Monte Carlo Integration – Part 4

In the last installment, we touched on the first rudiments of importance sample as a way of improving the Monte Carlo estimation of integrals. The key expression is that an integral $I$ can be estimated as

\[ I = \int\_a^b \, dx \, f(x) \approx \frac{1}{N} \sum\_i \frac{f(x)}{p(x)} \pm \frac{\sigma}{\sqrt{N}} \; , \]

where $p(x)$ is a known probability distribution and $N$ are the number of Monte Carlo trials (i.e. random samples).

The error in the estimate depends on two terms: the number of Monte Carlo trials $N$ and the sample standard deviation $\sigma$ pulled at random from $p(x)$. Although the error in the estimation can be made as small as possible by increasing the number of trials, the convergence is rather slow. The importance sampling strategy is to make each Monte Carlo trial count as much as possible by adapting $p(x)$ to the form of $f(x)$ such that the ratio of the two terms is a close as possible to being a constant.

The examples of how to implement this strategy from the last post were highly contrived to illustrate the method but are not representative of what is actually encountered in practice since the majority of integrands are complicated and often no able to even be represented in closed form.

In this post, we are going to extend this method a bit by considering examples where the ratio of $f(x)/p(x)$ is not a constant and finish by discussing some of the high-powered algorithms that can be plugged-in (with some effort) into complicated situations where the estimation of high-dimensional integrals is required.

Our first example will be to estimate the integral

\[ I = \int\_{0}^{10} \, dx \, e^{-x^2} \; . \]

Up to an overall scaling this integral is proportional to the [error function](https://en.wikipedia.org/wiki/Error_function) defined as

\[ erf(z) = \frac{2}{\sqrt{\pi}} \int\_0^z \, dx e^{-x^2} \; . \]

While the error function is special function that can’t be expressed as a finite number of elementary terms, it has been studied and cataloged and, like the trigonometric functions, can be thought of as being a known function with optimized ways of getting a numerical approximation. Our original integral is then

\[ I = \frac{\sqrt{pi}}{2} erf(10) \approx = 0.886226925452758 \; , \]

which we take as the exact, “truth” value.

To start we are going to estimate this integral using two different probability distributions: the uniform distribution, to provide the naïve Monte Carlo baseline for comparison, and the exponential distribution $p(x) ~ \exp(-x)$. While the exponential distribution doesn’t exactly match the integrand (i.e. $exp(-x^2)/exp(-x) \neq \textrm{constant}$), it falls off similarly

A picture containing shape

Description automatically generated

and will provide a useful experiment to set expectations of what the performance will be when $p(x)$ only approximately conforms to $f(x)$. Using the exponential distribution will also provide another change to use the variable substitution method for generating probability distributions. We’ll start with this step.