2 Importance Sampling

Can we do better than the simple Monte Carlo estimator of

$$\theta = E[g(X)] = \int g(x)f(x)dx \approx \frac{1}{m}\sum_{i=1}^{m} g(X_i)$$

where the variables $X_1, \ldots, X_m$ are randomly sampled from f?

Yes!!

Goal: estimate integrals with lower variance than the simplest Monte Carlo approach.

To accomplish this, we will use *importance sampling*.

2.1 The Problem

If we are sampling an event that doesn't occur frequently, then the naive Monte Carlo estimator will have high variance.

Example 2.1 Monte Carlo integration for the standard Normal cdf. Consider estimating $\Phi(-3)$ or $\Phi(3)$.

We want to improve accuracy by causing rare events to occur more frequently than they would under the naive Monte Carlo sampling framework, thereby enabling more precise estimation.

2.2 Algorithm

Consider a density function $f(x)$ with support $\mathcal{X}$. Consider the expectation of $g(X)$,

$$\theta = E[g(X)] = \int_{\mathcal{X}} g(x)f(x)dx.$$

Let $\phi(x)$ be a density where $\phi(x) > 0$ for all $x \in \mathcal{X}$. Then the above statement can be rewritten as

An estimator of ϕ is given by the *importance sampling algorithm*:

1.

2.

For this strategy to be convenient, it must be

Example 2.2 Suppose you have a fair six-sided die. We want to estimate the probability that a single die roll will yield a 1.

2.3 Choosing ϕ

In order for the estimators to avoid excessive variability, it is important that $f(x)/\phi(x)$ is bounded and that ϕ has heavier tails than f.

Example 2.3

Example 2.4

A rare draw from ϕ with much higher density under f than under ϕ will receive a huge weight and inflate the variance of the estimate.

Strategy –

Example 2.5

The importance sampling estimator can be shown to converge to θ under the SLLN so long as the support of ϕ includes all of the support of f.

2.4 Compare to Previous Monte Carlo Approach

Common goal –

Step 1 Do some derivations.

a. Find an appropriate f and g to rewrite your integral as an expected value.

b. For **importance sampling** only,

   Find an appropriate ϕ to rewrite θ as an expectation with respect to ϕ.

Step 2 Write pseudo-code (a plan) to define estimator and set-up the algorithm.

- For **Monte Carlo integration**

  1.

  2.

- For **importance sampling**

  1.

  2.

Step 3 Program it.

2.5 Extended Example

In this example, we will estimate $\theta = \int_0^1 \frac{e^{-x}}{1+x^2} dx$ using MC integration and importance sampling with two different importance sampling distributions, ϕ.