Table 11.3 Comparison of the Monte Carlo estimates of the integral (11.46) using the uniform probability density $p(x) = 1$ and the nonuniform probability density $p(x) = Ae^{-x}$. The normalization constant A is chosen such that $p(x)$ is normalized on the unit interval. The value of the integral to five decimal places is 0.74682. The estimate F_n, variance σ of f/p, and the probable error $\sigma/n^{1/2}$ are shown. The CPU time (in seconds) is shown for comparison only. (The number of samples was chosen so that the error estimates are comparable.)

	$p(x) = 1$	$p(x) = Ae^{-x}$
n (samples)	5×10^6	4×10^5
F_n	0.74684	0.74689
σ	0.2010	0.0550
$\sigma/\sqrt{n}$	0.00009	0.00009
Total CPU time (s)	20	2.5
CPU time per sample (s)	4×10^{-6}	6×10^{-6}

$f(x)$ is large. A suitable choice of $p(x)$ would make the integrand $f(x)/p(x)$ slowly varying, and hence reduce the variance. Because we cannot evaluate the variance analytically in general, we determine σ *a posteriori*.

As an example, we again consider the integral (see Problem 11.10d)

$$F = \int_0^1 e^{-x^2}\, dx. \tag{11.46}$$

The estimate of F with $p(x) = 1$ for $0 \leq x \leq 1$ is shown in the second column of Table 11.3. A simple choice for the weight function is $p(x) = Ae^{-x}$, where A is chosen such that $p(x)$ is normalized on the unit interval. Note that this choice of $p(x)$ is positive definite and is qualitatively similar to $f(x)$. The results are shown in the third column of Table 11.3. We see that although the computation time per sample for the nonuniform case is larger, the smaller value of σ makes the use of the nonuniform probability distribution more efficient.

Problem 11.15 Importance sampling

(a) Choose $f(x) = \sqrt{1-x^2}$ and consider $p(x) = A(1-x)$ for $x \geq 0$. What is the value of A that normalizes $p(x)$ in the unit interval [0, 1]? What is the relation for the random variable x in terms of r for this form of $p(x)$? What is the variance of $f(x)/p(x)$ in the unit interval? Evaluate the integral $\int_0^1 f(x)\, dx$ using $n = 10^6$ and estimate the probable error of your result.

(b) Choose $p(x) = Ae^{-\lambda x}$ and evaluate the integral

$$\int_0^{\pi} \frac{1}{x^2 + \cos^2 x}\, dx. \tag{11.47}$$

Determine the value of λ that minimizes the variance of the integrand. ■

Problem 11.16 An adaptive approach to importance sampling

An alternative approach is to use the known values of $f(x)$ at regular intervals to sample more often where $f(x)$ is relatively large. Because the idea is to use $f(x)$ itself to determine the probability of sampling, we only consider integrands that are nonnegative. To compute a rough estimate of the relative values of $f(x)$, we first compute its average value by taking k equally spaced points s_i and computing the sum

$$S = \sum_{i=1}^{k} f(s_i). \tag{11.48}$$

This sum divided by k gives an estimate of the average value of f in the interval. The approximate value of the integral is given by $F \approx Sh$, where $h = (b-a)/k$. This approximation of the integral is equivalent to the rectangular or midpoint approximation, depending on where we compute the values of $f(x)$. We then generate n random samples as follows. The probability of choosing subinterval (bin) i is given by the probability

$$p_i = \frac{f(s_i)}{S}. \tag{11.49}$$

Note that the sum of p_i over all subintervals is normalized to unity.

To choose a subinterval with the desired probability, we generate a random number uniformly in the interval $[a, b]$ and determine the subinterval i that satisfies the inequality (11.28). Now that the subinterval has been chosen with the desired probability, we generate a random number x_i in the subinterval $[s_i, s_i + h]$ and compute the ratio $f(x_i)/p(x_i)$. The estimate of the integral is given by the following considerations. The probability p_i in (11.49) is the probability of choosing the subinterval i, not the probability $p(x)\Delta x$ of choosing a value of x between x and $x + \Delta x$. The latter is p_i times the probability of picking the particular value of x in subinterval i:

$$p(x_i)\Delta x = p_i \frac{\Delta x}{h}. \tag{11.50}$$

Hence, we have that

$$F_n = \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)}{p(x_i)} = \frac{h}{n}\sum_{i=1}^{n} \frac{f(x_i)}{p_i}. \tag{11.51}$$

Apply this method to estimate the integral of $f(x) = \sqrt{1-x^2}$ in the unit interval. Under what circumstances would this approach be most useful? ■

11.7 ■ METROPOLIS ALGORITHM

Another way of generating an arbitrary nonuniform probability distribution was introduced by Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller in 1953. The *Metropolis* algorithm is a special case of an importance sampling procedure in which certain possible sampling attempts are rejected (see Appendix 11C). The Metropolis method is useful for computing averages of the form

$$\langle f \rangle = \frac{\int f(x)p(x)\, dx}{\int p(x)\, dx}, \tag{11.52}$$