Advanced Data Analysis

Introduction to Bayesian Inference via Markov Chain Monte Carlo

Matt Wand

Square Bracket Notation

$[x]$ is density of x

$[y|x]$ is density of y given x

Inverse Gamma (IG) Notation

$$
\begin{aligned}
& [x] \sim \text{IG}(A, B) \\
\iff \quad & [1/x] \sim \text{Gamma}(A, 1/B) \\
\iff \quad & [x] = \frac{B^A}{\Gamma(A)} x^{-A-1} \exp(-B/x), \quad x > 0.
\end{aligned}
$$

Bayesian Inference

Data: y
Parameter: θ

Bayesian inference is based on the posterior density

$$[\theta \mid y]$$

Bayesian Inference for Normal Data

Consider the Bayesian version of
inference for a univariate normal sample:

$$y_1, \ldots, y_n \overset{\text{ind.}}{\sim} N(\mu, \sigma^2).$$

A Bayesian model is:

$$[y|\mu, \sigma^2] \sim N(\mu 1, \sigma^2 I)$$
$$[\mu] \sim N(0, \sigma_\mu^2)$$
$$[\sigma^2] \sim \text{IG}(A, B)$$

Posterior Distributions

$$[\mu|y] = \frac{e^{-\mu^2/(2\sigma_\mu^2)}\Gamma(A + \frac{n}{2})}{\sqrt{2\pi}\sigma_\mu(B + \frac{1}{2}\|y - \mu 1\|^2)^{A+\frac{n}{2}} \int_0^\infty (\sigma^2 + \sigma_\mu^2)^{-n/2}(\sigma^2)^{-A-1} e^{-\frac{\|y\|^2}{2(\sigma^2+\sigma_\mu^2)} - \frac{B}{\sigma^2}} d\sigma^2}$$

$$[\sigma^2|y] = \frac{(\sigma^2 + \sigma_\mu^2)^{-n/2}(\sigma^2)^{-A-1} e^{-\frac{\|y\|^2}{2(\sigma^2+\sigma_\mu^2)} - \frac{B}{\sigma^2}}}{\int_0^\infty (\sigma^2 + \sigma_\mu^2)^{-n/2}(\sigma^2)^{-A-1} e^{-\frac{\|y\|^2}{2(\sigma^2+\sigma_\mu^2)} - \frac{B}{\sigma^2}} d\sigma^2}$$

($\|v\| = \sqrt{v^T v}$ is norm of the vector v).

Even in the
univariate normal data setting

the posterior distributions involve
intractable integrals!!!

How About the Full Conditionals?

The full conditionals in the univariate normal example are:

$$[\mu|\sigma^2, y] \quad \text{and} \quad [\sigma^2|\mu, y]$$

$$[\mu|\sigma^2, y] \sim N\left(\frac{\overline{y}}{1 + \sigma^2/(n\sigma_\mu^2)}, \frac{\sigma^2}{n + \sigma^2/\sigma_\mu^2}\right)$$

$$[\sigma^2|\mu, y] \sim \text{IG}(A + \tfrac{n}{2}, B + \tfrac{1}{2}\|y - \mu 1\|^2)$$

Central MCMC Theoretical Result

Theory says that successive draws from

$$[\mu|\sigma^2, y] \quad \text{and} \quad [\sigma^2|\mu, y]$$

eventually leads to samples from

$$[\mu, \sigma^2|y]!$$

This is known as Gibbs sampling

and is a special case of
Markov Chain Monte Carlo (MCMC)

Live R Demonstration

We will now do a live

Gibbs sampling

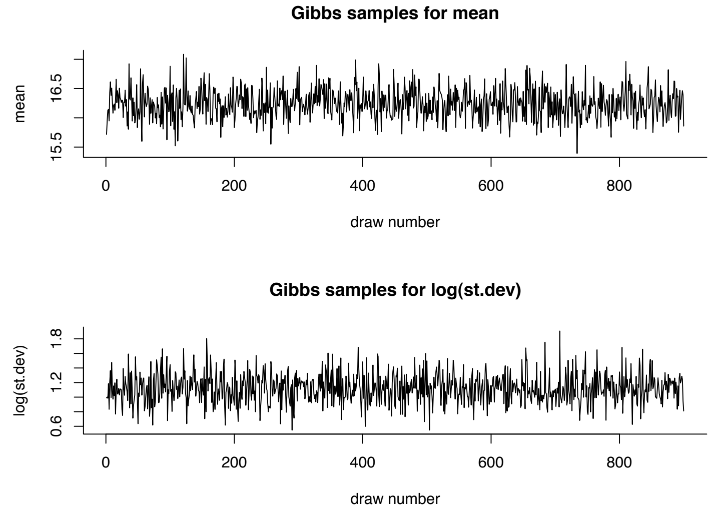
demonstration using

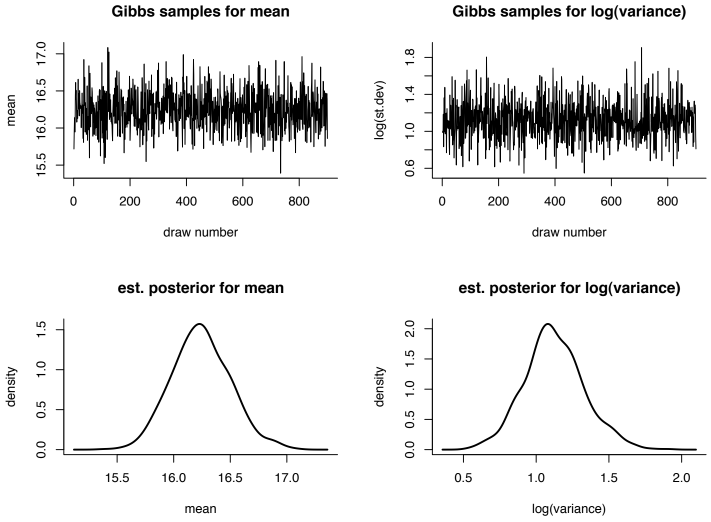
R

Illustration of MCMC

The following slides shows how MCMC works for the univariate normal example.

The first plots are successive draws from the full conditionals.





Practical MCMC

MCMC is a young (1990+) branch of Statistics and has several practical issues; e.g.

- partitioning of parameters,
- starting values,
- correlation between successive draws,
- convergence to required posteriors $\Longleftrightarrow$ length of burn-in.
- number of draws.

MCMC Software

The most sophisticated MCMC software is from

The BUGS Project (www.mrc-bsu.cam.ac.uk/bugs)

BUGS is an acronym for

Bayesian inference Using Gibbs Sampling

BUGS in Advanced Data Analysis

We will use BUGS via:

- The `WinBUGS` package available only on `Windows`.
- The `R` package `BRugs`; which allows `WinBUGS` to be called from inside `R`.