Importance Sampling for Seth

Evan Majic

December 2021

1 A Gentle Introduction

Suppose you want to compute the value of an integral. Suppose this integral is very nasty and you can't and or won't attempt to do it by hand. There are several ways of doing this, but the ways that are particularly popular and relevant to the kind of work I do are Monte Carlo methods (read: simulation methods). They involve using simulations to approximate the quantity of interest. One such method in this class is called **importance sampling**. The idea behind this is the following:

$$\int f(x)dx = \int \frac{f(x)}{g(x)} g(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} \frac{f(x_i)}{g(x_i)},$$

where $x_i \sim g$ (the xs are distributed according to g). A little more intuition on this may be needed to see why this works. Suppose that x takes on only M values. For each of those values, we have N_j observations. Then

$$\frac{1}{N}\sum_{i=1}^{N} \frac{f(x_i)}{g(x_i)} = \sum_{j=1}^{M} \frac{N_j}{N}\frac{f(x_j)}{g(x_j)}.$$

For a distribution that take on only a finite number of values $g(x_j) \approx \frac{N_j}{N}$ when we have a sufficient number of samples. Hence:

$$\sum_{j=1}^{M} \frac{N_j}{N}\frac{f(x_j)}{g(x_j)} \approx \sum_{j=1}^{M} g(x_j)\frac{f(x_j)}{g(x_j)} = \sum_{j=1}^{M} f(x_j).$$

That sort of looks like going to an integral! So, when we go from a distribution g with a finite number of values it can take to one with an infinite number, we converge to the integral.

2 A Simple Example

Say we want to estimate the integral of the following function over $[-1,1] \times [-1,1]$ (box of width 2, centered at origin)

$$f(x,y) = \begin{cases} 1; & x^2 + y^2 \leq 1 \\ 0; & x^2 + y^2 > 1 \end{cases}.$$

The value of this is obviously π. We can verify this computationally using the importance sampling method. Draw N values from the uniform distribution over $[-1,1] \times [-1,1]$, which has PDF $g(x,y) = \frac{1}{4}$, compute $f(x,y)$ and f/g and take the average. This yields:

$$4 \times \frac{1}{N}\sum_{i=1}^{N} f(x_i, y_i).$$

This should be just $4\times$ the average value of f over the square centered at the origin (probability that a random point in the square is within the circle). This is the same as

$$4 \times \frac{\pi}{4} = \pi.$$

So, our importance sampling method recovers the correct value of the integral! We can see how that works below.