Technical Computing for the Earth Sciences, Lecture 7:

Markov Chain Monte Carlo
The Metropolis Algorithm

EARS 80.03

“Monte Carlo”?

“Monte Carlo”?

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 247 SEPTEMBER 1949 *Volume 44*

THE MONTE CARLO METHOD

NICHOLAS METROPOLIS AND S. ULAM
Los Alamos Laboratory

We shall present here the motivation and a general description of a method dealing with a class of problems in mathematical physics. The method is, essentially, a statistical approach to the study of differential equations, or more generally, of integro-differential equations that occur in various branches of the natural sciences.

“Monte Carlo”?

JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION

Number 247 SEPTEMBER 1949 *Volume 44*

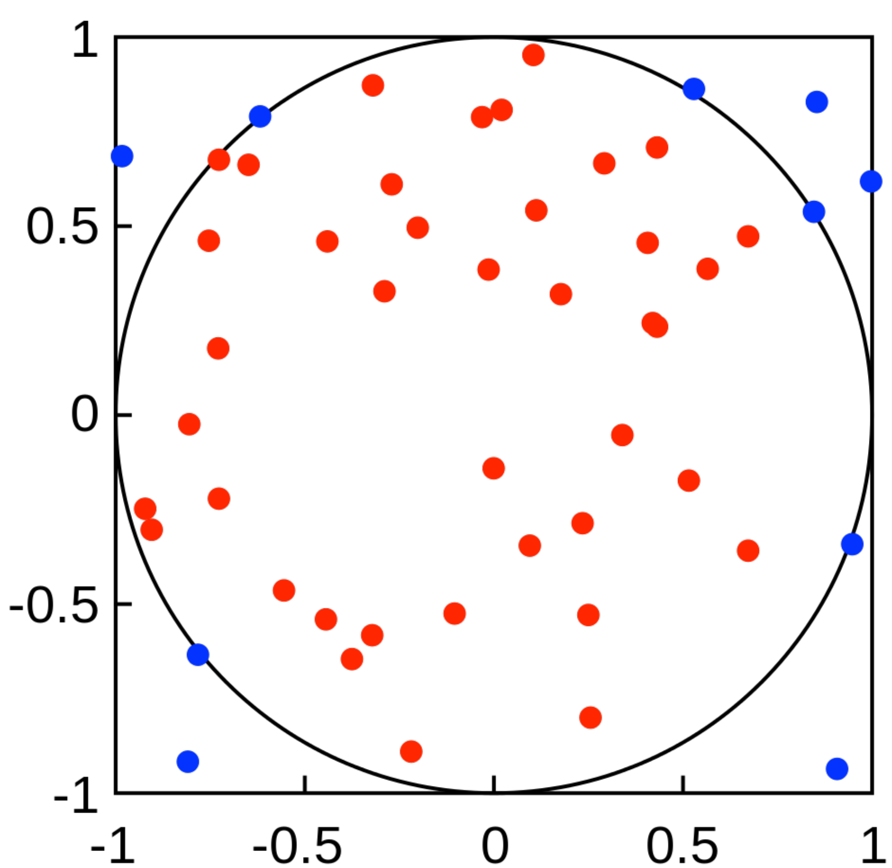
THE MONTE CARLO METHOD

NICHOLAS METROPOLIS AND S. ULAM

Los Alamos Laboratory

We shall present here the motivation and a general description of a method dealing with a class of problems in mathematical physics. The method is, essentially, a statistical approach to the study of differential equations, or more generally, of integro-differential equations that occur in various branches of the natural sciences.

Monte Carlo *integration*



“Monte Carlo”?

More broadly [wp]:

- “*Monte Carlo methods*, or *Monte Carlo experiments*, are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results.”
- “The underlying concept is to use randomness to solve problems that might be deterministic in principle”
- See in particular *Monte Carlo error propagation* [demo]

“Markov Chain”?

“A sequence of possible events in which the probability of each event depends only on the state attained in the previous event”

The Metropolis Algorithm

11.2 Metropolis and Metropolis-Hastings algorithms

The *Metropolis-Hastings algorithm* is a general term for a family of Markov chain simulation methods that are useful for sampling from Bayesian posterior distributions. We have already seen the Gibbs sampler in the previous section; it can be viewed as a special case of Metropolis-Hastings (as described in Section 11.3). Here we present the basic Metropolis algorithm and its generalization to the Metropolis-Hastings algorithm.

The Metropolis algorithm

The Metropolis algorithm is an adaptation of a random walk with an acceptance/rejection rule to converge to the specified target distribution. The algorithm proceeds as follows.

1. Draw a starting point θ^0, for which $p(\theta^0|y) > 0$, from a *starting distribution* $p_0(\theta)$. The starting distribution might, for example, be based on an approximation as described in Section 13.3. Or we may simply choose starting values dispersed around a crude approximate estimate of the sort discussed in Chapter 10.
2. For $t = 1, 2, \dots$:
   (a) Sample a *proposal* θ^* from a *jumping distribution* (or *proposal distribution*) at time t, $J_t(\theta^*|\theta^{t-1})$. For the Metropolis algorithm (but not the Metropolis-Hastings algorithm, as discussed later in this section), the jumping distribution must be *symmetric*, satisfying the condition $J_t(\theta_a|\theta_b) = J_t(\theta_b|\theta_a)$ for all θ_a, θ_b, and t.
   (b) Calculate the ratio of the densities,
   $$r = \frac{p(\theta^*|y)}{p(\theta^{t-1}|y)}. \tag{11.1}$$
   (c) Set
   $$\theta^t = \begin{cases} \theta^* & \text{with probability } \min(r, 1) \\ \theta^{t-1} & \text{otherwise.} \end{cases}$$

Gelman et al., p 278.

A note on probability notation

The Metropolis Algorithm

11.2 Metropolis and Metropolis-Hastings algorithms

The *Metropolis-Hastings algorithm* is a general term for a family of Markov chain simulation methods that are useful for sampling from Bayesian posterior distributions. We have already seen the Gibbs sampler in the previous section; it can be viewed as a special case of Metropolis-Hastings (as described in Section 11.3). Here we present the basic Metropolis algorithm and its generalization to the Metropolis-Hastings algorithm.

The Metropolis algorithm

The Metropolis algorithm is an adaptation of a random walk with an acceptance/rejection rule to converge to the specified target distribution. The algorithm proceeds as follows.

1. Draw a starting point θ^0, for which $p(\theta^0|y) > 0$, from a *starting distribution* $p_0(\theta)$. The starting distribution might, for example, be based on an approximation as described in Section 13.3. Or we may simply choose starting values dispersed around a crude approximate estimate of the sort discussed in Chapter 10.
2. For $t = 1, 2, \ldots$:
   (a) Sample a *proposal* θ^* from a *jumping distribution* (or *proposal distribution*) at time t, $J_t(\theta^*|\theta^{t-1})$. For the Metropolis algorithm (but not the Metropolis-Hastings algorithm, as discussed later in this section), the jumping distribution must be *symmetric*, satisfying the condition $J_t(\theta_a|\theta_b) = J_t(\theta_b|\theta_a)$ for all θ_a, θ_b, and t.

   (b) Calculate the ratio of the densities,

$$r = \frac{p(\theta^*|y)}{p(\theta^{t-1}|y)}. \tag{11.1}$$

   (c) Set

$$\theta^t = \begin{cases} \theta^* & \text{with probability } \min(r, 1) \\ \theta^{t-1} & \text{otherwise.} \end{cases}$$

Gelman et al., p 278.