STAT 547: Bayesian Workflow

Charles C. Margossian
University of British Columbia
Winter 2026
https://charlesm93.github.io/stat_547/

DRAFT
Syllabus is subject to change!

Overview. Bayesian inference refers to the use of conditional probability to learn about unknown parameters given data and a probabilistic model. *Bayesian workflow* is the tangled process through which we iteratively build, fit, and criticize many models for the purpose of model development and (exploratory) data analysis. It is in the context of this comprehensive workflow that we can best understand how useful various statistical methods are to the data analyst. In this course, we will study the methods and algorithms that enable scientists to deploy Bayesian workflow within a probabilistic software, such as Stan, PyMC, Pyro, Turing and TensorFlow Probability.

This is a graduate level course, primarily intended for MSc and PhD students interested in doing academic research on this topic. In addition to developing a foundation in Bayesian workflow, we will discuss shortcomings in existing methods and identify open research questions. The course will be grounded in examples from scientific applications.

Prerequisites. The prerequisites are:

- basic probability and statistics,
- applied mathematics: multivariate calculus, linear algebra, basic differential equations,
- coding experience in a high-level language, such as R and Python.

It will also be helpful for students to be familiar with Bayesian statistics, Markov chain Monte Carlo and the software Stan, although we will review these topics during the class.

Requirements and Grades. The requirements are weekly reader reports, one homework assignment [**I might break down the assignment into multiple small asignments**], and a final project.

- *Reader Reports.* Each week, write your thought about the reading. The reports can be up to one page. They are required but not graded and must be handed in class. The reports provide an opportunity to organize your thoughts and practice your scientific writing.
- *Homework.* There will be one homework assignment. This assignment is somewhat long and you should work on it throughout the course, as you learn new material.
- *Final project.* For the project, you will apply the concepts learned in class to an open-ended question. The project can be the application of a probabilistic model to analyze data, the analysis of a method or a theoretical result. You are welcomed to use the final project to advance your own research. The output of the project can be a short paper or a case-study. You will be graded on both the content and the writing quality.

Your grade decomposes as follows:

- Final Project: 50%
- Homework: 30%
- Reader Reports: 10% (the reports are not graded, you just have to hand them in.)
- Class Discussion: 10%

Syllabus.

- *Week 1*: Applied Bayesian analysis and introduction to Bayesian Workflow; Stan; succeptible-infected-recovered (SIR) example; review of MCMC.
- *Week 2*: posterior predictive checks; model comparison for SIR; predictions with calibrated uncertainty; Pareto-smoothed importance sampling.
- *Week 3*: Hamiltonian Monte Carlo; No-U-Turn Sampler. *Project Proposal due.*
- *Week 4*: Hierarchical models; diagnostics for MCMC; performance metrics; the many-short-chains regime. *Homework Assignment due.*
- *Week 5*: Variational inference; basic theory; practical implementations; pathfinder.
- *Week 6*: Open discussion. *Final Project due.*

Textbooks and references.

- *Bayesian Workflow*. Gelman, Vehtari, McElreath et al.
- *Stan User's Guide*. The Stan Development Team.
- *Getting Started with Bayesian Statistics using Stan and Python*. Bob Carpenter.

Reading. Each week, you will have several optional readings. You must read one of the proposed texts and write your reader report on it. You are welcomed to read more if you chose to.

Week 2.

- *Bayesian Workflow*, chapters 1 and 2
- Margossian and Gelman. "For how Many Iterations should we run MCMC?", *Handbook of MCMC*, second edition.
- Vehtari, Gelman, and Gabry. "Practical Bayesian Model Evaluation using Leave-One-Out Cross-Validation and WAIC", *Statistics and Computing*

Week 3.

- Neal. "MCMC using Hamiltonian dynamics", *Handbook of MCMC*

- Betancourt. "A Conceptual Introduction to Hamiltonian Monte Carlo"
- Hoffman and Gelman. "Adaptively Setting Path Lengths in Hamiltonian Monte Carlo"
- Hoffman, Radul and Sountsov. "An Adaptive MCMC Scheme for Setting Trajectory Lengths in Hamiltonian Monte Carlo." *AISTATS*

Week 4.

- Gelman, Vehtari, McElreath et al. *Bayesian Workflow*, chapter 11
- Betancourt. "A Short Review of Ergodicity and Convergence of Markov chain Monte Carlo Estimators"
- Vehtari et al. "Rank-normalization, folding, and localization: an improved $\widehat{R}$ for assessing convergence of MCMC" *Bayesian Analysis*
- Margossian et al. "Nested $\widehat{R}$: Assessing the Convergence of MCMC when running many short chains" *Bayesian Analysis*
- Biwas, Jacob and Vaneti. "Estimating Convergence of Markov chain Monte Carlo with L-lag Coupling"

Week 5.

- Blei, Kucukelbir and McAuliffe. "Variational Inference: A Review for Statisticians." *Journal of the American Statistical Association*.
- Yao et al. "Yes but did it work? Evaluating Variational Inference". *ICML*
- Huggins et al. "Validated variational inference via practical posterior error bounds" *AISTATS*
- Zhang et al. "Pathfinder: Parallel quasi-Newton variational inference" *Journal of Machine Learning Research*

Week 6.

- Gelman, Vehtari, McElreath et al. *Bayesian Workflow*, any one chapter in Part IV
- Gelman and Shalizi. "Philosophy and the Practice of Bayesian Analysis."