Bayesian Inference and Computation

Dr Mengchu Li and Dr Lukas Trottner

Semester 2, 2025

Contents

Practicalities **5**
0.1 Module Aims . . . 5
0.2 Module Structure . . . 6
0.3 Assessment . . . 6
0.4 Recommended Books and Videos . . . 6
0.5 Common Distributions . . . 6

1 Fundamentals of Bayesian Inference **9**
1.1 Statistical Inference . . . 9
1.2 Frequentist Theory . . . 9
1.3 Bayesian Paradigm . . . 9
1.4 Probability Basics and Exchangability . . . 9
1.5 Bayes' Theorem . . . 9

2 Programming in R **11**
2.1 Random Numbers, For Loops and R . . . 11
2.2 Functions in R . . . 11
2.3 Good Coding Practices . . . 11

3 Bayesian Inference **13**
3.1 The Binomial Distribution . . . 13
3.2 Reporting Conclusions from Bayesian Inference . . . 13
3.3 The Exponential Distribution . . . 13
3.4 The Normal Distribtuion . . . 13
3.5 Hierarchical Models . . . 13
3.6 Prediction . . . 13
3.7 Non-informative Prior Distibrutions . . . 13
3.8 Bernstein-von-Mises Theorem . . . 13
3.9 Lab . . . 13

4 Sampling **15**
4.1 Uniform Random Numbers . . . 15
4.2 Inverse Transform Sampling . . . 15

4.3 Rejection Sampling 15
4.4 Ziggurat Sampling 15
4.5 Approximate Bayesian Computation 15
4.6 Lab 15

5 Markov Chain Monte Carlo **17**
5.1 Properties of Markov Chains 17
5.2 Metropolis–Hastings 17
5.3 Gibbs Sampler 17
5.4 Metropolis-within-Gibbs 17
5.5 MCMC Diagnostics 17
5.6 Beyond MCMC 17
5.7 Lab 17

6 Markov Chain Monte Carlo **19**
6.1 Properties of Markov Chains 19
6.2 Metropolis–Hastings 19
6.3 Gibbs Sampler 19
6.4 Metropolis-within-Gibbs 19
6.5 MCMC Diagnostics 19
6.6 Beyond MCMC 19
6.7 Lab 19

7 Advanced Computation **21**
7.1 Gaussian Processes 21
7.2 Data Augmentation 21
7.3 Prior Ellicitation 21
7.4 Lab 21

Practicalities

0.1 Module Aims

Bayesian inference is a set of methods where the probability of an event occurring can be updated as more information becomes available. It is fundamentally different from frequentist methods, which are based on long running relative frequencies. This module gives an introduction to the Bayesian approach to statistical analysis and the theory that underpins it.

Students will be able to explain the distinctive features of Bayesian methodology, understand and appreciate the role of prior distributions and compute posterior distributions. It will cover the derivation of posterior distributions, the construction of prior distributions, and inference for missing data. Extensions are considered to models with more than a single parameter and how these can be used to analyse data. Computational methods have greatly advanced the use of Bayesian methods and this module covers, and allows students to apply, procedures for the sampling and analysis of intractable Bayesian problems.

By the end of the course, students should be able to:

1. Demonstrate a full and rigorous understanding of all definitions associated with Bayesian inference and understand the differences between the Bayesian and frequentist approaches to inference
2. Demonstrate a sound understanding of the fundamental concepts of Bayesian inference and computational sampling methods
3. Understand how to make inferences assuming various population distributions while taking into account expert opinion and the implications of weak prior knowledge and large samples
4. Demonstrate an understanding of the principles of Markov Chain Monte Carlo and be able to programme an MCMC algorithm
5. Engage in Bayesian data analysis in diverse situations drawn from physics, biological, engineering and other mathematical contexts.

0.2 Module Structure

The module is split between theory and computation. Each week will have three lectures, one computer lab and one guided study. In the labs, you will need to bring your own laptop.

0.3 Assessment

Assessment for this module is 50% via an exam and 50% via coursework assignments during the semester. The exam will last 1h 30m and take place during the summer exam period. There will be three coursework assignment – assignment 1 will be worth 10% of the final mark, with assignments 2 and 3 counting for 20% each. More details about the assignments will be made available during the semester.

0.4 Recommended Books and Videos

No books are required for this course and the whole material is contained in these notes. However, you may find it useful to use other resources in your studies. I recommend the following:

1. A First Course in Bayesian Statistical Methods - Peter D. Hoff. This is a short book that covers the basics of Bayesian inference and computation. To the point and well written, it's a useful place to look topics up.

2. Bayesian Data Analysis - Andrew Gelman, John Carlin, Hal Stern, David Dunson, Aki Vehtari, and Donald Rubin. This is a thorough book explaining everything you'd need to know to carry out Bayesian data analysis. It's a fairly long and in-depth book, but the authors are authoritative and give good advice throughout. Example code on the website is in R, Python and Stan.

3. Statistical Rethinking - Richard McElrath. This book provides a friendly intuitive understanding of Bayesian inference and computation. Aimed at social and natural scientists, it has less theory that the other two books but is perhaps more approachable. A set of video lectures for this book can be found on YouTube.

0.5 Common Distributions

For many Bayesian inference problems, it is useful to be able to identify probability density functions (for continuous random variables) and probability mass functions (for discrete random variables) up to proportionality. Some common density/mass functions are given below.

Normal distribution $N(\mu, \sigma^2)$

$$\pi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\} \qquad x \in \mathbb{R},$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$.

Beta distribution $\text{Beta}(\alpha, \beta)$

$$\pi(x \mid \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1} \qquad x \in [0, 1],$$

where $\alpha, \beta > 0$ and $B(\alpha, \beta)$ is the Beta function.

Gamma distribution $\text{Gamma}(\alpha, \beta)$

$$\pi(x \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \qquad x > 0,$$

where $\alpha, \beta > 0$ and $\Gamma(\alpha)$ is the Gamma function.

Exponential distribution $\text{Exp}(\lambda)$

$$f(x \mid \lambda) = \lambda e^{-\lambda x} \qquad x > 0,$$

where $\lambda > 0$.

Poisson distribution $\text{Poi}(\lambda)$

$$\pi(x = k \mid \lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad k \in \{1, 2, ...\},$$

where $\lambda > 0$.

Binomial distribution $\text{Bin}(N, p)$

$$\pi(x = k \mid N, p) = \binom{N}{k} p^k (1-p)^{N-k} \qquad k \in \{1, \dots, N\}$$

where $p \in [0, 1]$.

Chapter 1

Fundamentals of Bayesian Inference

Placeholder

1.1 Statistical Inference

1.2 Frequentist Theory

1.3 Bayesian Paradigm

1.4 Probability Basics and Exchangability

1.5 Bayes' Theorem

Chapter 2

Programming in R

Placeholder

2.1 Random Numbers, For Loops and R

2.2 Functions in R

2.2.1 Built in commands

2.2.2 User defined functions

2.3 Good Coding Practices

2.3.1 Code Style

Chapter 3

Bayesian Inference

Placeholder

3.1 The Binomial Distribution

3.2 Reporting Conclusions from Bayesian Inference

3.3 The Exponential Distribution

3.4 The Normal Distribtuion

3.5 Hierarchical Models

3.6 Prediction

3.7 Non-informative Prior Distibrutions

3.8 Bernstein-von-Mises Theorem

3.9 Lab

Chapter 4

Sampling

Placeholder

4.1 Uniform Random Numbers

4.2 Inverse Transform Sampling

4.3 Rejection Sampling

4.3.1 Rejection Sampling Efficiency

4.4 Ziggurat Sampling

4.5 Approximate Bayesian Computation

4.5.1 ABC with Rejection

4.5.2 Summary ABC with Rejection

4.6 Lab

Chapter 5

Markov Chain Monte Carlo

Placeholder

5.1 Properties of Markov Chains

5.2 Metropolis–Hastings

5.3 Gibbs Sampler

5.4 Metropolis-within-Gibbs

5.5 MCMC Diagnostics

5.6 Beyond MCMC

5.7 Lab

Chapter 6

Markov Chain Monte Carlo

Placeholder

6.1 Properties of Markov Chains

6.2 Metropolis–Hastings

6.3 Gibbs Sampler

6.4 Metropolis-within-Gibbs

6.5 MCMC Diagnostics

6.6 Beyond MCMC

6.7 Lab

Chapter 7

Advanced Computation

Placeholder

7.1 Gaussian Processes

7.1.1 Covariance Functions

7.1.2 Gaussian Process Regression

7.2 Data Augmentation

7.2.1 Imputing censored observations

7.2.2 Imputing Latent Variables

7.2.3 Grouped Data

7.3 Prior Ellicitation

7.3.1 Prior Summaries

7.3.2 Betting with Histograms

7.3.3 Prior Intervals

7.4 Lab

7.4.1 Gaussian Processes

7.4.2 Missing Data

Bibliography

Hairer, M. and Mattingly, J. C. (2011). Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel.

Meyn, S. and Tweedie, R. L. (2009). *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition. With a prologue by Peter W. Glynn.