Guest Lecture by William Wen**
- December 3: **Guest Lecture by Zhenke Wu**
- December 15: **Course project due**

Core Competencies

After taking this class, students are expected to be able to

- Understand basic theory for Bayesian statistics
- Use existing methods or develop new methods for Bayesian data analysis
- Interpret the Bayesian data analysis results to address the scientific questions
- Implement Bayesian computational algorithms using a statistical package
- Perform model checking and model diagnostics

Standards of Academic Act

- From the School of Public Health's Student Code of Conduct,

  "Student academic misconduct includes behavior involving plagiarism, cheating, fabrication, falsification of records or official documents, intentional misuse of equipment or materials, and aiding and abetting the perpetration of such acts. The preparation of reports, papers, and examinations, assigned on an individual basis, must represent each student's own effort. Reference sources should be indicated clearly. The use of assistance from other students or aids of any kind during a written examination, except when the use of books or notes has been approved by an instructor, is a violation of the standard of academic conduct."

- In the context of this course, any work the student hand-in should be his/her own and any material that is a transcript (or interpreted transcript) of work by others must be clearly labeled as such.

Student Well-being

- SPH faculty and staff believe it is important to support the physical and emotional well-being of our students.
- If you have a physical or mental health issue that is affecting your performance or participation in any course, and/or if you need help connecting with University services, please contact the instructor or the Office of Academic Affairs.
- Please visit https://sph.umich.edu/community/student-experience/health-well-being.html for more information.

Student Accommodations

- Students should speak with their instructors before or during the first week of classes regarding any special needs.
  - Students can also visit the Office of Academic Affairs for assistance in coordinating communications around accommodations.
  - Students seeking academic accommodations should register with Services for Students with Disabilities (SSD), which arranges reasonable and appropriate academic accommodations for students with disabilities.
  - Please visit `http://ssd.umich.edu/accommodations` for more information on student accommodations.
- Students who expect to miss classes, examinations, or other assignments as a consequence of their religious observance shall be provided with a reasonable alternative opportunity to complete such academic responsibilities.
  - It is the obligation of students to provide faculty with reasonable notice of the dates of religious holidays on which they will be absent.
  - Please visit `https://www.provost.umich.edu/calendar/religious_holidays21-22.html` for the complete University policy.

Basic Concepts in Bayesian Inference

How to interpret probability?

Long run frequencies

OR

Subjective degree of belief

The Classic Question

Setup: Flip a coin 10 times and observe 7 heads.

Question: What is the probability the coin is biased toward heads?

Two perspectives:

- Bayesian: Update beliefs about the coin after seeing data.
- Frequentist: Test fairness of the coin using repeated sampling ideas.

Bayesian Reasoning

Step 1: Prior belief. Before flipping, assume all biases are equally possible.

Step 2: Update with data. Imagine we had “1 pretend head and 1 pretend tail” as prior information.

$$\text{Prior counts: } (1,1) \quad \rightarrow \quad \text{After 7 heads, 3 tails: } (1+7,\, 1+3) = (8,4).$$

Step 3: Interpret. It’s as if we saw 8 heads and 4 tails total. Most of the belief mass is above 0.5.

$$\Pr(\text{coin favors heads} \mid \text{data}) \approx 89\%.$$

Frequentist Reasoning

Parameters are fixed: the coin is either fair or not.

Hypothesis test:

$$H_0 : p = 0.5 \quad \text{vs} \quad H_A : p > 0.5.$$

If the coin were fair, what is the chance of seeing 7 or more heads out of 10?

$$\text{Pr}(X \geq 7 \mid H_0) = \frac{176}{1024} \approx 17\%.$$

Interpretation: This is not small enough to reject fairness at the 5% level.

Key Contrast

- **Bayesian:** Talks directly about the probability the coin is biased ($\approx 89\%$).
- **Frequentist:** Talks about how unusual the data would be if the coin were fair (p-value ≈ 0.17).

Punchline: Same data, two very different answers!

What is the Bayesian Inference?

- Reallocation of credibility across possibilities

$$\text{Prior Knowledge} \xrightarrow[\quad\quad\quad\quad\quad]{\substack{\text{Data} \\ \text{Information}}} \text{Posterior Inference}$$

- Bayesian versus frequentist
  - Difference in the goal:
    - Bayesian: Quantify and analyze the subjective degree of belief
    - Frequentist: Create procedure that have frequency guarantees
  - Difference in the assumption of parameters
    - Frequentist: Data are random while parameters are fixed but unknown
    - Bayesian: Both data and parameters are random
  - Which school is correct?
    - Can Bayesian inference enjoy the good frequentist properties?
    - Can frequentist approach achieve Bayesian goal?

Bayes' Theorem

Thomas Bayes (1701–1761)

- Bayes's Theorem: Let A and B be two random events with $\Pr(B) > 0$. Then

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)}.$$

  How to prove it?

- Why Bayes Theorem is useful for Bayesian?
- Does a frequentist use Bayes Theorem or not?

An Example – Screening Tests

- Accuracy of a diagnostic tests is usually measured by
  - Sensitivity: probability that a diseased patient tests positive.

$$\Pr(\text{test}+ \mid \text{disease}+)$$

  - Specificity: probability that a healthy person tests negative.

$$\Pr(\text{test}- \mid \text{disease}-)$$

  - Positive predictive value (PPV): probability of a patient has disease given the test is positive.

$$\Pr(\text{disease}+ \mid \text{test}+)$$

  - Negative predictive value (NPV): probability of a person is healthy given the test is negative.

$$\Pr(\text{disease}- \mid \text{test}-)$$

An Example – Screening Tests (Continued)

- The partial accuracy information of a screening test for breast cancer is given

	Diagnostic Test	
Breast Cancer	Positive	Negative
Yes	0.82	0.18
No	0.01	0.99

- We know that the prevalence of breast cancer is 0.1, i.e. $\Pr(\text{disease}+) = 0.1$.
- Can we compute the PPV and NPV based on the Bayes Theorem?

$$
\begin{aligned}
PPV &= \Pr(\text{disease}+ \mid \text{test}+) \\
&= \frac{\Pr(\text{test}+ \mid \text{disease}+)\Pr(\text{disease}+)}{\Pr(\text{test}+)} \\
&= \frac{\Pr(\text{test}+ \mid \text{disease}+)\Pr(\text{disease}+)}{\Pr(\text{test}+ \mid \text{disease}+)\Pr(\text{disease}+) + \Pr(\text{test}+ \mid \text{disease}-)\Pr(\text{disease}-)} \\
&= \frac{0.82(0.1)}{0.82(0.1) + 0.01(0.9)} \\
&= 0.90
\end{aligned}
$$

Let's compute NPV together.

Bayesian data analysis

- Making Bayesian inferences from data
- Probability models are used for all quantities – both observed and unobserved (those we wish to learn something about).
- Three basic steps to an applied Bayesian analysis
  - Set up a full probability model
    - A joint probability distribution for all observable and unobservable quantities in a problem.
    - The model should be consistent with knowledge about the underlying scientific problem and the data collection process
  - Condition on the observed data
    - Involves calculating and interpreting the appropriate posterior distribution
    - The conditional probability distribution of the unobserved quantities of interest given the observed data and perhaps nuisance parameters
  - Evaluate the fit of the model
    - Does the model fit the data?
    - Are substantive conclusions reasonable?
    - How sensitive are results to model assumptions?

Course Survey

https://forms.gle/Gg9YeidBJL7EpXuh9