

Monte Carlo Methods for Mortals

Santhosh Nadig

July 3, 2018

1 Introduction

Here's a scenario: assume that you are playing solitaire (the card game) and wondering what's the probability of a successful outcome? Well, this was a question that occurred to Stan Ulam (the Polish mathematician) who was recovering from illness and playing solitaire to while away his time. He tried to solve it using combinatorics and failed. Finally, he thought of simulating a very large number of games (with cards shuffled at random, of course) and just counting the number of successes. He, and his friend, Von Neumann, who knew a thing or two about computers and programming, set out to do this calculation and eventually used this method for scientific purposes.

This is the central idea of MC. Use randomness to "simulate" a large number of experiments. Infer / calculate a quantity based on the outcome of the experiments. More the number of experiments, the better!

What are Monte Carlo (MC) methods? An (over)simplified view is that they are tools for:

1. Estimating (multi-dimensional) integrals
2. Estimating probabilities
3. ... etc.

First, a note about estimating integrals. The usual deterministic methods of numerical integration (such as trapezoidal method) suffer from what's known as the "curse of dimensionality". That is, for a fixed error the number of function evaluations (of the integrand) grows exponentially with the "dimension of integration". In general, calculations that grow exponentially are frowned upon by engineers and scientists. At the risk of giving away the story-line, it suffices to say that MC methods do not suffer from the aforementioned "curse." I will (probably) give examples of this later on.

2 Estimating Probabilities

2.1 Random Variables

What is a random variable? As someone said, it is neither random nor variable. For example, let X denote the resistance of a resistor (assume fixed temperature and other conditions, just for argument's sake). X is an *unknown constant* (or, only known to the supreme fascist, SF). We do have, however, a series of measurements x_1, x_2, \dots from an ohm-meter. These measurements are called realizations/observations/samples of the random

variable. So, now we have a random variable X and its realizations (or samples) x_1, x_2, \dots and so on.

The expected value (mean, for ordinary mortals) of a random variable, X , is defined as

$$E[X] = \int xp(x)dx,$$

where $p(x)$ is the probability density function (pdf). Assume that $p(x)$ is a “normally distributed” (bell curve) with mean m and standard deviation σ . What it means is that if you plot a histogram of x_1, x_2, \dots , it resembles the bell shape centered around m and its width $\propto \sigma$.

So, in theory, we can “determine” the expected value simply by performing the integration. However, given N realizations x_1, \dots, x_N of X , we can “estimate” the expected value (and hence, the integral) as

$$E[X] = m = \int xp(x)dx \approx \frac{1}{N} \sum_{i=1}^N x_i.$$

The estimate gets better (closer to m) as the number of samples, N , increases.

The above idea might not be a revelation. But, let us assume that we have an integral of the kind:

$$\int g(x)f(x)dx,$$