(d) The *sample variance* $\tilde{\sigma}^2$ is given by

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n}[y_i - \langle y\rangle]^2}{n-1}. \tag{7.20}$$

The reason for the factor of $n-1$ rather than n in (7.20) is that to compute $\tilde{\sigma}^2$, we need to use the n values of x to compute the mean y, and thus, loosely speaking, we have only $n-1$ independent values of x remaining to calculate $\tilde{\sigma}^2$. Show that if $n \gg 1$, then $\tilde{\sigma}^2 \approx \sigma_y^2$, where σ_y^2 is given by

$$\sigma_y^2 = \langle y^2\rangle - \langle y\rangle^2. \tag{7.21}$$

(e) The quantity $\tilde{\sigma}$ is known as the *standard deviation of the mean.* That is, $\tilde{\sigma}$ gives a measure of how much variation we expect to find if we make repeated measurements of y. How does the value of $\tilde{\sigma}$ compare with your estimate of the variability in part (b)?

(f) What is the qualitative shape of the probability density $p(y)$ that you obtained in part (b)? What is the order of magnitude of the width of the probability?

(g) Verify from your results that $\tilde{\sigma} \approx \sigma_y \approx \sigma_x/\sqrt{n-1} \approx \sigma_x/\sqrt{n}$.

(h) To test the generality of your results, consider the exponential probability density

$$f(x) = \begin{cases} e^{-x} & x \ge 0 \\ 0 & x < 0. \end{cases} \tag{7.22}$$

Calculate $\langle x\rangle$ and σ_x analytically. Modify your Monte Carlo program and estimate $\langle y\rangle$, $\tilde{\sigma}$, σ_y, and $p(y)$. How are $\tilde{\sigma}$, σ_y, and σ_x related for a given value of n? Plot $p(y)$ and discuss its qualitative form and its dependence on n and on the number of measurements of y. ■

Problem 7.15 illustrates the *central limit theorem*, which states that the probability distribution of a sum of random variables, the random variable y, is a Gaussian centered at $\langle y\rangle$ with a standard deviation approximately given by $1/\sqrt{n}$ times the standard deviation of $f(x)$. The requirements are that $f(x)$ has finite first and second moments, that the measurements of y are statistically independent, and that n is large. What is the relation of the central limit theorem to the calculations of the probability distribution in the random walk models that we have considered?

Problem 7.16 Generation of the Gaussian distribution

Consider the sum

$$y = \sum_{i=1}^{12} r_i, \tag{7.23}$$

where r_i is a uniform random number in the unit interval. Make many measurements of y and show that the probability distribution of y approximates the Gaussian distribution with mean value 6 and variance 1. What is the relation of this result to the central limit theorem? Discuss how to use this result to generate a Gaussian distribution with arbitrary mean and variance. This way of generating a Gaussian distribution is particularly useful when a "quick and dirty" approximation is appropriate. A better method for generating a sequence of random numbers distributed according to the Gaussian distribution is discussed in Section 11.5. ■

Many of the problems we have considered have revealed the slow convergence of Monte Carlo simulations and the difficulty of obtaining quantitative results for asymptotic quantities. We conclude this section with a cautionary note and consider a "simple" problem for which straightforward Monte Carlo methods give misleading asymptotic results.

***Problem 7.17 Random walk on lattices containing random traps**

(a) In Problem 7.10 we considered the mean survival time of a one-dimensional random walker in the presence of a periodic distribution of traps. Now suppose that the trap sites are distributed at *random* on a one-dimensional lattice with density $\rho = N/L$. For example, if $\rho = 0.01$, the probability that a site is a trap site is 1%. (A site is a trap site if $r \le \rho$, where, as usual, r is uniformly distributed in the interval $0 \le r < 1$.) If a walker is placed at random at any nontrapping site, determine its mean survival time τ, that is, the mean number of steps before a trap site is reached. Assume that the walker has an equal probability of moving to a nearest neighbor site at each step and use periodic boundary conditions; that is, the lattice sites are located on a ring. The major complication is that it is necessary to perform *three* averages: the distribution of traps, the origin of the walker, and the different walks for a given trap distribution and origin. Choose reasonable values for the number of trials associated with each average and do a Monte Carlo simulation to estimate the mean survival time τ. If τ exhibits a power law dependence on ρ, for example $\tau \approx \tau_0\rho^{-z}$, estimate the exponent z.

(b) A seemingly straightforward extension of part (a) is to estimate the survival probability S_N after N steps. Choose $\rho = 0.5$ and do a Monte Carlo simulation of S_N for N as large as possible. (Published results are for $N = 3 \times 10^4$, on lattices large enough that a walker doesn't reach the boundary, and with about 54,000 trials.) Assume that the asymptotic form of S_N for large N is given by

$$S_N \sim e^{-bN^{\alpha}}, \tag{7.24}$$

where the exponent α is the quantity of interest, and b is a constant that depends on ρ. Are your results consistent with this form? Is it possible to make a meaningful estimate of the exponent α?

(c) It has been proven that the asymptotic N dependence of S_N has the form (7.24) with $\alpha = 1/3$. Are your Monte Carlo results consistent with this value of α? The object of part (b) is to convince you that it is not possible to use simple Monte Carlo methods directly to obtain the correct asymptotic behavior of S_N. The difficulty is that we are trying to estimate S_N in the asymptotic region where S_N is very small, and the small number of trials in this region prevents us from obtaining meaningful results.

(d) One way to reduce the number of required averages is to determine exactly the probability that the walker is at site i after N steps for a given distribution of trap sites. The method is illustrated in Figure 7.4. The first line represents a given configuration of traps distributed randomly on a one-dimensional lattice. One walker is placed at each nontrap site; trap sites are assigned the value 0. Because each walker moves