NERS 544 Monte Carlo Methods Lecture 3

Error Estimation (Chapter 5)

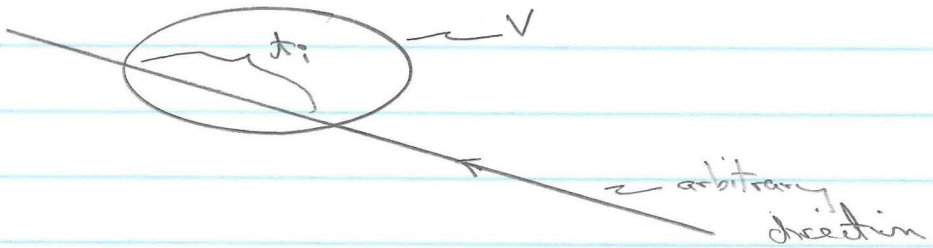
Ideally ⇒ cross sections are "perfect"
sampling methods are "exact"

MC produces results that have a variance that depends on N, the number of histories.

A "history" is a counter for each independent invocation, typically, in our field, the sampling of a single particle from a source

Each history may use many random numbers and arithmetical operations.

For concreteness, let us imagine that we have some volume, V, that is impinged upon by an external, isotropic source.



There is no scattering inside or outside V

For convex volume (any two points on the surface can be connected by a straight line that is contained within the volume, it is known that

$$\bar{t} = \frac{4V}{S} \qquad V = \text{volume}(L^3), \quad S = \text{surface area}(L^2)$$

It is a mathematically interesting question to understand what the underlying probability distribution is, namely

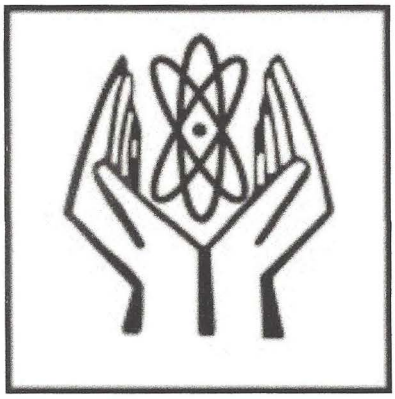
$$\bar{t} = \int_0^{t_{max}} dt\, t\, p(t) \qquad (1)$$

where $t=0$ corresponds to a grazing trajectory & t_{max} is the longest possible chord length.

For closed bodies, it is clear that $\langle t^n \rangle$ is finite $\forall\ n \in \mathbb{N}$ & $n \geq 0$

The probability distributions are known in only a few special case.

See: (next page)



So, with little math guidance we undertake a Monte Carlo experiment

Procedure (Direct Error estimation)

1) Estimate the mean value of t, $\langle t \rangle$

cf (5.12)

$$\langle t \rangle = \frac{1}{N} \sum_{i=1}^{N} t_i$$

(N: sum over histories; t_i: chord length of ith history)

Can we trust this number?

If μ, the true mean of t exists, namely

$$\mu = \int_0^{t_{max}} dt\, t\, p(t)$$

then

$$\lim_{N \to \infty} P\left\{ \left| \frac{1}{N} \sum_{i=1}^{N} t_i - \mu \right| \right\} \to 0$$

i.e. in the large N limit, $\langle t \rangle \to \mu$

This is the "law of large numbers"

This law does not give a sense of how good $\langle t \rangle$ is for finite N.

A much stronger statement can be made if σ^2 exists

2) Estimate the variance

$$s_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \langle x \rangle)^2$$

$$= \frac{1}{N-1} \sum_{i=1}^{N} (x_i^2 - \langle x \rangle^2)$$

Note s_x^2 is an "estimated" variance
σ^2 is the true variance

3) Calculate the estimated variance of the mean

$$s_{\bar{x}}^2 = \frac{s_x^2}{N}$$

Quote your result as

$\langle x \rangle \pm s_{\bar{x}}$ } 67% confidence level (traditional)

or $\langle x \rangle \pm 2 s_{\bar{x}}$ } 95% confidence level

How good is this?

Central limit theorem

$$P\left(\frac{\langle x\rangle - \mu}{\sigma/\sqrt{N}} < \beta\right) \Rightarrow \Phi(\beta)$$

$\Rightarrow$ For large N, $\langle x\rangle$ follows a normal distribution with a Gaussian width $\frac{\sigma}{\sqrt{N}}$

Try this with the "coin flip" pdf

$$p(x) = \frac{1}{2}[\delta(x) + \delta(x-1)] \quad \forall \; |x| < \infty$$

Binary statistics

$$s_x^2 = \frac{1}{N-1} \sum_{i=1}^{N} \left(x_i^2 - \langle x \rangle^2 \right)$$

If $x_i = 0$ or 1 only

$$s_x^2 = \frac{1}{N-1} \left(\underbrace{\sum_{i=1}^{N} (x_i^2)}_{= N\langle x \rangle} - N\langle x \rangle^2 \right)$$

$$= \frac{N}{N-1} \left(\langle x \rangle - \langle x \rangle^2 \right)$$

$$s_x^2 = \frac{\langle x \rangle (1 - \langle x \rangle)}{N-1}$$

- Estimate the variance associated with the distribution of the x_i:

$$s_x^2 = \frac{1}{N-1}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{N-1}\sum_{i=1}^{n}(x_i^2 - \overline{x}^2) \tag{5.13}$$

- The estimated variance of $\overline{x}$ is the standard variance of the mean:

$$s_{\overline{x}}^2 = \frac{s_x^2}{N} \tag{5.14}$$

  It is the error in $\overline{x}$ we are seeking, not the "spread" of the distribution of the x_i.

- Report the final result as $x = \overline{x} \pm s_{\overline{x}}$.

Remarks:

The true mean and variances are not available to us, however. We must estimate them. The estimated mean $\overline{x}$ calculated in eq. 5.12 is an estimate for the true mean μ and the estimated variance s_x^2 calculated in eq. 5.13 is an estimate for the true variance σ_x^2. The appearance of the $N-1$ in the denominator in the expression for the estimated variance for the mean implicit in eq. 5.14 is often introduced as a "degrees of freedom" arguments. However, it can be derived by considering the difference between "sample" variance and its relation to the true variance. A derivation is given at the end of this chapter.

5.2 Batch statistics error estimation

In many cases, the estimation of means and variances using the methods described in the previous section is not feasible. A score may be a complicated object or there may be many geometrical volume elements to consider. It may require a lot of effort to stop after each history and compute the x_i's and x_i^2's. The "direct error estimation" should be considered first and if it is not feasible, then an alternative that is nearly as good is "batch statistics error estimation". We present a "cook book" recipe as before.

- Split the N histories into n statistical batches of N/n histories each. Note that n must be "large enough" as discussed earlier. A standard choice is $n = 30$. The accumulated quantity for each of these batches is called $x_j = \sum_{i=1}^{n} x_i$ for the j'th statistical batch.

- Calculate the mean value of x:

$$\overline{x} = \frac{1}{N}\sum_{j=1}^{n} x_j \tag{5.15}$$

- Estimate the variance associated with the distribution of the x_i:

$$s_x^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j - \overline{x})^2 = \frac{1}{n-1}\sum_{j=1}^{n}(x_j^2 - \overline{x}^2) \tag{5.16}$$

- The estimated variance of $\overline{x}$ is the standard variance of the mean:

$$s_{\overline{x}}^2 = \frac{s_x^2}{n} \tag{5.17}$$

- Report the final result as $x = \overline{x} \pm s_{\overline{x}}$.

Remarks:

We use eqs. 5.15–5.17 with n fixed (at say, 30) because it gives a reasonable estimate of the error in $\overline{x}$. Any large number will do, as long as we are within the range of applicability of the Central limit theorem. In reality, decisions based upon the error in $\overline{x}$ are usually subjective in nature. There is some evidence that the calculated statistic depends weakly on the choice of n. Therefore, it is important to report how your statistics were done when you publish your Monte Carlo results.

5.3 Combining errors of independent runs

For m independent Monte Carlo runs, it is easy to derive the following relation:

$$\overline{x} = \sum_{k=1}^{m} \left(\frac{N_k}{N}\right) \overline{x}_k \tag{5.18}$$

where $\overline{x}_k$ is the value of $\overline{x}$ for the k^{th} run and N_k is the number of histories in the k^{th} run. The total number of histories is given by:

$$N = \sum_{k=1}^{m} N_k \tag{5.19}$$

Then, assuming 1^{st}-order propagation of independent errors, it is also easy to derive:

$$s_{\overline{x}}^2 = \sum_{k=1}^{m} \left(\frac{N_k}{N}\right)^2 s_{\overline{x}_k}^2 \tag{5.20}$$

where $s_{\overline{x}_k}^2$ is the estimated variance in $\overline{x}_k$.

Example: For m = 2:

$$\overline{x} = \left(\frac{N_1}{N}\right) \overline{x}_1 + \left(\frac{N_2}{N}\right) \overline{x}_2 \tag{5.21}$$

$$N = N_1 + N_2 \tag{5.22}$$

$$s_{\overline{x}} = \sqrt{\left(\frac{N_1}{N}\right)^2 s_{\overline{x}_1}^2 + \left(\frac{N_2}{N}\right)^2 s_{\overline{x}_2}^2} \tag{5.23}$$

Remarks:

This method of combining errors effectively increases the value of n, the number of statistical batches used in the calculation. In view of the fact that the calculated statistics are thought to depend weakly on n, it is preferable (but only marginally so for the sake of consistency) to combine the x_i's (the raw data) into the standard number of statistical batches. This is easy to do by initialising the data arrays to the results of the previous run before the start of a new run.

5.4 Error estimation for binary scoring

In the special case where the scoring quantity is binary (*e.g.* a particle enters a detector's sensitive region or not, *e.g.* a particle backscatters or not) there is a significant simplification. Consider N events where the scoring variable x_i is either 0 or 1. The mean is calculated according to eq. 5.12. Starting from eqs. 5.12–5.14 it is easy to show that:

$$s_{\overline{x}}^2 = \frac{\overline{x}(1-\overline{x})}{N-1}. \tag{5.24}$$

Hence, knowing the mean value gives one knowledge of its estimated variance! One is not required to gather these events in statistical bins.

Combining the results of independent runs is particularly simple. The mean is calculated from eq. 5.12 but the use of the variance combination relation, eq. 5.20, is inaccurate unless the N_j are all "large" ($N_j \gg 1$). It is best to resort to eq. 5.24 which has no such restriction.

5.5 Relationships between S_x^2 and s_x^2, $S_{\overline{x}}^2$ and $s_{\overline{x}}^2$

Consider N independent measurements of x, $x_i : i = 1, 2, \cdots N$. The x_i are distributed according to some parent distribution with a mean μ and a variance σ^2. The estimated mean $\overline{x}$ and the estimated variance s_x^2 are approximations (lower-case letters) that are made from a finte data set for the true mean μ and the true variance σ^2. The "sample" mean, $\overline{X}$ and the "sample" variance are related to the estimate. What is the connection?

The sample mean and the estimated mean are the same

$$\overline{X} = \overline{x} = \frac{1}{N}\sum_{i=1}^{N} x_i \approx \mu \ . \tag{5.25}$$

Since the x_i are independent, their errors combine in quadrature:

$$s_{\overline{x}}^2 = \sum_{i=1}^{N}\left(\frac{\partial \overline{x}}{\partial x_i}\right)^2 \sigma^2 \approx \sum_{i=1}^{N}\left(\frac{\partial \overline{x}}{\partial x_i}\right)^2 s_x^2 \ . \tag{5.26}$$

Chapter 6

Oddities: Random number and precision problems

Now that we understand about random number generators, sampling and error estimation, it is time for a brief respite to consider some of the oddities one might encounter during Monte Carlo calculations. These oddities are related to artefacts associated with random number generation and machine precision.

6.1 Random number artefacts

Consider the determination of the value of π one obtains by throwing random "darts" at a circle inscribed within a square. This is depicted in Figure 6.1. The ratio of the number of darts within the circle to the total number of darts within the square should be $\pi/4$.

For a small number of iterations, the result converges as expected. This is shown in Figure 6.2 where the ratio $4N_{\rm in}/(N\pi)$ is plotted up to 10^4 cycles along with the 1-σ error bars predicted by the "binary statistics" method. The estimated mean goes over and under the theoretical prediction, takes an excursion in the overprediction direction and eventually begins to settle down.

However, some difficulties are evident for large cycles as shown in Figures 6.3. and 6.4

There are some classic signals indicated in Figure 6.3 that the random number generator is cycling. The first piece of evidence is that the result exhibits a periodic structure. The random number generator employed in this study is a multiplicative congruential random number generator (MCRNG) with a sequence length of $2^{30} = 1,073,741,824$. Since two random numbers are employed, the periodic structure occurs over a period of $2^{29} = 536,870,912$. Another curious anomaly is that the result is close to unity (actually to within a few parts

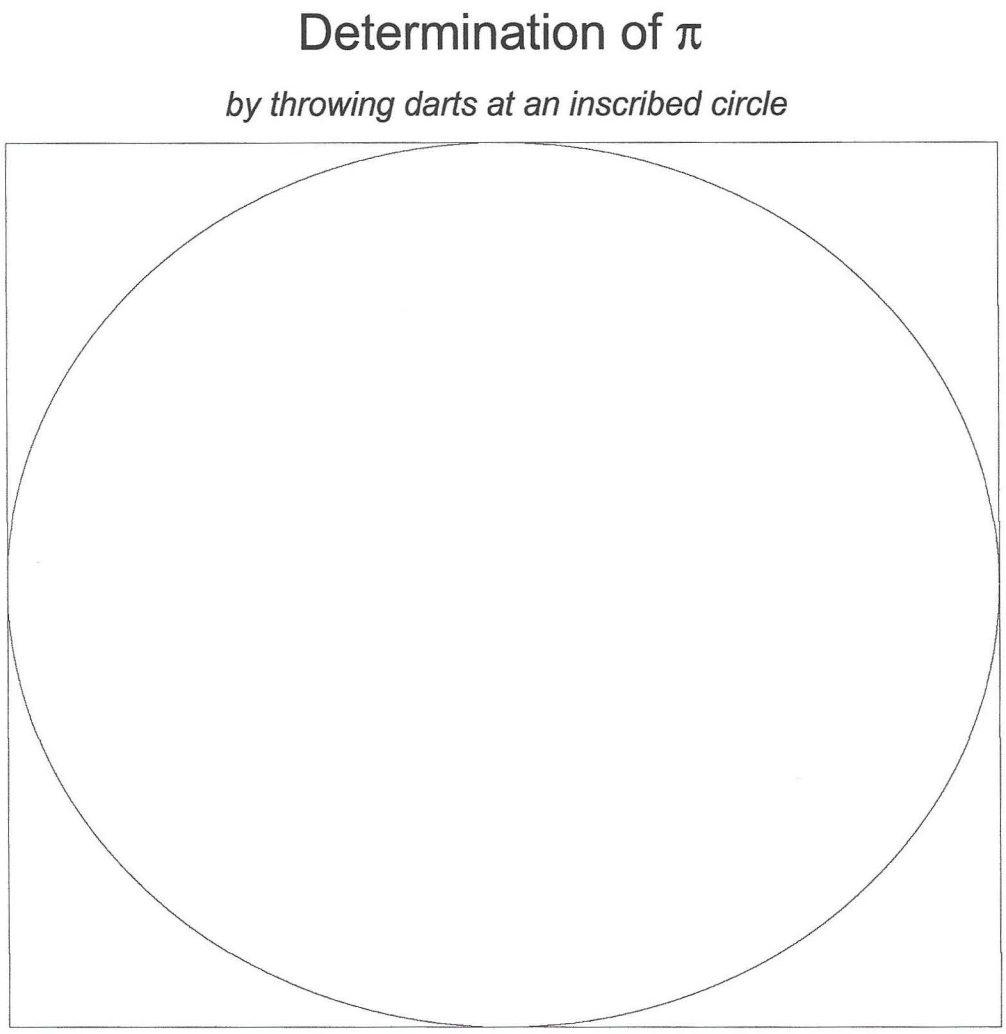


Figure 6.1: Random darts are thrown at a square with an inscribed circle. The ratio of the number of darts within the circle to the total number of darts within the square should be $\pi/4$.

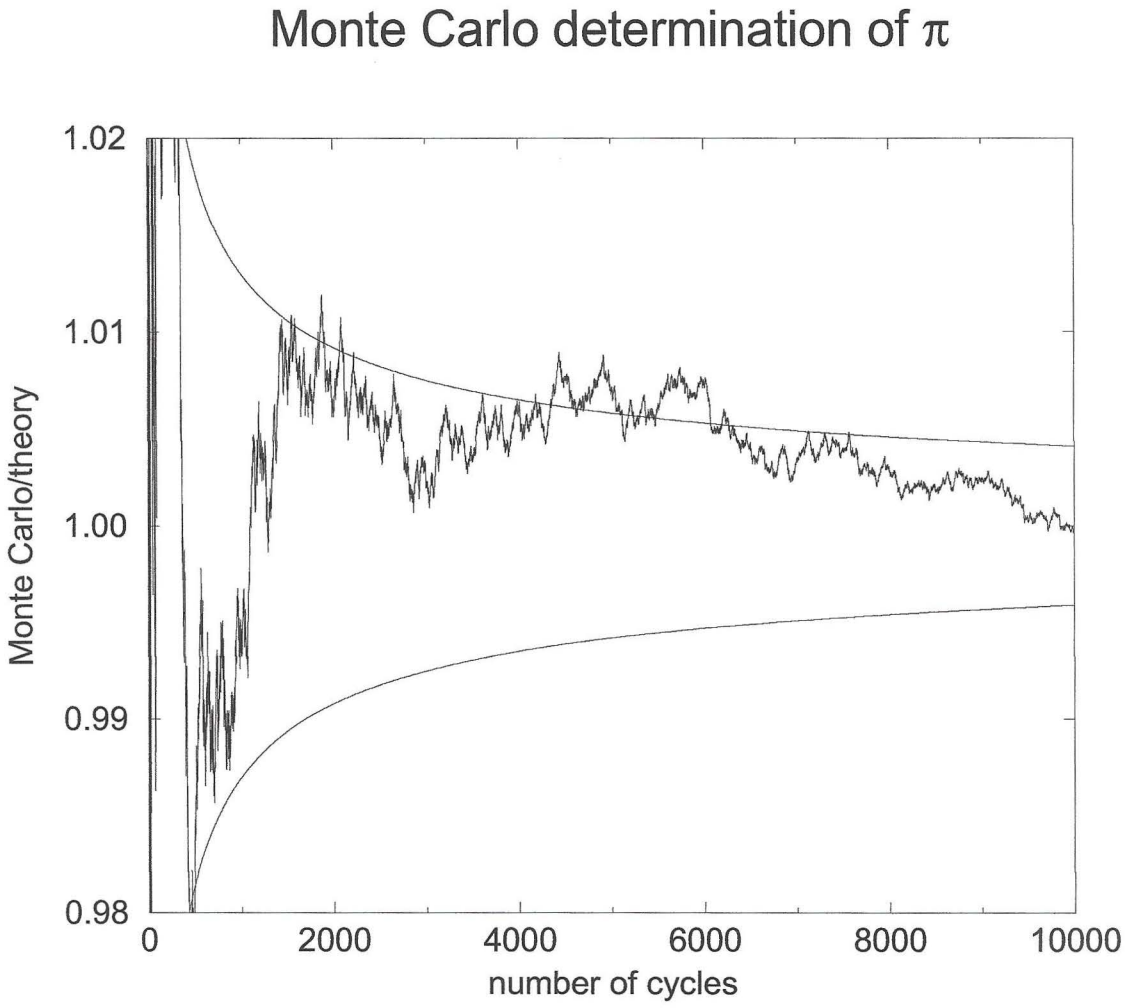


Figure 6.2: Random darts are thrown at a square with an inscribed circle. The Monte Carlo prediction divided by the theoretical prediction with the associated $1 \pm \sigma$ prediction.

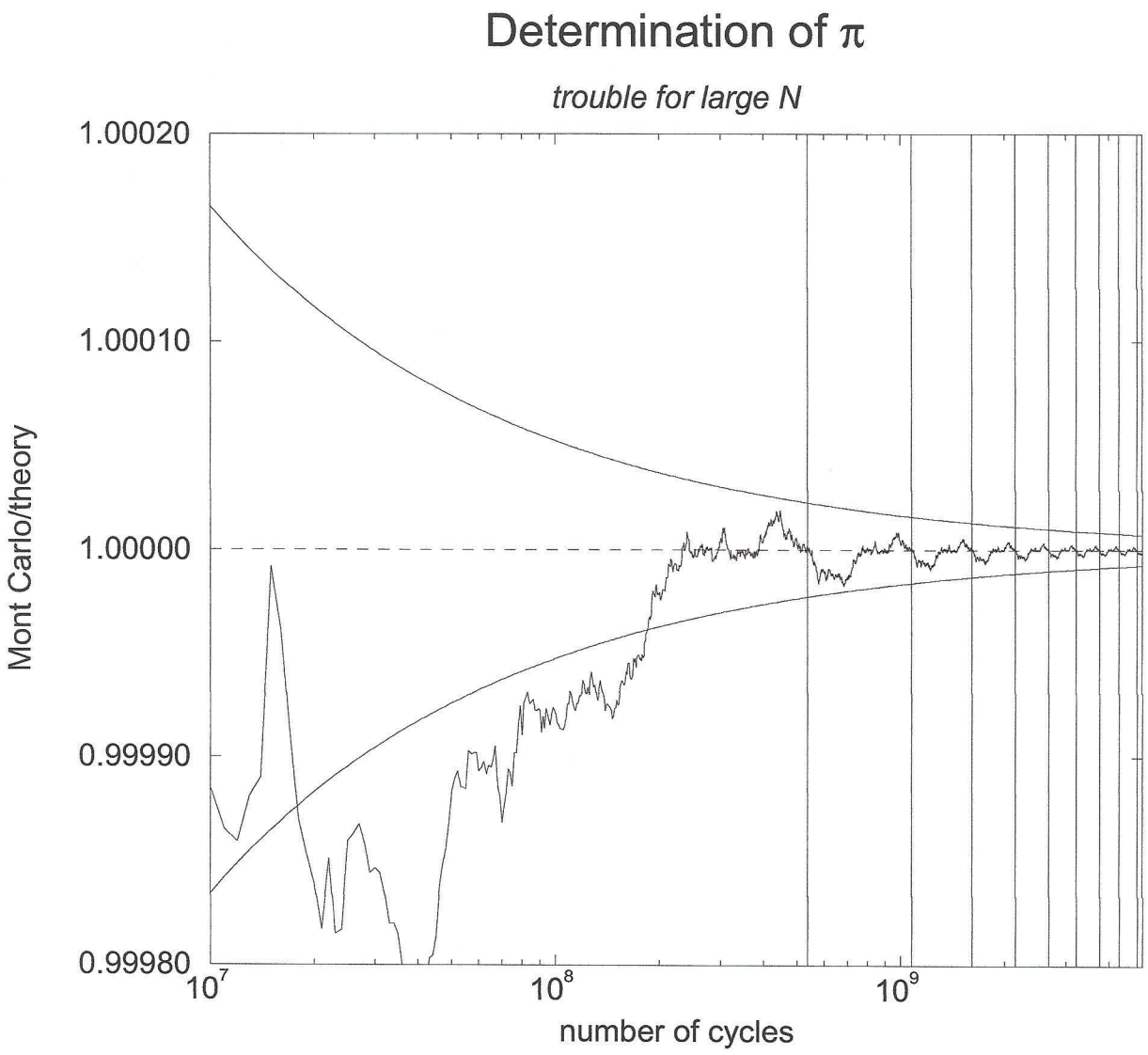


Figure 6.3: Random darts are thrown at a square with an inscribed circle. Large cycle behaviour of the Monte Carlo π experiment. The vertical lines are drawn where the random number generator begins a new cycle.

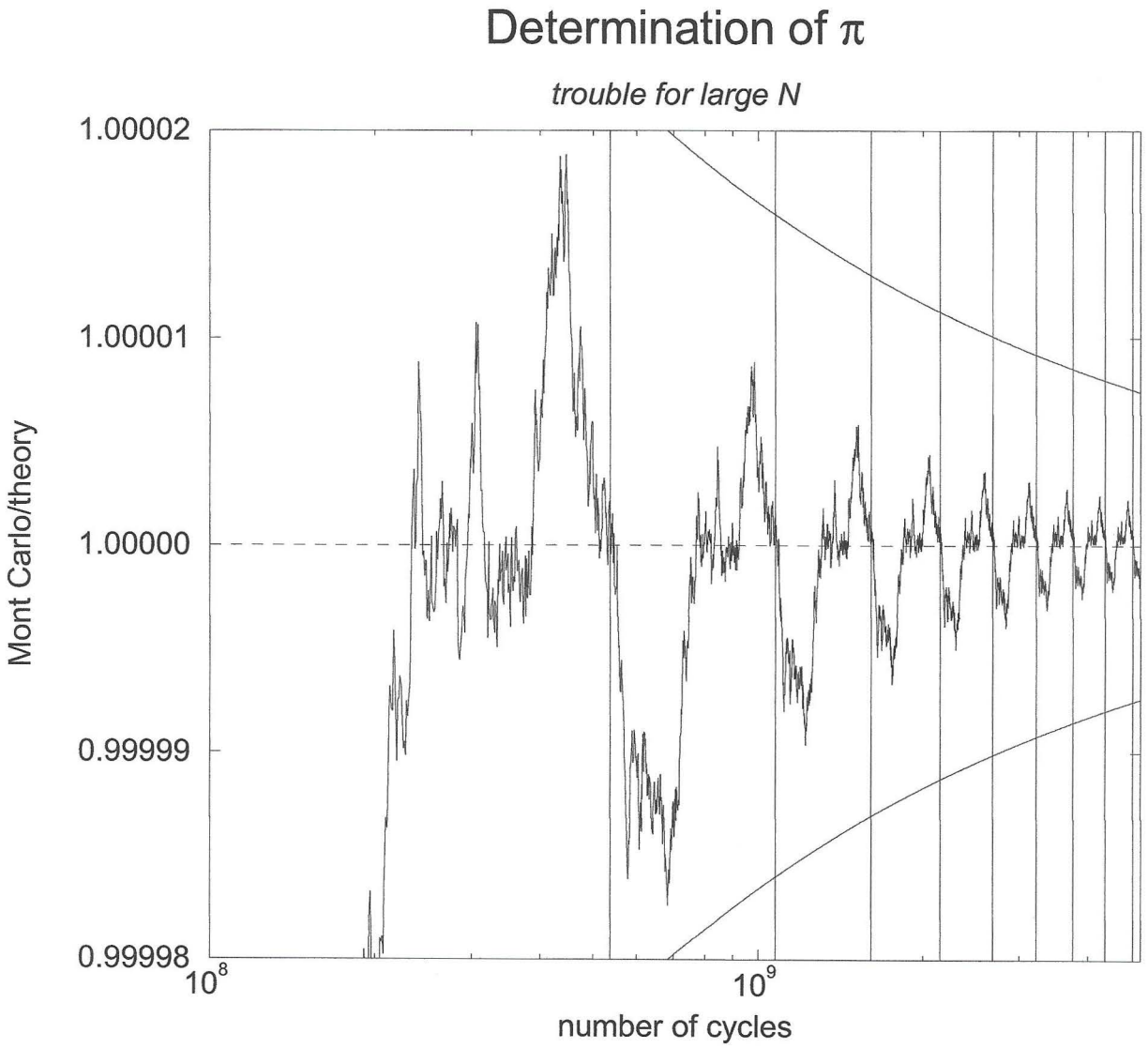


Figure 6.4: A zoom-in on the large cycle behaviour of previous figure.

in 10^7) at the point where the cycle restarts[1]. As a result, the calculated value appears to be well below the 1σ bounds predicted by the Central Limit theorem in the range shown. These are all strong signals that the random number has been "looped". It is never wise to use more than a fraction, say $1/10^{\text{th}}$ of the sequence. Note that the latter half of the sequence anti-correlates with the first half. This could lead to spurious results if a sequence is exhausted.

It is also false to conclude: *"Despite the periodic structure, the result converges to the correct answer."* **Wrong!** We just happened to be lucky in this case! The result converged to about $1+5\times10^{-7}$ after one complete cycle, nearly the correct answer but not **the** correct answer. The "error term" after one cycle just happens to be very small for this application. If we ran this application for about 12×10^9 cycles, we would note a "false convergence" to $1+5\times10^{-7}$ whereas the 1σ bounds would be smaller and converging on unity.

An example of "false convergence" is given in Figure 6.5 which is the same example except that a large number of random numbers were thrown away after each sample of π, as if to simulate many random numbers being employed in a different aspect of a calculation. Although the example is somewhat extreme, it depicts clearly an anomalous result that will never converge to the correct answer.

The object of this lesson is to warn against using random number generators beyond a fraction of their sequence length.

The signals that you have cycled the random number generator are:

- The tally exhibits a period structure.
- The tally converges in a way that is contrary to Central Limit predictions, assuming that the second moment of the tally exists.
- The presence of false convergence, which may be very difficult to detect.

[1]This is due to 2D space being nearly uniformly filled by this MCRNG.

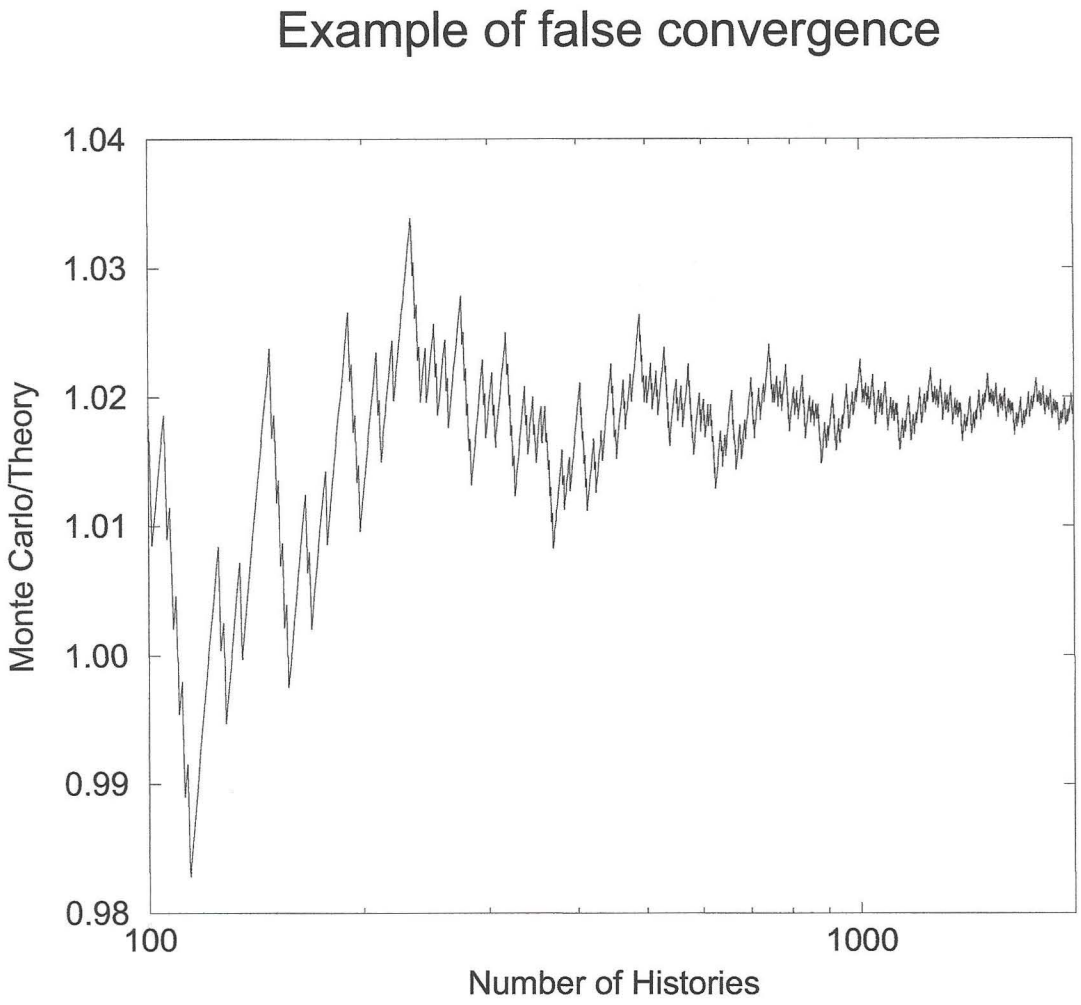


Figure 6.5: An example of false convergence.

6.2 Accumulation errors

Consider the summation:

$$s = \sum_{i=1}^{N} (1/N) \ . \tag{6.1}$$

Of course, mathematically the result is $s \equiv 1$. Numerically, however, it is a different story. The result of s *vs.* N is give in Figure 6.6.

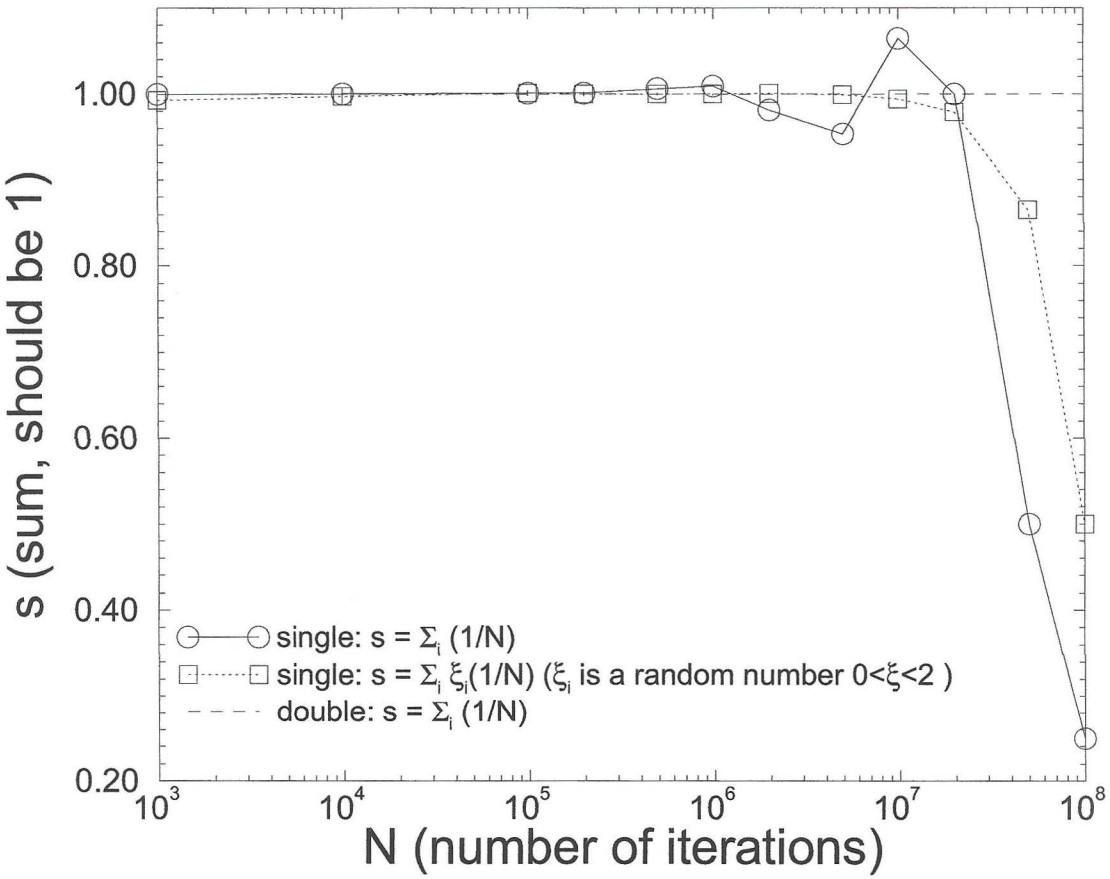


Figure 6.6: An example of constant and random accumulation errors in single-precision arithmetic.

We note that an accumumulation error is seen starting from about 10^6 iterations when the

accumulation is done using single precision Fortran, a 32-bit representation of floating-point numbers. The shape of this curve is the result of constant accumulated round-off error, can be positive or negative, but eventually underestimates due to truncation error. The precise shape of the artefact is probably machine dependent. The reason for the underestimate at large value of N is because $1 + 10^{-8} \equiv 1$ in single precision arithmetic.

Another expression of the similar thing is:

$$s = \sum_{i=1}^{N} (2r/N) \ . \tag{6.2}$$

where r is a random number uniformly distributed on $[0, 1]$. Since $\langle r \rangle = 1/2$, $s \equiv 1$ mathematically as well. However, a numerical evaulation exhibits some accumulation error starting from about 10^7 iterations. This is also seen in Figure 6.6. The shape of this curve is the result of random accumulation error and is probably common to all single-precision architectures. A double precision accumulation is shown as well. For double precision, $1 + 10^{-8} = 1.00000001$ and no accumulation error is evident. Double precision errors would start at about 10^{15} to 10^{16} iterations, a realm where no application has dared to go (yet).

The obvious solution to this problem is: Use double precision! However, there are good reasons for using single precison numbers. On some architectures, single precision arithmetic is faster than double precision. (There are counter examples to this as well!) Double precision numbers also take more computer storage. Fetching and storing them can take longer than for single precision numbers. A good rule is to develop your application in double precision. Then, if fast exacution or computer storage become critical to your application, consider single precision for some, if not all of your calculation. However, you must be aware of the shortcomings (pun intended!) of single precision variables and use them with caution.