STAT 238: BAYESIAN STATISTICS

INSTRUCTOR: ADITYA GUNTUBOYINA
COURSE OUTLINE FOR SPRING 2026
UNIVERSITY OF CALIFORNIA, BERKELEY

- **Instructor**: Aditya Guntuboyina. Email: `aditya@stat.berkeley.edu` and Website: `www.stat.berkeley.edu/~aditya`
- **Lectures**: 1 pm to 1:59 pm on Mondays, Wednesdays and Fridays in 9 Lewis Hall.
- **Office Hours**: 2-3 pm on Mondays and Fridays in 443 Evans Hall.
- **GSI**: Reece D Huff. Email: `rdhuff@berkeley.edu`
- **GSI Lab Section**: 9 am - 10:59 am or 12 pm - 1:59 pm in 340 Evans Hall on Tuesdays
- **GSI Office Hours**: 3-5 pm on Tuesdays and 11 am - 12 pm on Wednesdays in 444 Evans Hall

About the course: This course develops Bayesian statistics from first principles of probability, emphasizing interpretation, coherence, and contrasts with frequentist inference, then moves through core parametric estimation and testing problems (including empirical Bayes), linear and generalized linear models, model selection, and high-dimensional regression. It introduces nonparametric regression via Gaussian processes and kernels, with connections to RKHS theory and applications such as Bayesian optimization. A substantial component focuses on Bayesian computation—Monte Carlo methods, MCMC (including Langevin and Hamiltonian dynamics), variational inference, and modern neural-network–enhanced sampling and approximation techniques. The course culminates in advanced applications, including hierarchical and multilevel models, sparse and variable-selection methods, mixture and latent variable models, Bayesian neural networks and trees, probabilistic numerics, optimization, and inverse problems.

Prerequisites: Undergraduate probability at the level of STAT 134, DATA 140, or EE 126 is required. In addition, undergraduate statistics at the level of STAT 133 and STAT 135 is required.

Programming Language: You are free to use any language (e.g., R, Python, Julia, Matlab etc.) for working on your homework. We will be using Python code in class and the lab sections.

Text: There is no required textbook for the class. A variety of references will be provided, and you are welcome to consult any of them.

I will provide materials for each lecture (including slides or typed lecture notes, and code) which will be posted *after* the lecture.

Ed Discussion: I have created a site for this class at Ed Discussion and we will use this platform for Q & A.

Homework assignments: There will be six homework assignments. Your lowest homework score will be dropped when calculating the final grade, and all remaining homework assignments will be weighted equally.

The homework assignments Will be posted on bcourses according to the following schedule. Solutions will need to be uploaded on Gradescope.

- Homework One - will be posted on Jan 30 and due on Feb 13
- Homework Two - will be posted on Feb 13 and due on Feb 27
- Homework Three - will be posted on Feb 27 and due on March 13
- Homework Four - will be posted on March 13 and due on April 03
- Homework Five - will be posted on April 03 and due on April 17
- Homework Six - will be posted on April 17 and due on May 01

You have a total of 120 late hours that you can apply to your homework for the entire semester. No points will be awarded for any homework which brings the total late hours to more than 120.

Final Project: There will be a final project. I will post a list of suggested topics by April 1. Each student will work individually on one topic (students may propose a topic outside the list with prior approval).

Each student is required to record a 20–25 minute presentation explaining their project and findings, and to submit a Jupyter notebook that fully reproduces all numerical analyses. Both the presentation video and the notebook must be uploaded to bCourses by 11:59 PM on May 11.

Assessment: Your final score for the class will be calculated as

$$50\% \text{ Homework} + 50\% \text{ Final Project}.$$

Each homework assignment is worth an equal amount (we will drop the lowest homework score).

Grade Complaints: If you have a complaint against an assigned homework or project grade and want to talk to me about it, first send me a written request through email explaining your case clearly.

Academic Integrity: Collaboration in small groups on homework problems is encouraged, but your write-up must be entirely your own, and you may not read or copy another student's solutions. You may consult books, online resources, and generative AI tools (e.g., ChatGPT or other LLMs) to support your learning; if you do, you must (i) acknowledge all such sources/tools in your write-up and (ii) ensure you fully understand and can independently reproduce and explain every step of your solution. Copying text, code, or solutions—verbatim or paraphrased—from any source (including AI tools) without clear attribution, or submitting work you do not understand, is cheating. Students found to be cheating risk failing the course and being referred to the Office of Student Conduct.

How to cite AI briefly (example): "Consulted ChatGPT for hints on Problem 4 (a)"

Students with disabilities: If you need accommodations for any physical, psychological, or learning disability, please get in touch with me so that we can make the necessary arrangements.