11.4 ■ MONTE CARLO ERROR ANALYSIS

Both the classical numerical integration methods and the Monte Carlo methods yield approximate answers whose accuracy depends on the number of intervals or on the number of samples, respectively. So far, we have used our knowledge of the exact values of various integrals to determine that the error in the Monte Carlo methods approaches zero as $n^{-1/2}$ for large n, where n is the number of samples. In the following, we will learn how to estimate the error when the exact answer is unknown. Our main result is that the $n^{-1/2}$ dependence of the error is a general result and is independent of the nature of the integrand and, most importantly, independent of the number of dimensions.

As before, we first determine the error for an explicit example. Consider the Monte Carlo evaluation of the integral of $f(x) = 4\sqrt{1-x^2}$ in the interval [0, 1] (see Problem 11.7). Our result for a particular sequence of $n = 10^4$ random numbers using the sample mean method is $F_n = 3.1489$. How does this result for F_n compare with your result found in Problem 11.7 for the same value of n? By comparing F_n to the exact result of $F = \pi \approx 3.1416$, we find that the error associated with $n = 10^4$ samples is approximately 0.0073. How do we know if $n = 10^4$ samples are sufficient to achieve the desired accuracy? We cannot answer this question definitively because if the actual error were known, we could correct F_n by the required amount and obtain F. The best we can do is to calculate the *probability* that the true value F is within a certain range centered about F_n.

We know that if the integrand was a constant, then the error would be zero; that is, F_n would equal F for any n. This limiting behavior suggests that a possible measure of the error is the *sample variance* $\tilde{\sigma}^2$ defined by

$$\tilde{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}[f(x_i) - \langle f\rangle]^2, \tag{11.19}$$

where

$$\langle f\rangle = \frac{1}{n}\sum_{i=1}^{n} f(x_i). \tag{11.20}$$

The reason for the factor of $1/(n-1)$ in (11.19) rather than $1/n$ is similar to the reason for the expression $1/\sqrt{n-2}$ in the error estimates of the least squares fits (see (7.43)). To compute $\tilde{\sigma}^2$, we need to use n samples to compute the mean $\langle f\rangle$, and, loosely speaking, we have only $n-1$ independent samples remaining to calculate $\tilde{\sigma}^2$. Because we will always be considering values of $n \gg 1$, we will replace $\tilde{\sigma}^2$ by the *variance* σ^2, which is given by

$$\sigma^2 = \langle f^2\rangle - \langle f\rangle^2, \tag{11.21}$$

where

$$\langle f^2\rangle = \frac{1}{n}\sum_{i=1}^{n}[f(x_i)]^2. \tag{11.22}$$

For our example and the same sequence of random numbers that we used to obtain $F_n = 3.1489$, we obtain the standard deviation $\sigma = 0.8850$. Because this value of σ is two orders of magnitude larger than the actual error, we conclude that σ is not a direct measure of the error.

Table 11.2 Examples of Monte Carlo measurements of the mean value of $f(x) = 4\sqrt{1-x^2}$ in the interval [0, 1]. The actual error is given by the difference $|F_n - \pi|$. The standard deviation σ is estimated using (11.21).

n	F_n	Actual Error	σ	$\sigma/\sqrt{n}$
10^2	3.0692	0.0724	0.8550	0.0855
10^3	3.1704	0.0288	0.8790	0.0270
10^4	3.1489	0.0073	0.8850	0.0089

Another clue to finding an appropriate measure of the error can be found by increasing n and seeing how the actual error decreases as n increases. In Table 11.2 we see that as n is increased from 10^2 to 10^4, the actual error decreased by a factor of approximately 10, that is, as $\sim 1/n^{1/2}$. We also see that the actual error is approximately given by $\sigma/\sqrt{n}$. In Appendix 11B we show that the *standard error of the means* σ_m is given by

$$\sigma_m = \frac{\sigma}{\sqrt{n-1}} \tag{11.23a}$$

$$\approx \frac{\sigma}{\sqrt{n}}. \tag{11.23b}$$

The interpretation of σ_m is that if we make many independent measurements of F_n, each with n data points, then the *probable error* associated with any single measurement is σ_m. The more precise interpretation of σ_m is that F_n, our estimate for the mean, has a 68% chance of being within σ_m of the "true" mean and a 97% chance of being within $2\sigma_m$. This interpretation assumes a Gaussian distribution of the various measurements.

The quantity F_n is an estimate of the average value of the data points. As we increase n, the number of data points, we do not expect our estimate of the mean to change much. What changes, as we increase n, is our *confidence* in our estimate of the mean. Similar considerations hold for our estimate of σ, which is why σ is not a direct measure of the error.

The error estimate in (11.23) assumes that the data points are independent of each other. However, in many situations the data is correlated, and we have to be careful about how we estimate the error. For example, suppose that instead of choosing n random values for x, we instead start with a particular value x_0 and then randomly add increments such that the ith value of x is given by $x_i = x_{i-1} + (2r-1)\delta$, where r is uniformly distributed between 0 and 1, and $\delta = 0.01$. Clearly, the x_i are now correlated. We can still obtain an estimate for the integral, but we cannot use $\sigma/\sqrt{n}$ as the estimate for the error because this estimate would be smaller than the actual error. However, we expect that two data points, x_i and x_j, will become uncorrelated if $|j-i|$ is sufficiently large. How can we tell when $|j-i|$ is sufficiently large? One way is to group the data by averaging over m data points. We take $f_1^{(m)}$ to be the average of the first m values of $f(x_i)$, $f_2^{(m)}$ to be the average of the next m values, and so forth. Then we compute $\sigma_s/\sqrt{s}$, where $s = n/m$ is the number of $f_i^{(m)}$ data points, each of which is an average over m of the original data points, and σ_s is the standard deviation of the s data points. We do this grouping for different values of m (and s) and