Monte Carlo Integration

Christoph Heindl
https://github.com/cheind/monte-carlo-integration

July 12, 2022

1 Introduction

The purpose of this document is to present Monte Carlo integration, a method to approximate the value of the definite integral

$$\int_a^b g(x)\,dx. \tag{1}$$

In Monte Carlo, we transform the above integral into an expectation over a random variable. We then approximate the value of the expectation using an unbiased estimator. Assume x is random $x \sim X$ with PDF $f_X(x)$. Multiply and divide by $f_X(x)$ to get

$$\int_a^b \frac{g(x)}{f_X(x)} f_X(x)\,dx.$$

Assume that $f_X(x)$ is zero everywhere but in $[a, b]$, so we can change the bounds of integration to

$$\int_{-\infty}^{\infty} \frac{g(x)}{f_X(x)} f_X(x)\,dx. \tag{2}$$

Using LOTUS[1] we identify Equation 2 to be the expectation

$$\mathbb{E}\left[\frac{g(X)}{f_X(X)}\right]. \tag{3}$$

An unbiased expectation estimator [TAB17] is given by

$$\frac{1}{N}\sum_{i=1}^{N} \frac{g(\hat{x}_i)}{f_X(\hat{x}_i)}, \tag{4}$$

where $\hat{x}_i \overset{iid}{\sim} X$ is the i-th random value. As N increases our approximation approaches the true value of the integral. To summarize

$$\int_a^b g(x)\,dx = \mathbb{E}\left[\frac{g(X)}{f_X(X)}\right] \approx \frac{1}{N}\sum_{i=1}^{N} \frac{g(\hat{x}_i)}{f_X(\hat{x}_i)}. \tag{5}$$

Further, if we assume X is distributed uniformly on $[a, b]$ with $f_X(x) = \frac{1}{b-a}$ we get

$$\int_a^b g(x)\,dx \approx \frac{1}{N}\sum_{i=1}^{N} g(\hat{x}_i)(b-a). \tag{6}$$

Geometrically, this equals taking the average area of rectangles having length $b - a$ and height $g(\hat{x}_i)$.

[1] https://en.wikipedia.org/wiki/Law_of_the_unconscious_statistician

2 Multivariate Case

For k-dimensional integrals a straightforward extension of Equation 5 exists. Consider a double integral of the form

$$\int_a^b \int_c^d g(x, y)\, dydx.$$

Let X and Y be random variables whose joint PDF is $f_{X,Y}$. For similar reasons as in the univariate case, we identify the above integral as the expectation of

$$\mathbb{E}_{X,Y}\left[\frac{g(X, Y)}{f_{X,Y}(X, Y)}\right]. \tag{7}$$

An unbiased estimator for the exception in Equation 8 is given by

$$\frac{1}{N}\sum_{i=1}^{N}\frac{g(\hat{x}_i, \hat{y}_i)}{f_{X,Y}(\hat{x}_i, \hat{y}_i)}. \tag{8}$$

If X and Y are independent, i.e. $f_{X,Y} = f_X(x)f_Y(y)$ and the distribution is uniform we arrive at

$$\frac{(b-a)(d-c)}{N}\sum_{i=1}^{N} g(\hat{x}_i, \hat{y}_i), \tag{9}$$

which is a natural extension of Equation 6.

2.1 Integrals with Dependent Limits and Indicator Functions

If the limits of integration dependent on a integration variable, result Equation 8 is not straightforward to apply. We simply cannot expect $f_{X,Y} = f_X(x)f_Y(y)$ to hold. Consider the integral

$$\int_a^b \int_{l(x)}^{u(x)} g(x, y)\, dydx,$$

where the limits of the inner integral are functions of the outer variable. If we manage to rewrite the limits to be unconditional, we are in good shape to apply Equation 8.

Let c be the minimum possible value of $l(x)$ and likewise d be the maximum possible value of $u(x)$. Next, introduce an indicator function to determine if y is within the original bounds

$$\mathbb{I}_{x,y} := \begin{cases} 1 & y \in [l(x), u(x)] \\ 0 & \text{else} \end{cases},$$

and rewrite the integral as

$$\int_a^b \int_c^d g(x, y)\mathbb{I}_{x,y}\, dydx.$$

The adapted version of Equation 8 is

$$\frac{1}{N}\sum_{i=1}^{N}\frac{g(\hat{x}_i, \hat{y}_i)\mathbb{I}_{\hat{x}_i,\hat{y}_i}}{f_{X,Y}(\hat{x}_i, \hat{y}_i)}. \tag{10}$$

If X and Y are independent, i.e. $f_{X,Y} = f_X(x)f_Y(y)$ and the distribution is uniform we arrive at

$$\frac{(b-a)(d-c)}{N}\sum_{i=1}^{N} g(\hat{x}_i, \hat{y}_i)\mathbb{I}_{\hat{x}_i,\hat{y}_i}. \tag{11}$$

Note, the above frames the problem of computing the area of circle as the problem of iterating a rectangle and summing up all square patches for which the indicator function yields 1. This is less efficient than just accessing the squares that make up the circle, but leads to much simpler algorithms and estimators. In general, given a shape $\mathcal{S}$, an associated indicator function $\mathbb{I}$ that determines whether a point $\mathbf{b}$ is inside $\mathcal{S}$, and a bounding rectangle $\mathcal{B} \in \mathcal{R}^K$ with known volume $V_{\mathcal{B}}$, we approximate volume $V_{\mathcal{S}}$ by

$$\begin{aligned} V_{\mathcal{S}} &= \int_{\mathcal{B}} \cdots \int \mathbb{I}(b^1 \dots b^k)\, db^1 \dots db^k \\ &\approx \frac{V_{\mathcal{B}}}{N} \sum_{i=1}^{N} \mathbb{I}(\hat{b}_i^1 \dots \hat{b}_i^k) \end{aligned} \tag{12}$$

using iid uniformly distributed samples $\{\mathbf{b}_i\}_{i<=N} \in \mathcal{B}$. In particular, for the case $K = 2$, the area of $A_{\mathcal{S}}$ is approximated by

$$\begin{aligned} A_{\mathcal{S}} &= \iint_{\mathcal{B}} \mathbb{I}(x, y)\, dydx \\ &\approx \frac{A_{\mathcal{B}}}{N} \sum_{i=1}^{N} \mathbb{I}(\hat{x}_i, \hat{y}_i) = \frac{N_{\text{inside}}}{N} A_{\mathcal{B}}, \end{aligned} \tag{13}$$

where N_{inside} is the number of inside points and A_B is the area of the bounding region. In other words, to approximate the area $A_{\mathcal{S}}$, multiply the ratio of inside to total samples with the volume of the bounding rectangle.

3 Example - Circle Area

3.1 Univariate Approach

The area of a unit circle can be written as the integral of its circumference

$$A_{\text{circle}} = \int_0^1 2\pi r\, dr.$$

To estimate this integral we proceed as follows

1. Generate N iid samples from uniform $[0, 1]$ distribution.
2. Compute $g(\hat{x}_i) = 2\pi\hat{x}_i$ for every sample.
3. Use Equation 6 to estimate the integral.

3.2 Multivariate Approach

In this section still find the area of the circle, but consider a different integral. Consider a quadrant of a unit circle as shown in Figure 1. The area of this quadrant can be written as the following double integral with dependent bounds

$$A_q = \int_0^1 \int_0^{\sqrt{1-x^2}} dydx.$$

Conceptually, this sums the areas of all squares with side-lengths dx,dy within the circle area. While x runs freely from zero to one, the upper limit of y depends on the current x-position via $x^2 + y^2 = 1$ (solve for y as a function of x). Following the approach in Section 2.1, we let our indicator function be

$$\mathbb{I}_{x,y} := \begin{cases} 1 & x^2 + y^2 \le 1 \\ 0 & \text{else} \end{cases}.$$

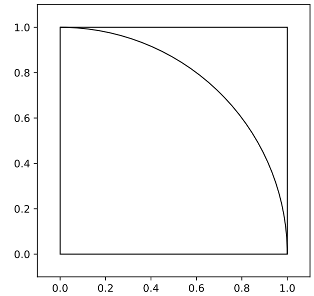

Figure 1: Quadrant of circle with bounding square.

The upper integral limit for y is maximized when $x = 0 \implies y = 1$ and we therefore set $d = 1$ which gives

$$A_q = \int_0^1 \int_0^1 \mathbb{I}_{x,y} \, dy dx.$$

Assuming a uniform distribution, we can use Equation 11 to get our integral estimator

$$A_q \approx \frac{(1-0)(1-0)}{N} \sum_{i=1}^{N} \mathbb{I}_{\hat{x}_i, \hat{y}_i} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{I}_{\hat{x}_i, \hat{y}_i}.$$

By symmetry the area of the circle is given by $A_{\text{circle}} = 4A_q$.

4 Remarks

Up to this point we made extensively use of the uniform distribution for $f_X(x)$, because its easy to visualize and leads to simple geometric concepts. Keep in mind, though, that other distributions are suited equally well as long the PDF is zero outside the integration bounds $[a, b]$. In fact the uniform might be a bad choice (in terms of approximation error and convergence) if most of the 'mass' of $g(x)$ is concentrated in few sub-regions of $[a, b]$.

5 Todo

Just a few reminders for myself of what's missing:

- Prove that the Monte Carlo estimator is an unbiased estimator.
- Add importance sampling to speed up convergence.
- Provide an estimate for the approximation error.

References

[TAB17] MARCO. TABOGA. *Lectures on probability theory and mathematical statistics*. Createspace, 2017.