Bayesian Learning

Contents

The course gives a solid introduction to the Bayesian statistical inference, with special emphasis on models and methods in computational statistics and machine learning.

- We will get off to a shocking start by introducing a very different probability concept than the one you are probably (😉) used to: subjective probability.
- We will then move on to the mathematics of the prior-to-posterior updating in basic statistical models, such as the Bernoulli, normal and multinomial models.
- Bayesian prediction and decision making under uncertainty is carefully explained, and you will hopefully see why Bayesian methods are so useful in modern application where so much focuses on prediction and decision making.
- The really interesting stuff starts to happen when we study regression and classification. You will learn how prior distributions can be seen as regularization that allows us to use flexible models with many parameters without overfitting the data and improving predictive performance.
- A new world will open up when we learn how complex models can be analyzed with simulation methods like Markov Chain Monte Carlo (MCMC), Hamiltonian Monte Carlo (HMC) and approximate optimization methods like Variational Inference (VI).
- You will also get a taste for probabilistic programming languages for Bayesian learning, in particular the popular Stan language in R.
- Finally, we'll round off with an introduction to Bayesian model selection and Bayesian variable selection.

Course literature and Schedule

The course will use the following book as the main course literature:

- Gelman, Carlin, Stern, Dunson, Vehtari, Rubin (2014). Bayesian Data Analysis. Chapman & Hall/CRC: Boca Raton, Florida. 3rd edition.
- Some chapters from a book that I am writing for a courses of this kind.
- Additional course material linked from this page, such as articles and tutorial.

The course schedule on TimeEdit is here: Schedule.

Computer labs

- The three computer labs are central to the course. Expect to allocate substantial time for each lab. Many of the exam questions will be computer based, so working on the labs will also help you with the exam.
- You are strongly encouraged to do the labs in R, but any programming language is ok to use.
- The labs should be done in **pairs** of students.
- Each lab report should be submitted as a PDF along with the .R file with code. Submission is done through Athena.
- There is only two hours of supervised time allocated to each lab. The idea is that you:
    - should start working on the lab before the computer session
    - so that you are in a position to ask questions at the session
    - and then finish up the report after the lab session.
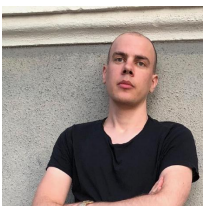
Teachers

Mattias Villani, Lecturer

Professor, Natural Born Bayesian 😜
Department of Statistics, Stockholm University
Division of Statistics and Machine Learning, Linköping University

Oscar Oelrich, Math exercises and Computer labs
PhD Candidate, specializing in Bayesian Statistics
Department of Statistics, Stockholm University

Lectures

Lecturer: Mattias Villani

Lecture 1 - Subjective probability. Likelihood. Bayesian updating. The Bernoulli model
Reading: BDA Ch. 1, 2.1-2.4 | Slides
Code: Beta density | Bernoulli model

Lecture 2 - Gaussian model. The Poisson model. Conjugate priors. Prior elicitation

Reading: BDA Ch. 2.-2.9 | Slides

Code: One-parameter Gaussian model

Python Notebook: Poisson eBay data in Notebook and HTML

Lecture 3 - Multi-parameter models. Marginalization. Multinomial model. Multivariate normal model

Reading: BDA Ch. 3 | Slides

Code: Two-parameter Gaussian model

R Notebook: Multinomial survey data in Notebook and HTML

Lecture 4 - Monte Carlo simulation. Prediction. Making Decisions

Reading: BDA Ch. 9.1-9.2. | Slides

Code: Prediction with two-parameter Gaussian model

Lecture 5 - Regression. Regularization priors

Reading: BDA Ch. 14 and Ch. 20.1-20.2 | Slides

Lecture 6 - Classification. Posterior approximation.

Reading: BDA Ch. 16.1-16.3 | Slides

Code: Logistic and Probit Regression

Lecture 7 - Gibbs sampling and Data augmentation

Reading: BDA Ch. 10-11 | Slides

Code: Gibbs sampling for a bivariate normal | Gibbs sampling for a mixture of normals

Extra material: Illustration of Gibbs sampling when parameters are strongly correlated.

Lecture 8 - MCMC and Metropolis-Hastings

Reading: BDA Ch. 11 | Slides

Code: Simulating Markov Chains

Lecture 9 - HMC and Variational inference

Reading: BDA Ch. 12.4 and Ch. 13.7 |

Lecture 10 - Probabilistic programming for Bayesian learning

Reading: BDA Ch. X | Slides RStan vignette

Code: RStan - Three Plants | RStan - Bernoulli model | RStan - Logistic regression | RStan - Logistic regression with random effects | RStan - Poisson model

Lecture 11 - Bayesian model comparison

Reading: BDA Ch. 7 | Slides

Code: Comparing models for count data

Lecture 12 - Bayesian variable selection and model averaging

Reading: Article on Bayesian variable selection | Slides

Mathematical exercises

Assistant: Oscar Oelrich

Exercise session 1: Problem set

Exercise session 2: Problem set

Exercise session 3: Problem set

Computer labs

Assistant: Oscar Oelrich

Computer Lab 1 - Exploring posterior distributions in one-parameter models by simulation and direct numerical evaluation.
Lab: Lab 1
Submission: Athena.

Computer Lab 2 - Polynomial regression and classification with logistic regression.
Lab: Lab 2 | Temp in Linköping dataset | WomenWork dataset
Submission: Athena.

Computer Lab 3 - Gibbs sampling, Metropolis-Hastings and Stan.
Lab: Lab 3 | Rainfall dataset | eBay dataset | How to code up Random Walk Metropolis | campylobacter dataset
Submission: Athena.

Examination

The course examination consists of:

- Written lab reports (deadlines given in Athena)
- Take home exam (first exam on Jan 14-15)

R

- The main page with links to downloads for the programming language R
- RStudio - a very nice developing environment for R.
- Short introduction to R | A little longer introduction | John Cook's intro to R for programmers.