Metropolis-Hastings exercises

Josep Fortiana
Facultat de Matemàtiques i Informàtica, UB

2017-11-26

Exercise 1. *Prepare and run an MCMC simulation for a binomial model* $X \sim \mathrm{B}(n, \theta)$*, where* n *is a positive integer and the probability parameter* θ *takes values in* $(0, 1)$ *with a prior* $\mathrm{Beta}(\alpha, \beta)$*.*

For instance, we can retake the drug response data from some lessons ago. There, from prior knowledge, we decided $\alpha = 9.2$*,* $\beta = 13.8$ *and we learnt of an experiment with* $n = 20$ *patients where* $x = 15$ *of them responded to the drug treatment.*

Generate a simulated sample for the posterior pdf and, from it, the posterior expectation, variance, mode, and a 95% credible interval. Compare these results, obtained with with several chain lengths, burn-in interval (the discarded initial segment in the chain after which samples are assumed to be distributed following the stationary pdf), and candidate generation densities with the corresponding theoretical ones, derived from the exact posterior pdf, which we know is a $\mathrm{Beta}(\alpha + x, \beta + n - x)$*.*

Obtain a simulated sample for the prior predictive pmf for x*, for a sample size* $n = 20$*. Plot the estimated pmf and compare it with the corresponding theoretical object. Do the same for the posterior predictive pmf for the number* y *of patients from a new batch of* $m = 18$ *patients, including the observed evidence that* $x = 15$ *out of* $n = 20$ *patients have responded to the drug treatment.*

Exercise 2. *Prepare and run an MCMC simulation for a simple least squares regression model following the explanatory text and R scripts in Florian Hartig's blog.*

Exercise 3 (Example in Section 6.7 of Jim Albert (2009), Bayesian computation with R, 2nd ed. - Learning about a normal population from grouped data.)**.** *Suppose a random sample is taken from a normal population with mean* μ *and standard deviation* σ*. But one only observes the data in "grouped" form, where the frequencies of the data in bins are recorded. For example, suppose one is interested in learning about the mean and standard deviation of the heights (in inches) of men from a local college. One is given the summary frequency data shown in Table 6.1. One sees that 14 men were shorter than 66 inches, 30 men had heights between 66 and 68 inches, and so on.*

Data: Grouped frequency data for heights of male students at a college

Height interval (in.)	*Frequency*
less than 66	14
between 66 *and* 68	30
between 68 *and* 70	49
between 70 *and* 72	70
between 72 *and* 74	33
over 74	15

Exercise 4 (Bolstad (2010), *Understanding computational Bayesian Statistics,* Chap. 6, Example 8)**.** *The object of this exercise is to build Metropolis-Hastings algorithms with target pdf:*

$$g(\theta|y) \propto 0.8 \times \exp\{-\frac{1}{2}\,\theta^2\} + 0.2 \times \frac{1}{2}\,\exp\{-\frac{1}{2}\,\frac{(\theta-3)^2}{2^2}\},$$

a mixture of a N(0, 1) *and a* N(3, 4)*. As a guide, you might want to see the function* `normMixMH` *from the* `Bolstad2` *package.*

1. *Previous work: g is written as a non-normalized pdf, since that is all we need for MH. However, the normalization constant is known ($1/\sqrt{2\pi}$). Furthermore, we know how to simulate a mixture. Then, plot this pdf, generate a sequence of (independent) random numbers $\sim g$, comparing the histogram with the pdf. Find the sample size needed to obtain a good (or, at least, acceptable) proximity to the target pdf.*

2. *Write a Metropolis-Hastings code with the candidate proposal kernel:*

$$k(\theta'|\theta) = \exp\{-\frac{1}{2}(\theta'-\theta)^2\}$$

   The probability of acceptance is:

$$\min\left\{1, \frac{h_x(\theta')\,k(\theta|\theta')}{h_x(\theta)\,k(\theta'|\theta)}\right\} = \min\left\{1, \frac{h_x(\theta')}{h_x(\theta)}\right\}.$$

   The symmetry of the candidate generation implies this M-H algorithm reduces to the simple Metropolis algorithm. Use $\theta = 2$ as the start value in the trajectory. Test the algorithm with several values of chain length and burn-in *initial discarded segment.*

3. *Write a Metropolis-Hastings code with the candidate proposal kernel:*

$$k(\theta'|\theta) = q(\theta') = \exp\{-\frac{1}{2}\,(\theta'/3)^2\},$$

   a $N(0, 3^2)$ *pdf. M-H algorithm using such a candidate proposal kernel is called* independent candidate M-H algorithm*. With it, the probability of acceptance is:*

$$\min\left\{1, \frac{h_x(\theta')\,k(\theta|\theta')}{h_x(\theta)\,k(\theta'|\theta)}\right\} = \min\left\{1, \frac{h_x(\theta')\cdot q(\theta)}{h_x(\theta)\cdot q(\theta')}\right\}.$$

   Use $\theta = 0.4448$ as the start value in the trajectory. Test the algorithm with several values of chain length and burn-in *initial discarded segment. Discuss traceplot, ACF and histogram of the chains tested.*

Exercise 5 (Modeling Data with Cauchy Errors - Example in Jim Albert (2009), Bayesian Computation with R, 2nd ed., Section 6.9.)**.** *In this example we study a famous dataset by Charles Darwin (1876), The Effects of Cross- and Self-fertilisation in the Vegetable Kingdom, resulting from an experiment to examine the superiority of cross-fertilized plants over self-fertilized plants.*

n=15 pairs of plants were used. Each pair consisted of one cross-fertilized plant and one self-fertilized plant which germinated at the same time and grew in the same pot. The plants were measured at a fixed time after planting and the difference in heights between the cross- and self-fertilized plants are recorded in eighths of an inch.

There are two outliers, which prevent a gaussian model for this data. We use, instead, a Cauchy model.