Syllabus

François Leblanc (francois.leblanc@tbs-sct.gc.ca)
Chris Lavoie (chris.lavoie@canada.ca)

16-05-2019

Course Info

- Time: Friday 11 - 12AM
- Website: bayes.leblancfg.com
- Join us online by WebEx: TODO
- Or join us in person: Boardroom TBD

Required Text

> **Doing Bayesian Data Analysis, 2nd Edition: A tutorial with R, JAGS, and Stan**, by John K. Kruschke (2015), published by Academic Press / Elsevier, ISBN-13: 978-0124058880.

The author has a 30% discount when the book is purchased through Academic Press / Elsevier. Enter code `MATH318` at checkout.

Prerequisites

- **Familiarity** with R, and **working knowledge** of R or Python. Your relative course load will likely depend on this.
- Knowledge of basic calculus.

Course Description

"Classical" statistics, encapsulating well-known methods such as t-tests, ANOVA, etc. are from the frequentist school of statistical thought. The basic idea of frequentist statistics is that the world is described by parameters that are fixed and unknown. Since these parameters are unknown, we do not know their exact values. Since they are fixed, however, we cannot discuss them in probabilistic terms.

It turns out that we can use probabilities not only to express the chance that something will occur, but we can also use them to express the extent to which we believe something. This is the Bayesian approach.

Bayesian statistics are rippling through everything from physics to cancer research, ecology to psychology, law to politics, even sports analytics. Enthusiasts say they are allowing scientists to solve problems that would have been considered impossible just 20 years ago. It is proving especially useful in approaching complex problems, such as in the search for the crashed Air France Flight 447 in 2011.

Data gathering is frequently expensive compared with data analysis. It is sensible then that hard-won data be inspected from many different viewpoints. In the selection of viewpoints, Bayesian methods allow greater emphasis to be given to scientific interest and less to mathematical convenience. This course is designed to provide an introduction to fundamental conceptual, computational, and practical methods of Bayesian data analysis.

The course will provide a basic introduction to Bayesian concepts and methods with an emphasis on the data analysis. We will discuss model choice, including the assessment of prior distributions. We will discuss how to conduct inference in a Bayesian setting, through posterior means, credible intervals and hypothesis testing.

The analyses will be performed using the freely available software Python and R, and we will spend some time during our first meeting to set up our environments. JAGS and STAN are the sampling libraries used in the textbook we'll be following, but it has already been "translated" into other implementations: BUGS, brms, and PyMC3.

Tentative Schedule

1. May 31st 2019 - Introduction, Probability, and Setting Up
    - Chapters 2-4
    - Assignment: Setting up environment, Ex. 2.1, 2.2, 4.3, 4.6
2. June 7th 2019 - Bayes' rule, Inferences with binomials
    - Chapters 5-6
    - Assignment: Ex. 5.1 - 5.4, 6.1, 6.2, 6.4, 6.5
3. June 14th 2019 - Sampling methods: Markov Chain Monte Carlo (MCMC) and JAGS
    - Chapters 7-8
    - Assignment: *TBD*
4. June 21st 2019 - Hierarchical models
    - Chapter 9
    - Assignment: *TBD*
5. June 28th 2019 - Model checking and comparison, Reporting a Bayesian Analysis
    - Chapters 10, 25.1
    - Assignment: *TBD*
6. July 5th 2019 - Null Hypothesis Testing, Goals, Power & Sample Size
    - Chapters 11 - 13

- Assignment: *TBD*

7. July 12th 2019 - Linear Regression
   - Chapter 16
   - Assignment: *TBD*
8. June 21st 2019 - Generalized linear models
   - Chapter 15
   - Assignment: *TBD*
9. July 19th 2019 - Logistic Regression
   - Chapters 21 - 22
   - Assignment: *TBD*
10. July 26th 2019 - Special Topic - Principled Bayesian Workflow by Dr Keith O'Rourke (Health Canada)