Gibbs sampler in Python

With R and OpenBUGS comparisons

Tyler Olson Alex Zajichek

Fall 2016

December 5, 2016

Table of contents

Introduction
- Gibbs sampling

Models
- Multi-parameter Normal model with conjugate prior
  - Full - conditional derivations
  - Results
- Gamma-poisson hierarchical model
  - Full - conditional derivations
  - Metropolis-Hastings algorithm
  - Results

Gibbs sampling

Goal: Sample from joint posterior distribution

Gibbs sampling

Goal: Sample from joint posterior distribution

- Reduce a single multi-dimensional problem into multiple univariate problems

Gibbs sampling

Goal: Sample from joint posterior distribution

- Reduce a single multi-dimensional problem into multiple univariate problems
- Full conditional distributions for model parameters

Gibbs sampling

Goal: Sample from joint posterior distribution

- Reduce a single multi-dimensional problem into multiple univariate problems
- Full conditional distributions for model parameters
    1. Unnormalized joint-posterior distribution

Gibbs sampling

Goal: Sample from joint posterior distribution

- Reduce a single multi-dimensional problem into multiple univariate problems
- Full conditional distributions for model parameters
  1. Unnormalized joint-posterior distribution
  2. Factor out terms containing parameter of interest

Gibbs sampling

Goal: Sample from joint posterior distribution

- Reduce a single multi-dimensional problem into multiple univariate problems
- Full conditional distributions for model parameters
    1. Unnormalized joint-posterior distribution
    2. Factor out terms containing parameter of interest
    3. Product of terms is proportional to full conditional distribution

Gibbs sampling

Goal: Sample from joint posterior distribution

- Reduce a single multi-dimensional problem into multiple univariate problems
- Full conditional distributions for model parameters
    1. Unnormalized joint-posterior distribution
    2. Factor out terms containing parameter of interest
    3. Product of terms is proportional to full conditional distribution
    4. If possible, identify parametric family

[1] Cowles, 2013

Multi-parameter Normal model with conjugate prior

Multi-parameter Normal model with conjugate prior

$$(\mu, \sigma^2) \quad \sim \quad IG(\alpha, \beta) \times N(\mu_0, \frac{\sigma^2}{\kappa})$$

Multi-parameter Normal model with conjugate prior

$$
\begin{aligned}
(\mu, \sigma^2) &\sim IG(\alpha, \beta) \times N(\mu_0, \frac{\sigma^2}{\kappa}) \\
\bar{y}|\mu, \sigma^2 &\sim N(\mu, \frac{\sigma^2}{n})
\end{aligned}
$$

Joint-posterior distribution

Joint-posterior distribution

$$p(\mu, \sigma^2 | \boldsymbol{y}) \propto p(\mu, \sigma^2, \boldsymbol{y})$$

Joint-posterior distribution

$$\begin{aligned} p(\mu, \sigma^2 | \boldsymbol{y}) &\propto p(\mu, \sigma^2, \boldsymbol{y}) \\ &= p(\bar{y} | \mu, \sigma^2) p(\mu | \sigma^2) p(\sigma^2) \end{aligned}$$

Joint-posterior distribution

$$
\begin{aligned}
p(\mu, \sigma^2 | \boldsymbol{y}) &\propto p(\mu, \sigma^2, \boldsymbol{y}) \\
&= p(\bar{y} | \mu, \sigma^2) p(\mu | \sigma^2) p(\sigma^2) \\
&= \frac{\sqrt{n}}{\sqrt{2\pi}\sigma} e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2} \frac{\sqrt{\kappa}}{\sqrt{2\pi}\sigma} e^{\frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2} \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{1}{(\sigma^2)^{\alpha+1}} e^{\frac{-\beta}{\sigma^2}}
\end{aligned}
$$

Derivation of full-conditional for μ

Derivation of full-conditional for μ

$$p(\mu|\sigma^2, \boldsymbol{y}) \quad \propto \quad e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2 + \frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2}$$

Derivation of full-conditional for μ

$$
\begin{aligned}
p(\mu|\sigma^2, \boldsymbol{y}) &\propto e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2 + \frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}-\mu)^2 + \kappa(\mu-\mu_0)^2)}
\end{aligned}
$$

Derivation of full-conditional for μ

$$
\begin{aligned}
p(\mu|\sigma^2, \boldsymbol{y}) &\propto e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2+\frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}-\mu)^2+\kappa(\mu-\mu_0)^2)} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}^2-2\bar{y}\mu+\mu^2)+\kappa(\mu^2-2\mu\mu_0+\mu_0^2))}
\end{aligned}
$$

Derivation of full-conditional for μ

$$
\begin{aligned}
p(\mu|\sigma^2, \boldsymbol{y}) &\propto e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2 + \frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}-\mu)^2 + \kappa(\mu-\mu_0)^2)} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}^2 - 2\bar{y}\mu + \mu^2) + \kappa(\mu^2 - 2\mu\mu_0 + \mu_0^2))} \\
&\propto e^{\frac{-(n+\kappa)}{2\sigma^2}(\mu^2 - 2\mu\frac{n\bar{y}+\kappa\mu_0}{n+\kappa})}
\end{aligned}
$$

Derivation of full-conditional for μ

$$
\begin{aligned}
p(\mu|\sigma^2, \boldsymbol{y}) &\propto e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2 + \frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}-\mu)^2 + \kappa(\mu-\mu_0)^2)} \\
&= e^{\frac{-1}{2\sigma^2}(n(\bar{y}^2 - 2\bar{y}\mu + \mu^2) + \kappa(\mu^2 - 2\mu\mu_0 + \mu_0^2))} \\
&\propto e^{\frac{-(n+\kappa)}{2\sigma^2}(\mu^2 - 2\mu\frac{n\bar{y}+\kappa\mu_0}{n+\kappa})} \\
&\propto e^{\frac{-(n+\kappa)}{2\sigma^2}(\mu^2 - 2\mu\frac{n\bar{y}+\kappa\mu_0}{n+\kappa} + (\frac{n\bar{y}+\kappa\mu_0}{n+\kappa})^2)} \\
\mu|\sigma^2, \boldsymbol{y} &\sim N\left(\frac{n\bar{y}+\kappa\mu_0}{n+\kappa}, \frac{\sigma^2}{n+\kappa}\right)
\end{aligned}
$$

Derivation of full-conditional for σ^2

Derivation of full-conditional for σ^2

$$p(\sigma^2|\mu, \boldsymbol{y}) \quad \propto \quad \frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\frac{\sqrt{\kappa}}{\sqrt{2\pi}\sigma}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{1}{(\sigma^2)^{\alpha+1}}e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2+\frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2+\frac{-\beta}{\sigma^2}}$$

Derivation of full-conditional for σ^2

$$
\begin{aligned}
p(\sigma^2|\mu, \boldsymbol{y}) &\propto \frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\frac{\sqrt{\kappa}}{\sqrt{2\pi}\sigma}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{1}{(\sigma^2)^{\alpha+1}} e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2+\frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2+\frac{-\beta}{\sigma^2}} \\
&\propto \frac{1}{(\sigma^2)^{\alpha+2}} e^{\frac{-1}{\sigma^2}(\frac{n}{2}(\bar{y}-\mu)^2+\frac{\kappa}{2}(\mu-\mu_0)^2+\beta)}
\end{aligned}
$$

Derivation of full-conditional for σ^2

$$
\begin{aligned}
p(\sigma^2|\mu, \boldsymbol{y}) &\propto \frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\frac{\sqrt{\kappa}}{\sqrt{2\pi}\sigma}\frac{\beta^\alpha}{\Gamma(\alpha)}\frac{1}{(\sigma^2)^{\alpha+1}}e^{\frac{-n}{2\sigma^2}(\bar{y}-\mu)^2+\frac{-\kappa}{2\sigma^2}(\mu-\mu_0)^2+\frac{-\beta}{\sigma^2}} \\
&\propto \frac{1}{(\sigma^2)^{\alpha+2}}e^{\frac{-1}{\sigma^2}\left(\frac{n}{2}(\bar{y}-\mu)^2+\frac{\kappa}{2}(\mu-\mu_0)^2+\beta\right)} \\
\sigma^2|\mu, \boldsymbol{y} &\sim IG\left(\alpha+1, \frac{n}{2}(\bar{y}-\mu)^2+\frac{\kappa}{2}(\mu-\mu_0)^2+\beta\right)
\end{aligned}
$$

Results

Results

	Python ($\approx$ .763 sec)			R ($\approx$ .496 sec)			OpenBUGS		
Parameter	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
μ	50.335	0.629	50.336	50.349	0.630	50.355	50.330	0.634	50.330
σ^2	11.939	1.976	11.736	11.920	1.948	11.720	11.93	1.98	11.730

Gamma-poisson hierarchical model

Gamma-poisson hierarchical model

$$x_i|\theta_i, t_i, \alpha, \beta \sim Poisson(\theta_i t_i)$$

Gamma-poisson hierarchical model

$$
\begin{aligned}
x_i|\theta_i, t_i, \alpha, \beta &\sim Poisson(\theta_i t_i) \\
\theta_i|\alpha, \beta &\sim Gamma(\alpha, \beta)
\end{aligned}
$$

Gamma-poisson hierarchical model

$$
\begin{aligned}
x_i|\theta_i, t_i, \alpha, \beta &\sim \mathit{Poisson}(\theta_i t_i) \\
\theta_i|\alpha, \beta &\sim \mathit{Gamma}(\alpha, \beta) \\
\alpha &\sim \mathit{Exponential}(\lambda_1)
\end{aligned}
$$

Gamma-poisson hierarchical model

$$\begin{aligned}
x_i|\theta_i, t_i, \alpha, \beta &\sim Poisson(\theta_i t_i)\\
\theta_i|\alpha, \beta &\sim Gamma(\alpha, \beta)\\
\alpha &\sim Exponential(\lambda_1)\\
\beta &\sim Exponential(\lambda_2)
\end{aligned}$$

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

$$p(\alpha, \beta, \boldsymbol{\theta}|\boldsymbol{x}) \quad \propto \quad p(\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{x})$$

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

$$\begin{aligned} p(\alpha, \beta, \boldsymbol{\theta} | \boldsymbol{x}) &\propto p(\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{x}) \\ &= p(\boldsymbol{x} | \alpha, \beta, \boldsymbol{\theta}) p(\alpha, \beta, \boldsymbol{\theta}) \end{aligned}$$

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

$$\begin{aligned} p(\alpha, \beta, \boldsymbol{\theta}|\boldsymbol{x}) &\propto p(\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{x}) \\ &= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\alpha, \beta, \boldsymbol{\theta}) \\ &= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\boldsymbol{\theta}|\alpha, \beta)p(\alpha, \beta) \end{aligned}$$

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

$$\begin{aligned}
p(\alpha, \beta, \boldsymbol{\theta}|\boldsymbol{x}) &\propto p(\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{x}) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\alpha, \beta, \boldsymbol{\theta}) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\boldsymbol{\theta}|\alpha, \beta)p(\alpha, \beta) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\boldsymbol{\theta}|\alpha, \beta)p(\alpha)p(\beta)
\end{aligned}$$

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

$$\begin{aligned}
p(\alpha, \beta, \boldsymbol{\theta}|\boldsymbol{x}) &\propto p(\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{x}) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta}) p(\alpha, \beta, \boldsymbol{\theta}) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\alpha, \beta) p(\alpha, \beta) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\alpha, \beta) p(\alpha) p(\beta) \\
&= \prod_{i=1}^{n} \frac{e^{-\theta_i t_i} (\theta_i t_i)^{x_i}}{x_i} \times \prod_{i=1}^{n} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta_i^{\alpha-1} e^{-\beta\theta_i} \times \lambda_1 e^{-\lambda_1 \alpha} \lambda_2 e^{-\lambda_2 \beta}
\end{aligned}$$

Joint-posterior distribution

If $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ and $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, then

$$
\begin{aligned}
p(\alpha, \beta, \boldsymbol{\theta}|\boldsymbol{x}) &\propto p(\alpha, \beta, \boldsymbol{\theta}, \boldsymbol{x}) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\alpha, \beta, \boldsymbol{\theta}) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\boldsymbol{\theta}|\alpha, \beta)p(\alpha, \beta) \\
&= p(\boldsymbol{x}|\alpha, \beta, \boldsymbol{\theta})p(\boldsymbol{\theta}|\alpha, \beta)p(\alpha)p(\beta) \\
&= \prod_{i=1}^{n} \frac{e^{-\theta_i t_i}(\theta_i t_i)^{x_i}}{x_i} \times \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)}\theta_i^{\alpha-1}e^{-\beta\theta_i} \times \lambda_1 e^{-\lambda_1\alpha}\lambda_2 e^{-\lambda_2\beta} \\
&= \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n}\lambda_1 e^{-\lambda_1\alpha}\lambda_2 e^{-\lambda_2\beta} \times \prod_{i=1}^{n} \frac{e^{-\theta_i(t_i+\beta)}\theta^{x_i+\alpha-1}t_i^{x_i}}{x_i}
\end{aligned}
$$

Derivation of full-conditional for $\boldsymbol{\theta}$

Derivation of full-conditional for $\boldsymbol{\theta}$

$$p(\theta_i | \boldsymbol{\theta}_{-i}, \alpha, \beta, \boldsymbol{x}) \quad \propto \quad e^{-\theta_i(t_i+\beta)} \theta_i^{x_i+\alpha-1}$$

Derivation of full-conditional for $\boldsymbol{\theta}$

$$
\begin{aligned}
p(\theta_i|\boldsymbol{\theta}_{-i}, \alpha, \beta, \boldsymbol{x}) &\propto e^{-\theta_i(t_i+\beta)}\theta_i^{x_i+\alpha-1} \\
\theta_i|\boldsymbol{\theta}_{-i}, \alpha, \beta, \boldsymbol{x} &\sim Gamma(x_i+\alpha, t_i+\beta)
\end{aligned}
$$

Derivation of full-conditional for β

Derivation of full-conditional for β

$$p(\beta|\alpha, \boldsymbol{\theta}, \boldsymbol{x}) \quad \propto \quad \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i t_i - \theta_i \beta}$$

Derivation of full-conditional for β

$$\begin{aligned} p(\beta|\alpha, \boldsymbol{\theta}, \boldsymbol{x}) &\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i t_i - \theta_i \beta} \\ &\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i \beta} \end{aligned}$$

Derivation of full-conditional for β

$$
\begin{aligned}
p(\beta|\alpha, \boldsymbol{\theta}, \boldsymbol{x}) &\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i t_i - \theta_i \beta} \\
&\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i \beta} \\
&= \beta^{n\alpha} e^{-\lambda_2 \beta} e^{-\sum_{i=1}^{n} \theta_i \beta}
\end{aligned}
$$

Derivation of full-conditional for β

$$
\begin{aligned}
p(\beta|\alpha, \boldsymbol{\theta}, \boldsymbol{x}) &\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i t_i - \theta_i \beta} \\
&\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i \beta} \\
&= \beta^{n\alpha} e^{-\lambda_2 \beta} e^{-\sum_{i=1}^{n} \theta_i \beta} \\
&= \beta^{n\alpha} e^{-\beta(\lambda_2 + \sum_{i=1}^{n} \theta_i)}
\end{aligned}
$$

Derivation of full-conditional for β

$$
\begin{aligned}
p(\beta|\alpha, \boldsymbol{\theta}, \boldsymbol{x}) &\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i t_i - \theta_i \beta} \\
&\propto \beta^{n\alpha} e^{-\lambda_2 \beta} \prod_{i=1}^{n} e^{-\theta_i \beta} \\
&= \beta^{n\alpha} e^{-\lambda_2 \beta} e^{-\sum_{i=1}^{n} \theta_i \beta} \\
&= \beta^{n\alpha} e^{-\beta(\lambda_2 + \sum_{i=1}^{n} \theta_i)} \\
\beta|\alpha, \boldsymbol{\theta}, \boldsymbol{x} &\sim Gamma(n\alpha + 1, \lambda_2 + \sum_{i=1}^{n} \theta_i)
\end{aligned}
$$

Derivation of full-conditional for α

Derivation of full-conditional for α

$$p(\alpha|\beta, \boldsymbol{\theta}, \boldsymbol{x}) \propto \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} e^{-\lambda_1 \alpha} \prod_{i=1}^{n} \theta_i^{x_i+\alpha-1}$$

Derivation of full-conditional for α

$$
\begin{aligned}
p(\alpha|\beta,\boldsymbol{\theta},\boldsymbol{x}) &\propto \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} e^{-\lambda_1\alpha} \prod_{i=1}^{n} \theta_i^{x_i+\alpha-1} \\
&\propto \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} e^{-\lambda_1\alpha} \prod_{i=1}^{n} \theta_i^{\alpha}
\end{aligned}
$$

Metropolis-Hastings algorithm

Metropolis-Hastings algorithm

If f(x) is the unnormalized density of interest, x_i is the value at the i^{th} iteration, and $q(x|x_{i-1})$ is the chosen proposal function, then

Metropolis-Hastings algorithm

If f(x) is the unnormalized density of interest, x_i is the value at the i^{th} iteration, and $q(x|x_{i-1})$ is the chosen proposal function, then

1. Sample $x^* \sim q(x|x_{i-1})$

Metropolis-Hastings algorithm

If f(x) is the unnormalized density of interest, x_i is the value at the i^{th} iteration, and $q(x|x_{i-1})$ is the chosen proposal function, then

1. Sample $x^* \sim q(x|x_{i-1})$
2. Calculate the acceptance probability

$$p = P(\text{accept } x^*) = min\left(1, \frac{f(x^*)q(x_{i-1}|x^*)}{f(x_{i-1})q(x^*|x_{i-1})}\right)$$

Metropolis-Hastings algorithm

If f(x) is the unnormalized density of interest, x_i is the value at the i^{th} iteration, and $q(x|x_{i-1})$ is the chosen proposal function, then

1. Sample $x^* \sim q(x|x_{i-1})$
2. Calculate the acceptance probability

$$p = P(\text{accept } x^*) = min\left(1, \frac{f(x^*)q(x_{i-1}|x^*)}{f(x_{i-1})q(x^*|x_{i-1})}\right)$$

3. If $Y \sim Bernoulli(p)$, then

$$x_i = \begin{cases} x^* & \text{if Y=1} \\ x_{i-1} & \text{if Y=0} \end{cases}$$

[2] Niemi, 2013

Metropolis-Hastings algorithm

Our Implementation:

Metropolis-Hastings algorithm

Our Implementation:

For each iteration of the Gibbs sampler:

- For i in 1:1000

Metropolis-Hastings algorithm

Our Implementation:

For each iteration of the Gibbs sampler:

- For i in 1:1000
    - Proposal function:

$$x|x_{i-1} \sim Gamma(\alpha = x_{i-1}, \beta = 1)$$

Results

Results

	Python ($\approx$ 1 hr)			R (18.24 min)			OpenBUGS		
Parameter	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
α	0.780	0.307	0.739	0.789	0.319	0.749	0.795	0.295	0.754
β	1.193	0.635	1.092	1.208	0.661	1.093	1.218	0.634	1.116
θ_1	0.060	0.025	0.057	0.061	0.025	0.057	0.060	0.025	0.057
θ_2	0.106	0.081	0.086	0.106	0.080	0.087	0.106	0.081	0.087
θ_3	0.091	0.038	0.086	0.090	0.038	0.085	0.090	0.038	0.085
θ_4	0.116	0.030	0.113	0.116	0.030	0.114	0.116	0.030	0.113
θ_5	0.593	0.310	0.544	0.590	0.307	0.534	0.590	0.306	0.539
θ_6	0.606	0.137	0.594	0.606	0.137	0.596	0.608	0.136	0.598
θ_7	0.808	0.640	0.651	0.825	0.660	0.659	0.822	0.658	0.655
θ_8	0.837	0.680	0.650	0.829	0.658	0.662	0.829	0.665	0.655
θ_9	1.512	0.739	1.392	1.487	0.718	1.362	1.477	0.715	1.356
θ_{10}	1.955	0.414	1.922	1.950	0.417	1.920	1.949	0.419	1.918

Results

	Python ($\approx$ 1 hr)			R (18.24 min)			OpenBUGS		
Parameter	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
α	0.780	0.307	0.739	0.789	0.319	0.749	0.795	0.295	0.754
β	1.193	0.635	1.092	1.208	0.661	1.093	1.218	0.634	1.116
θ_1	0.060	0.025	0.057	0.061	0.025	0.057	0.060	0.025	0.057
θ_2	0.106	0.081	0.086	0.106	0.080	0.087	0.106	0.081	0.087
θ_3	0.091	0.038	0.086	0.090	0.038	0.085	0.090	0.038	0.085
θ_4	0.116	0.030	0.113	0.116	0.030	0.114	0.116	0.030	0.113
θ_5	0.593	0.310	0.544	0.590	0.307	0.534	0.590	0.306	0.539
θ_6	0.606	0.137	0.594	0.606	0.137	0.596	0.608	0.136	0.598
θ_7	0.808	0.640	0.651	0.825	0.660	0.659	0.822	0.658	0.655
θ_8	0.837	0.680	0.650	0.829	0.658	0.662	0.829	0.665	0.655
θ_9	1.512	0.739	1.392	1.487	0.718	1.362	1.477	0.715	1.356
θ_{10}	1.955	0.414	1.922	1.950	0.417	1.920	1.949	0.419	1.918

- Runtime: find optimal proposal function

References

[1] Cowles, M. K. (2013). Applied bayesian statistics: With R and OpenBUGS examples. New York: Springer.

[2] Niemi, J. (2013, March 3). Video. Retrieved from https://www.youtube.com/watch?v=VGRVRjr0vyw