1 Abstract

- We will evaluate the same definite integral
- We use importance sampling to improve its efficiency
- We use inverse transform for exact sampling

2 Problem

Our goal is to compute

$$\alpha = \int_0^1 h(x)dx$$

where

$$h(x) = 100 \cdot I_{(0,1/100]}(x) + 1 \cdot I_{(1/100,1)}(x).$$

Of course, the exact value shall be

$$\alpha = 1.98.$$

Pretended not to know the exact value, we have used OMC with exact sampling of uniform random variable, denoted by omc_integral(n).

Next, we are going to improve the efficiency of omc_integral(n) by using importance sampling. We also extend our skill on exact sampling by using inverse transform.

3 Analysis

3.1 Importance sampling

Recall that, to estimate the above integral α, we use the uniform random variable X, whose density is $p(x) = I_{(0,1)}(x)$, and write

$$\alpha = \mathbb{E}[h(X)|X \sim p] = \int_0^1 h(x)p(x)dx.$$

Naturally, one can sample iid uniform random numbers by computer, denoted by

$$\{\ X_i \sim p : i = 1, 2, \ldots, n\},$$

then taking their average for its approximation of α, i.e.

$$\hat{\alpha}_n = \frac{1}{n}\sum_{i=1}^{n} h(X_i).$$

Example 1 *Compute MSE of $\hat{\alpha}_n$.*

Solution. Since it is unbiased, MSE is the same as Variance of $\hat{\alpha}_n$, and it is again equal to $1/n$ of

$$Var[h(X)|X \sim p].$$

Therefore, it is $\frac{100.99}{n}$. □

IS considers, with a smart choice of a pdf p_1,

$$\alpha = \int_0^1 h(x)\frac{p(x)}{p_1(x)}p_1(x)dx = \mathbb{E}\Big[h(X)\frac{p(X)}{p_1(X)}\Big|X \sim p_1(x)\Big]$$

Since we observe that the interval $(0, 1/100)$ is much more *important* than $(1/100, 1)$, our choice of p_1 is the following:

$$p_1(x) = \frac{1}{C}(2 \cdot I_{(0,1/100]}(x) + 1 \cdot I_{(1/100,1)}(x)),$$

where $C = 101/100$ is the normalizing constant to make p_1 to be a valid pdf.

Pseudocode is_integral(n):

- Generate iid p_1 samples, denoted by

$$\{X_i : i = 1, 2, \ldots, n\}.$$

- Compute the average of the integrand h adjusted by likelyhood ratio (also referred to radon-nikodym derivative) p/p_1, i.e.

$$\bar{\alpha}_n = \frac{1}{n}\sum_{i=1}^{n} h(X_i) \cdot \frac{p(X_i)}{p_1(X_i)}.$$

Example 2 *Prove that MSE of $\hat{\alpha}_n$ is $51.4999/n$.*

3.2 Inverse transform method

To implement the is_integral(n), we shall generate p_1 samples. But this is not directly available by python. Inverse transform method provides exact sampling as long as the inverse of CDF is explicitly available. Its theoretic basis is given next.

Proposition 1 *Suppose X has its CDF F and F^{-1} exists, then $F^{-1}(U) \sim X$, where $U \sim U(0, 1)$.*

Proof:

$$\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x) = F(x).$$

□

Pseudocode it_sampling(F^{-1}, n):

- Generate iid $U(0, 1)$ random variables

$$\{Y_i : i = 1, \ldots, n\}.$$

- Compute

$$\{X_i = F^{-1}(Y_i) : i = 1, \ldots, n\}.$$

4 Exercises

1. Find F_1, the cdf of p_1.
2. Find F_1^{-1}.
3. Implement it_sampling()
4. Implement is_integral()
5. Demonstrate the convergence rate of is_integral()
6. Could you find a pdf p_2 better than p_1?