Bayesian Data Analysis with WinBugs

Matthew S. Johnson [1]
IMPS 2005
Tilburg University
Tilburg, The Netherlands
July 4, 2005

Bayesian modeling and inference differs from classical inference in that it assumes a joint probability model for both the observed data and the unknown parameters. Bayesian inference is then made by examining the conditional distribution of the model parameters given the observed data. Therefore we can make intuitive probabilistic statements about the unknown parameters like, “The posterior probability that the parameter is greater than zero is 0.95.”

Because the Bayesian model assumes a joint distribution for the data and the parameters, Bayesian methods are flexible enough to handle both simple and complex problems. The *prior* distribution placed on the model parameters can be utilized to incorporate expert knowledge about the parameters into the model. Simple prior distributions can be used to limit the range of values a parameter can take. Multilevel prior specifications can be used to define hierarchical models. The flexibility of the Bayesian method makes it particularly useful for the analysis of psychometric data.

One of the difficulties of Bayesian data analysis was that for all but the simplest problems it was impossible to determine the posterior distribution of the model parameters analytically. For slightly more complex models some posterior quantities could be approximated, but still the list of models for which these approximations worked well was rather small. The development and improvement of Monte Carlo techniques has recently made the posterior distributions of very complicated Bayesian models easy to approximate. The WinBugs software provides users with a simple tool to perform these Markov chain Monte Carlo simulations.

This workshop, outlined in greater detail below, will: (a) introduce the concepts behind a Bayesian data analysis; (b) discuss Markov chain Monte Carlo (MCMC) algorithms, including how to monitor convergence and how to perform Bayesian inference using the output; (c) teach participants how to use the WinBugs software; and (d) perform hands-on Bayesian data analyses of selected psychometric problems.

The tentative outline for the workshop is summarized below. Interested individuals are invited to contact me by email if there is a topic of interest not mentioned below that they would like to have covered during the workshop.

[1] Matthew Johnson is Assistant Professor of Statistics at Baruch College of the City University of New York. He can be contacted by email at `Matthew_Johnson@baruch.cuny.edu`

09.00-10.30. *Bayesian Methodology–Introduction/Review*.

This component of the workshop will introduce participants to the basics concepts required for Bayesian data analysis. Each of the topics below will be illustrated using the normal and/or binomial distributions.

1. The statistical (data) model and the likelihood
2. Prior distributions
    (a) Conjugate priors
    (b) Informative priors
    (c) Non-informative priors including Jeffreys' prior
    (d) Hyper-priors and Hierarchical Models
3. Directed Acyclic-Graphs (DAGs)
4. The posterior distribution
5. Bayesian Inference
    (a) Bayes Factors
    (b) Confidence Regions

10.30-10.50. Coffee Break

10.50-12.20. *Approximating posterior distributions using Markov chain Monte Carlo Algorithms.*

In most applications it is impossible to determine the posterior distribution analytically. Markov chain Monte Carlo (MCMC) algorithms allow us to approximate the posterior distribution using samples from a Markov chain.

This component of the workshop will introduce participants to the basics of MCMC and WinBugs. The topics will be illustrated using the same simple examples from component I as well as a simple normal measurement error model.

1. Overview of MCMC methods
2. The most common algorithms–Gibbs sampling, Metropolis-Hastings, Adaptive and non-adaptive rejection sampling.
3. Using WinBugs
    (a) Syntax for defining Bayesian models in WinBugs
    (b) Using the WinBugs *Doodle* utility to define the model with a DAG.
4. Output Analysis
    (a) Monitoring convergence–Burn-in, Monte Carlo errors, and the length of the chain.
    (b) Point estimates–Posterior mean and median
    (c) Posterior variance and standard deviation
    (d) Approximating equal-tailed credible intervals

12.20-13.40 Lunch

13.40-15.10 *Psychometric examples using WinBugs.*

In this third component of the workshop I will work with the participants to develop Bayesian models for a number of popular models from psychometrics, and then use WinBugs to approximate the posterior distributions of the model parameters.

Below is a tentative list of models for discussion. If there is a model that is not listed below that you would like to discuss please do not hesitate to contact me at Matthew_Johnson@baruch.cuny.edu.

1. Models for paired comparisons.
2. IRT models.
   (a) 1-, 2-, and 3-PL and normal ogive models
   (b) Models for polytomous data–Partial credit model, Graded response model, etc.
3. Unfolding models.
4. Latent class models and mixture models.
5. Factor analysis models–exploratory and confirmatory.

15.10-15.30 Coffee

15.30-16.30 *Continuing the psychometric examples and further topics.*

After the discussion and implementation of Bayesian psychometric models in WinBugs we will discuss some advanced MCMC topics. Possible topics for discussion include:

1. Approximating Bayes factors with MCMC output.
2. Goodness of fit and model checking with posterior predictive checks.

16.30-18.00 Reception