Bayesian Statistics Modelling

Chendi Wang (Vrije Universiteit Amsterdam)

September 12th – 14th

Teacher e-mail:
chendi.wang@vu.nl

Language: English
Schedule: Afternoon, from 15 to 19 hours.

Module overview

This course aims at introducing the theoretical and applied principles of Bayesian statistics specifically geared toward students in political science. Despite the usefulness of the Bayesian approach, it is less taught among students of political science compared with the Frequentist paradigm. The Bayesian framework is particularly useful for the type of data that political scientists encounter. It is a formal method for combining prior information with observed quantitative information; it offers a more general way to deal with issues of model identification, and it allows researchers to fit very realistic, sometimes complicated, models.

We will first discuss the differences between Bayesian and Frequentist approaches and the advantages of using the Bayesian approach for social scientists. Second, the course will cover the theoretical foundation of Bayesian statistics and introduce stochastic simulation methods for inference (Markov chain Monte Carlo).

Third, we will mainly focus on using Bayesian models in political science data analysis. Students will learn how to do standard statistical models from a Bayesian perspective. The topics include linear, logit/probit, poisson/negative binomial, hierarchical models, etc. Additional topics, such as measurement models and Bayesian model averaging will be introduced in later sessions.

The goal of the course is to teach students to understand and apply various Bayesian methods for answering research questions in quantitative political science. The students will also gain proficiency in data analytic skills by using the open source statistical programming language R, in addition to the theoretical material.

This course is for students in various social science disciplines. It is an advanced quantitative methods courses designed for students who have already obtained basic quantitative training.

Pre-requisites

Students who intend to enrol in this course are required to have basic knowledge in statistical analysis including linear regression models and hypothesis testing and have a working understanding of matrix algebra and some exposure to models with limited dependent variables (e.g. binary). Successful completion of Introduction to Quantitative Methods will suffice.

Computing requirement

This workshop requires researchers to mostly use JAGS or BUGS. If time permits, a gentle introduction to Stan will also be provided. We access JAGS and OpenBUGS through R, and therefore basic knowledge of R language is required. Most models in the lecture will be built on JAGS or OpenBUGS. The languages of these two programs are almost identical. JAGS is recommended, as it runs on Macs, Windows as well as Linux. Instructions about how to install JAGS will be distributed online prior to the workshop.

Objectives

After taking this course, students are expected to:

- Understand the fundamental differences and similarities between frequentist and Bayesian approaches to inference.
- Formulate linear and generalised linear models in the Bayesian framework using JAGS or Stan.
- Estimate and interpret linear and generalised linear models in the Bayesian framework.
- Formulate and estimate Bayesian hierarchical models.
- Compare and asses Bayesian models.
- Apply the Bayesian methods to political science research questions.

Recommended books

The following two books are recommended for your own reference. A long reading list of useful materials on Bayesian modelling will also be distributed prior to the workshop.

Gelman, A. & Hill, J. (2007). Data analysis using regression and multilevel/hierarchical models. New York, NY: Cambridge University Press.

Gill, J. (2014). Bayesian Methods A Social and Behavioral Sciences Approach. (3rd ed.). BocaRaton, FL: Chapman and Hall/CRC.

For researchers who are not familiar with R, the following:

Fox, J. & Weisberg, S. (2011). An R Companion to Applied Regression. (2nd ed.). Thousand Oaks: Sage.

Tutorials

Tutorials wuth guided hands-on coding Will also be offered during the course after each topic.

Class schedule

The sessions will be held in the afternoon, from 15 pm to 19 pm.

Session 1 (September 12). Introduction:

- Probability.
- Bayes´Rule.
- Priors and Posteriors.
- MCMC Algorithms.
- Convergence Diagnostics.

- JAGS/BUGS Language.
- The Linear Model.

Session 2 (September 13):

- Models for binary Outcomes.
- Models for Discrete Choice Outcomes.
- Models for Count Outcomes.

Session 3 (September 14):

- Measurement Models.
- Hierarchical/Multilevel Models.
- Advanced topics: Bayes Factors and Model Comparison (time permitting).