Monte Carlo Integration

Stefan Kapetanović

June 2025

Abstract

This is an explanation of Monte Carlo integration and the concepts required for its definition.

1 Expected value

1.1 Definition

Given a discrete random variable X with possible values from the set $\{x_1, x_2, ..., x_n\}$ and a probability mass function p (sometimes called the discrete density function), the expected value is defined as

$$E[X] = \sum_{i=1}^{n} x_i p(x_i)$$

If the number of possible values for X is countably infinite, then the same definition is valid for the infinite n, assuming some condition (for example absolute convergence) holds, which ensures the validity of such sum

$$E[X] = \sum_{i=1}^{\infty} x_i p(x_i)$$

If the number of possible values for X is uncountable, then we can approach the definition of expected value in the following way. First, we discretize the space of values for the random variable X. In general, this space is some measurable space, however, we will see it as a subset of $\mathbb{R}$. We discretize it by forming intervals $[x_i, x_{i+1}]$. The reason for doing this is that it allows us to express the current continuous problem using the previous discrete case. Now, we define a function q that will assign a value to each interval $[x_i, x_{i+1}]$, such that it represents the probability that X takes any value in that interval. Intuitively, q behaves just like p in the sense that it places a certain mass on some value/values. This way, we can use q to approximate expected value for a continuous random variable X

$$E[X] = \sum_{i=1}^{n} [x_i, x_{i+1}] q([x_i, x_{i+1}])$$

If we now start shrinking all intervals so that the size of the largest one always goes to 0, while still covering the whole range of values for X, we have

$$E[X] = \lim_{\substack{n \to \infty \\ |x_{i+1} - x_i| \to 0}} \sum_{i=1}^{n} [x_i, x_{i+1}] q([x_i, x_{i+1}]) = \int_{-\infty}^{+\infty} x p(x) dx$$

where $p(x)$ is the limit of q in the mantioned sense and represents the probability density function.

1.2 Sidetrack

Unlike with p in the discrete case, we can't just plug in some value of random variable X (like $p(x_1)$) and expect to get the probability that X takes that value. Instead, since it is a density, we need to multiply it with the length of the range of values, thus making the larger error for larger range. Under the integral, the range is assumed to be infinitesimal. We can also conclude that the probability in the continuous case is represented by the area (integral) of the probability density function p and not by its direct value at some specific point. Therefore, the probability that X takes some value in the interval $[a, b]$ is

$$P(a \leqslant X \leqslant b) = \int_a^b p(x)dx$$

In relation to this, we can introduce a concept called cumulative distribution function F which describes how probability accumulates as we go over possible values for X up to some point x

$$F(x) = P(X \leqslant x) = \int_{-\infty}^x p(t)dt$$

We can use this concept to represent the probability that X takes values in some interval $(a, b]$ (left side excluded since we defined F with $\leqslant$)

$$P(a < X \leqslant b) = P(X \leqslant b) - P(X \leqslant a) = F(b) - F(a)$$

If the random variable X has a discrete probability for some specific value a, we can then express the probability that X takes that value by starting with $F(a)$ and then subtracting all accumulated probability from the left side up to that point i.e.

$$P(X = a) = F(a) - \lim_{x \to a^-} F(x)$$

However, if the cumulative distribution function is continuous at that point, then the probability $P(X = a)$ is 0. This means that if we work under the assumption that F is continuous, it is irrelevant whether we include or exclude the edge values when finding $P(a \leqslant X \leqslant b)$. Based on this, we can conclude that the probability density function and cumulative distribution function are related like this

$$P(a \leqslant X \leqslant b) = \int_a^b p(x)dx = F(b) - F(a)$$

This gives us the following

$$F(x) = \int_{-\infty}^x p(t)dt$$

$$p(x) = \frac{dF}{dx}$$

Furthermore, since we know that the probability that X takes any value must be 1, we know that the following must hold (same holds for the discrete case)

$$\int_{-\infty}^{+\infty} p(x)dx = 1$$

$$\lim_{x \to \infty} F(x) = 1$$