# Monte Carlo Methods — Why Do They Work? — Part II

In Chapters 13, 15 and 16 Donovan and Mickey give us three specific Monte Carlo methods to consider:
* Chapter 13: The Metropolis algorithm, published 1953, but developed during the Manhattan project
* Chapter 15: The Metropolis-Hastings algorithm, published 1970, where Hastings made a critical improvement to the algorithm
* Chapter 16: The "Gibbs Sampling" (GS) algorithm, published 1984 by Geman and Geman, who were smoothing images

In the previous "Why Do They Work?" write-up, I showed why the Metropolis algorithm worked. In this write-up, I am going to show that Metropolis-Hastings works. I was surprised that the proof was a little more difficult than I had imagined, and getting the proof right told me I have to revise Problem Set 17 a little.

## The Metropolis-Hastings Algorithm

Let us summarize the Metropolis-Hastings algorithm. The situation we want to apply it to is still the same:
* We have $n$ bins and a set of desired probabilities $p_i$, where $i$ runs from 1 to $n$.
* In our quarterly iPhone sales example, $n$ was 4 and the four $p_i$'s were 0.1, 0.2, 0.4, and 0.3.

Then the core of the algorithm (repeated *ad nauseam)* was:

**Step 1:** You are in bin $i$. You choose a random move from $p_i$ to $p_j$. In Metropolis, the random move was to the nearest neighbor, and you had 0.5 chance of going to either of the two nearest neighbors. In Metropolis-Hastings, the probability of the random move is just denoted $g(j \mid i)$, which can be non-zero for non-nearest neighbors, and furthermore the probabilities $g(j \mid i)$ don't even have to be symmetric. E.g., it does ***not*** have to be the case that $g(i \mid j) = g(j \mid 1)$.

**Step 2:** Compute the appropriate ratio. The appropriate ratio for going from $i \to j$ is $\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}$. Now as you can see, unlike my proof for Metropolis, in Metropolis-Hastings, the ratio can be greater than 1. So that ratio has no chance of being conceptualized as a probability. In fact, if the appropriate ratio for going from $i \to j$ is $\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}$ and is less than 1, then the appropriate ratio for going from $j \to i$ is $\frac{p_i}{p_i} \frac{g(j \mid i)}{g(i \mid j)}$ which is exactly the inverse, will be greater than 1. So we **clamp** the ratio as follows: the **clamped** appropriate ratio for going from $i \to j$ is $\min\left(\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}, 1\right)$.

**Step 3:** Generate a random number between 0 and 1. If the number is less than clamped appropriate ratio, move to the proposed bin, and make a tally there. Otherwise stay in the current bin and make another tally in the current bin.

## Thinking Probabilistically Instead of Algorithmically

Suppose after running the algorithm for a while, **the chance** that the algorithm is currently in bin $i$ is $q_i$ where $i = 1, \ldots, n$. Note that the $q_i$ do not have to be related to the $p_i$. That relationship is what we are trying to prove! In fact, at the beginning, we start the algorithm off somewhere, and of course the $q_i$ are initially concentrated wherever we chose to start. We want to derive something about the $q_i$ once they settle down into an equilibrium.

Let's think about the $j \to i$ transition and the $i \to j$ transition. Again, these do not have to be neighbors. The chance of being in bin $j$ and going to bin $i$ is:

$q_j * g(i \mid j) *$ the clamped appropriate ratio for $j \to i$

While the chance of being in bin $i$ and going to bin $j$ is:

$q_i * g(j \mid i) *$ the clamped appropriate ratio for $i \to j$

## The Principle of Detailed Balance

When the $q_j$ settle down, all the flows out of any bin $i$ must equal all the flows into that same bin. So we could write down:

$$\sum_j q_i * g(j \mid i) * \text{the appropriate ratio for } i \to j$$
$$= \quad \sum_j q_j * g(i \mid j) * \text{the appropriate ratio for } j \to i$$

That would be called "The Principle of Balance." But we are going to demand something stronger, which is "The Principle of Detailed Balance." We are going to demand that the flow from bin $i$ to bin $j$ is equal to the flow from bin $j$ to bin $i$. Certainly if we demand that, then the Principle of Balance is also true. So, we demand it, and see where it leads. In other words, we see if we can satisfy:

$$q_i * g(j \mid i) * \text{the appropriate ratio for } i \to j$$
$$= \quad q_j * g(i \mid j) * \text{the appropriate ratio for } j \to i$$

Now let's stick in the clamped appropriate ratios, and the detailed balance condition becomes:

$$q_i * g(j \mid i) * \min\left(\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}, 1\right)$$
$$= \quad q_j * g(i \mid j) * \min\left(\frac{p_i}{p_j} \frac{g(j \mid i)}{g(i \mid j)}, 1\right)$$

Remember that at least one of $\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}$ or $\frac{p_i}{p_j} \frac{g(j \mid i)}{g(i \mid j)}$ was greater than or equal to 1 (because they are inverses of each other!). Just suppose it is $\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}$ that is greater than or equal to 1. Well, in that case, $\min\left(\frac{p_j}{p_i} \frac{g(i \mid j)}{g(j \mid i)}, 1\right)$ is just 1 and $\min\left(\frac{p_i}{p_j} \frac{g(j \mid i)}{g(i \mid j)}, 1\right)$ simplifies to $\frac{p_i}{p_j} \frac{g(j \mid i)}{g(i \mid j)}$.

So in that case, the Principle of Detailed Balance becomes:

$$q_i * g(j \mid i) * 1$$
$$= \quad q_j * g(i \mid j) * \frac{p_i}{p_j} \frac{g(j \mid i)}{g(i \mid j)}$$

Look at the lovely cancellation of all the $g$ factors! And this same cancellation would have happened if it was $\frac{p_i}{p_j} \frac{g(j \mid i)}{g(i \mid j)}$ that was greater than 1. Either way, we have:

$$q_i = q_j \frac{p_i}{p_j} \qquad \text{or} \qquad \frac{q_i}{p_i} = \frac{q_j}{p_j}$$

## Concluding the Argument

Now once things have settled down, if the Principle of Detailed Balance is satisfied, ***it is true for every pair***. So that means this ratio is the same ratio for any $i$ and $j$. Let’s call the constant ratio $\rho$. We have shown that

$$\frac{q_j}{p_j} = \rho$$

or

$$q_j = \rho p_j$$

The very last thing to observe is that the sum of all the $p_j$ is 1, because the $p_j$ are a probability distribution. And the sum of all the $q_j$ is 1, because the $q_j$ are also a probability distribution. The only way the $p_j$ and the $q_j$ can both be normalized is if $\rho$ is 1.

We have shown that the Metropolis-Hastings version of Monte Carlo works! Every $q_j$ settles down to the corresponding $p_j$.

This whole proof was surprisingly complicated. It told me I have to modify Problem Set 17 to use the clamped appropriate ratio. So I will be handing out an adjusted version of Problem Set 17 this morning.

## Footnote

I don’t see that the Principle of Detailed Balance has to be satisfied. Certainly the Principle of Balance has to be satisfied. I am unclear whether this distinction is a red flag telling me about a potential issue, or whether I just need to think about it harder to see that it is not in fact an issue. I think the argument that it is not an issue would go something like:

The Principle of Detailed Balance did not have to be satisfied. We assumed that it was and saw where it took us. Since it yielded a solution for the $q_j$ and whatever the $q_j$ settle down to has to be unique, then we have found the unique solution. The unique solution happens to satisfy the Principle of Detailed Balance, but since it does, it of course also satisfies the Principle of Balance. If we had found that The Principle of Detailed Balance was not satisfiable, then we would have had to try the weaker Principle of Balance and see what we could prove from that.