

Math 153 - Lab 4 - Metropolis-Hastings

Recall that the Metropolis Hastings (M-H) ratio is given as

$$r(x, y) = \min\left\{1, \frac{f(y)g(x|y)}{f(x)g(y|x)}\right\}$$

and the algorithm functions as follows:

At time t , suppose you are at value x , i.e. $w_t = x$. You then draw an observation from $g(y|x)$ (which may or may not actually depend on x). You then assign $w_{t+1} = y$ with probability $r(x, y)$, otherwise you set $w_{t+1} = x$.

We have argued that w_t , $t = 0, 1, \dots$ will have stationary (limiting) distribution $f(\cdot)$.

1. (Non-Bayesian/Educational) - Suppose that your goal is to estimate the value of π , which you know shows up in the area calculation of the unit circle. So we want to generate points in the square $[-1, 1] \times [-1, 1]$, and check to see if they live in the circle. The proportion of points that live in the circle is converging to the probability, which is related to the area. But rather than just making *runif* calls to generate points, we decide to use a random walk chain. We will generate the i^{th} coordinate ($i = 1, 2$) $y^i \sim U(w_t^i - \epsilon, w_t^i + \epsilon)$ for $\epsilon > 0$, so our neighborhood is actually a square centered at our current location w_t .

There are 3 ingredients here. Use the following plots and output to talk about them and how they relate to each other.

- i) the choice of ϵ
- ii) the choice of w_0 (try the chain starting at (0,0) and compare it to the chain that starts at (1,1))
- iii) the length of the run, which we might call n .
 - a) To start, suppose that naively, you mix an acceptance algorithm with a markov chain, so rather than rejecting (if your point isn't within the stationary support), you just stay. Use $\epsilon < 2$. What is the problem here, and why? (Why is it important that you do the unit circle and not just a quarter of the unit circle and the unit square?)
 - b) Create plots of the movement of the chain (if your chain is a $n \times 2$ matrix, just plot it with $t = 'l'$).
 - c) Create plots of the estimate, $\hat{\pi}_t$ against t and draw correspondence with the plots in b.
 - d) After figuring out the issues regarding iii), guess what the idea of 'burn-in', which involves not using all your values, is to fix this problem.

2. (Bayesian (but unnecessary)) - Our goal is to summarize the posterior distribution. Let's do it in a situation where the answers are known. Use M-H to summarize the posterior distribution of a binomial experiment with beta prior (and thus beta posterior). Using only the beta pdf (*dbeta*) compare inference (posterior mean/median/credible interval) when your prior is good, and when your prior is bad. Use your prior as the candidate density (so you are using

an independence chain). Use the same kinds of plots above to support your discussion of what is going on here.