Computer Physics Communications 208 (2016) 29–34

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

First- and second-order error estimates in Monte Carlo integration

R. Bakx, R.H.P. Kleiss *, F. Versteegen

Radboud University Nijmegen, Institute for Mathematics, Astrophysics and Particle Physics, Heyendaalseweg 135, NL-6525 AJ Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 2 September 2015
Received in revised form
18 April 2016
Accepted 11 July 2016
Available online 17 July 2016

Keywords:
Monte Carlo integration
Error estimates
Central limit theorem
Chan–Golub algorithm

ABSTRACT

In Monte Carlo integration an accurate and reliable determination of the numerical integration error is essential. We point out the need for an independent estimate of the error on this error, for which we present an unbiased estimator. In contrast to the usual (first-order) error estimator, this second-order estimator can be shown to be not necessarily positive in an actual Monte Carlo computation. We propose an alternative and indicate how this can be computed in linear time without risk of large rounding errors. In addition, we comment on the relatively very slow convergence of the second-order error estimate.

1. Monte Carlo integration and its errors

It does not need to be stressed that in numerical integration, including Monte Carlo (MC) integration [1], determination or estimate of the integration error made is essential. The Central Limit Theorem (CLT) practically ensures that if the number N of MC points is sufficiently large the numerical value of the MC integral – itself a stochastic variable – will have a Gaussian distribution around the true integral value, with a standard deviation that can itself also be estimated: this is the *first-order error*. The results of MC integrations are therefore usually reported as

"result" $\pm$ "error"

with the understanding that the "error" value quoted is the Gaussian's standard deviation. In this way one can, for instance, assign confidence levels when comparing the integration result with a measurement. However, since the Gaussian distribution is quite steep, a modest change in the value of the error can change the confidence levels considerably. It is therefore preferable to also have a *second-order error* that estimates how well the first-order error was computed. The better way to report the result of a MC integration is then

"result" $\pm$ ("first-order error" $\pm$ "second-order error") .

A first attempt to implement such a method was presented in [2]. However, in that paper no explicit form of the second-order error estimator was presented, nor were its numerical stability properties and its convergence behaviour discussed: also it was (wrongly) stated that the second-order error was the square root of the estimator, while it ought to be the fourth root. The present paper addresses and corrects these issues. In what follows we shall arrive at an estimator for the second-order error that, like the first-order one, can be evaluated in linear time *i.e.* at essentially no extra CPU cost. We shall also discuss several of its numerical aspects, and suggest an improvement.

2. Error estimators

We will start by defining some mathematical tools. We consider an integral over an integration region Γ of an integrand $f(x)$, with $x \in \Gamma$. We have at our disposal a set of MC integration points x_j , $j = 1, 2, \ldots, N$, assumed to be iid (Independent, Identically Distributed) with a probability distribution $P(x)$ in Γ. We define

$$J_p = \int_\Gamma dx\, P(x)\; w(x)^p, \quad w(x) = \frac{f(x)}{P(x)}, \tag{1}$$

so that $J_1 = \int dx\; f(x)$, the integral we want to compute. The numbers $w_j \equiv w(x_j)$ are called the *weights* of the points. We see that J_p is nothing but the expectation value of $w(x)^p$:

$$\left\langle w^p \right\rangle = J_p. \tag{2}$$

Furthermore, we define the following multiple sums:

$$S_{p_1,p_2,\ldots,p_k} = \sum_{j_{1,2,\ldots,k}=1}^{N} w_{j_1}{}^{p_1}\, w_{j_2}{}^{p_2} \cdots w_{j_k}{}^{p_k} \tag{3}$$

* Corresponding author.
E-mail addresses: renz.bakx@student.ru.nl (R. Bakx), R.Kleiss@science.ru.nl (R.H.P. Kleiss), f.versteegen@student.ru.nl (F. Versteegen).

http://dx.doi.org/10.1016/j.cpc.2016.07.021
0010-4655/

with the condition that the indices $j_{1,2,\dots,k}$ are all *different*. As an example, the sum $S_{1,1}$ does not contain N^2 but $N^{\underline{2}} = N^2 - N$ terms. The *falling powers* are defined by

$$N^{\underline{p}} = N!/(N-p)! = N(N-1)(N-2)\cdots(N-p+1). \tag{4}$$

The simple sums S_p can be evaluated in linear time (that is, using N additions), but a multiple sum $S_{p_1,\dots,p_k}$ needs time of the order N^k. In calculating estimators we therefore want to use only simple sums. On the other hand, only the multiple sums have a simple expectation value:

$$\left\langle S_{p_1,p_2,\dots,p_k}\right\rangle = N^{\underline{k}} J_{p_1} J_{p_2}\cdots J_{p_k}. \tag{5}$$

We can relate simple and multiple sums to one another by the following obvious rule:

$$S_{p_1,p_2,\dots,p_k} S_q = S_{p_1+q,p_2,\dots,p_k} + S_{p_1,p_2+q,\dots,p_k} + \cdots + S_{p_1,p_2,\dots,p_k+q} + S_{p_1,p_2,\dots,p_k,q}. \tag{6}$$

We are now ready to construct the various estimators, starting with the well-known MC formulæ for clarity. For the integral we have

$$E_1 = \frac{1}{N} S_1, \tag{7}$$

since $\langle E_1\rangle = J_1$; moreover we see that this estimator is unbiased. For the variance of E_1 we have

$$\left\langle {E_1}^2\right\rangle - \langle E_1\rangle^2 = \frac{1}{N^2}\left\langle S_2 + S_{1,1}\right\rangle - {J_1}^2 = \frac{1}{N}\left(J_2 - {J_1}^2\right) = \frac{1}{N^2}\langle S_2\rangle - \frac{1}{N^{\underline{2}}N}\left\langle S_{1,1}\right\rangle \tag{8}$$

so that the appropriate estimator is

$$E_2 = \frac{S_2}{N^2} - \frac{S_{1,1}}{N^{\underline{2}}N} = \frac{1}{N^{\underline{2}}N}\Sigma_2, \quad \Sigma_2 = N S_2 - {S_1}^2. \tag{9}$$

The latter form is more suited to computation since it can be evaluated in linear time. From Eq. (8) we see that the first-order error, defined as ${E_2}^{1/2}$ decreases as $N^{-1/2}$, as is of course very well known. Moreover, the expected error is defined for all functions that are quadratically integrable, as is equally well known.

The *second-order error* should have as *its* expectation value the variance of E_2, which by the same methods as above can be shown to be

$$\left\langle {E_2}^2\right\rangle - \langle E_2\rangle^2 = \frac{1}{N^3}\left(J_4 - 4J_3J_1 + 3{J_2}^2 - 4\left(J_2 - {J_1}^2\right)^2\right) + \frac{2}{N^{\underline{2}}N^2}\left(J_2 - {J_1}^2\right)^2. \tag{10}$$

We see that the second-order error, defined as ${E_4}^{1/4}$ decreases, for large N, as $N^{-3/4}$. Moreover we see that the second-order error is only meaningful for integrands that are at least *quartically* integrable. The appropriate unbiased estimator with the correct expectation value is

$$E_4 = \frac{1}{N^{\underline{4}}N^3}\left(N^{\underline{2}}\Sigma_4 - 4{\Sigma_2}^2\right) + \frac{2}{N^{\underline{4}}N^{\underline{2}}N^2}{\Sigma_2}^2,$$
$$\Sigma_4 = N S_4 - 4 S_3 S_1 + 3 {S_2}^2. \tag{11}$$

An important observation here concerns the asymptotic behaviour of the relative errors. Whereas the relative first-order error, *i.e.* the ratio ${E_2}^{1/2}/E_1$, goes as $N^{-1/2}$ according to the 'standard' behaviour in MC, the relative second-order error ${E_4}^{1/4}/{E_2}^{1/2}$ only decreases as fast as $N^{-1/4}$. It will therefore take much longer for the error to be well-determined than for the integral itself.[1]

[1] Note that the relative errors as defined here are the *dimensionless* ratios, the only meaningful measures of performance of the computation.

A final point is in order. By the CLT we know that the distribution of E_1 in an ensemble of MC computations is normally distributed, which tells us the *meaning* of E_2, as discussed above. Since E_2 is not computed as a simple average, *its* distribution is not governed by the *same* CLT. Nevertheless, as is shown in the Appendix a good case can be made for it being also approximately normally distributed, so that the relation between E_4 and the confidence levels of E_2 can be treated in the usual manner. Below, we shall illustrate this with several examples.

3. Positivity and numerical stability

In principle, Eqs. (7), (9) and (11) are what is necessary to obtain the integral and its first- and second-order errors. However, a number of considerations must modify this picture. In the first place, the issue of positivity. Writing $w(x) = J_1 + u(x)$ so that $\int dx\, P(x)\, u(x) = 0$, we have

$$J_2 - {J_1}^2 = \int dx\; P(x)\, u(x)^2,$$
$$J_4 - 4J_3J_1 + 3{J_2}^2 = \int dx\; P(x)\, u(x)^4 + 3\left(\int dx\; P(x)\, u(x)^2\right)^2,$$
$$J_4 - 4J_3J_1 + 3{J_2}^2 - 4\left(J_2 - {J_1}^2\right)^2 = \frac{1}{2}\int dx\, dy\; P(x)\, P(y)\, (u(x)^2 - u(y)^2)^2, \tag{12}$$

so that the *expectation values* of $E_{2,4}$ are positive, as they should. In addition, since with the notation $W_j = E_1 + u_j$ the Σ_2 can be written as

$$\Sigma_2 = \frac{1}{2}\sum_{j,k}\left(u_j - u_k\right)^2, \tag{13}$$

also E_2 itself is strictly nonnegative in any actual MC calculation. For E_4 this does not hold, however. A counterexample can be constructed as follows. Let us assume that the MC weights w_j take on only the values 0 and 1, and that $E_1 = Nb$, $b \in [0, 1]$. We then have

$$\Sigma_2 = \Sigma_4 = N^2 a, \quad a = b - b^2 \in [0, 1/4]. \tag{14}$$

The value of E_4 now comes out as

$$E_4 = \frac{1}{N^{\underline{4}}}\left(\frac{N^{\underline{2}}}{N}a - \frac{4N^3 - 6N^2}{N^{\underline{2}}}a^2\right), \tag{15}$$

which is actually *negative* for

$$a > \frac{(N-1)^2}{N(4N-6)} = \frac{1}{4} - \frac{N-2}{2N(4N-6)}. \tag{16}$$

Although by small margin (surprisingly, in this counterexample, for $b \approx 1/2$), the positivity of E_4 cannot be guaranteed, so that ${E_4}^{1/4}$ may be undefined. As an improvement on this situation we propose to abandon the estimator E_4 in favour of

$$\hat{E}_4 = \frac{1}{N^{\underline{4}}N^3}\left(N^2\Sigma_4 - 4{\Sigma_2}^2\right). \tag{17}$$

This estimator has a slight (order $1/N$) bias, which ought to be acceptable since we are dealing with only the second-order error here; its advantage is that, since

$$N^2\Sigma_4 - 4{\Sigma_2}^2 = \frac{N^2}{2}\sum_{j,k}\left({u_j}^2 - {u_k}^2\right)^2, \tag{18}$$

it always evaluates to a nonnegative number.

The second issue is that of numerical stability. It is well known that already the evaluation of Σ_2 involves large cancellations