

That is, on each iteration the value of the estimator is between 0 and 0.001236. Since, in this case, θ is the probability that a binomial random variable with parameters 20, 0.4 is at least 16, it can be explicitly computed with the result $\theta = 0.000317$. Hence, the raw simulation estimator I , which on each iteration takes the value 0 if the sum of the Bernoullis with parameter 0.4 is less than 16 and takes the value 1 otherwise, will have variance

$$\text{Var}(I) = \theta(1 - \theta) = 3.169 \times 10^{-4}$$

On the other hand, it follows from the fact that $0 \leq \hat{\theta} \leq 0.001236$ that (see [Exercise 33](#))

$$\text{Var}(\hat{\theta}) \leq 2.9131 \times 10^{-7} \quad \blacksquare$$

Example 11.24 Consider a single-server queue in which the times between successive customer arrivals have density function f and the service times have density g . Let D_n denote the amount of time that the n th arrival spends waiting in queue and suppose we are interested in estimating $\alpha = P\{D_n \geq a\}$ when a is much larger than $E[D_n]$. Rather than generating the successive interarrival and service times according to f and g , respectively, they should be generated according to the densities f_{-t} and g_t , where t is a positive number to be determined. Note that using these distributions as opposed to f and g will result in smaller interarrival times (since $-t < 0$) and larger service times. Hence, there will be a greater chance that $D_n > a$ than if we had simulated using the densities f and g . The importance sampling estimator of α would then be

$$\hat{\alpha} = I\{D_n > a\} e^{t(S_n - Y_n)} [M_f(-t) M_g(t)]^n$$

where S_n is the sum of the first n interarrival times, Y_n is the sum of the first n service times, and M_f and M_g are the moment generating functions of the densities f and g , respectively. The value of t used should be determined by experimenting with a variety of different choices. ■

11.7 Determining the Number of Runs

Suppose that we are going to use simulation to generate r independent and identically distributed random variables $Y^{(1)}, \dots, Y^{(r)}$ having mean μ and variance σ^2 . We are then going to use

$$\bar{Y}_r = \frac{Y^{(1)} + \dots + Y^{(r)}}{r}$$

as an estimate of μ . The precision of this estimate can be measured by its variance

$$\begin{aligned} \text{Var}(\bar{Y}_r) &= E[(\bar{Y}_r - \mu)^2] \\ &= \sigma^2/r \end{aligned}$$

Hence, we would want to choose r , the number of necessary runs, large enough so that σ^2/r is acceptably small. However, the difficulty is that σ^2 is not known in advance. To get around this, you should initially simulate k runs (where $k \geq 30$) and then use the simulated values $Y^{(1)}, \dots, Y^{(k)}$ to estimate σ^2 by the sample variance

$$\sum_{i=1}^k (Y^{(i)} - \bar{Y}_k)^2 / (k - 1)$$

Based on this estimate of σ^2 the value of r that attains the desired level of precision can now be determined and an additional $r - k$ runs can be generated.

11.8 Generating from the Stationary Distribution of a Markov Chain

11.8.1 Coupling from the Past

Consider an irreducible Markov chain with states $1, \dots, m$ and transition probabilities $P_{i,j}$ and suppose we want to generate the value of a random variable whose distribution is that of the stationary distribution of this Markov chain. Whereas we could *approximately* generate such a random variable by arbitrarily choosing an initial state, simulating the resulting Markov chain for a large fixed number of time periods, and then choosing the final state as the value of the random variable, we will now present a procedure that generates a random variable whose distribution is *exactly* that of the stationary distribution.

If, in theory, we generated the Markov chain starting at time $-\infty$ in any arbitrary state, then the state at time 0 would have the stationary distribution. So imagine that we do this, and suppose that a different person is to generate the next state at each of these times. Thus, if $X(-n)$, the state at time $-n$, is i , then person $-n$ would generate a random variable that is equal to j with probability $P_{i,j}$, $j = 1, \dots, m$, and the value generated would be the state at time $-(n-1)$. Now suppose that person -1 wants to do his random variable generation early. Because he does not know what the state at time -1 will be, he generates a sequence of random variables $N_{-1}(i)$, $i = 1, \dots, m$, where $N_{-1}(i)$, the next state if $X(-1) = i$, is equal to j with probability $P_{i,j}$, $j = 1, \dots, m$. If it results that $X(-1) = i$, then person -1 would report that the state at time 0 is

$$S_{-1}(i) = N_{-1}(i), \quad i = 1, \dots, m$$

(That is, $S_{-1}(i)$ is the simulated state at time 0 when the simulated state at time -1 is i .)

Now suppose that person -2 , hearing that person -1 is doing his simulation early, decides to do the same thing. She generates a sequence of random variables $N_{-2}(i)$, $i = 1, \dots, m$, where $N_{-2}(i)$ is equal to j with probability $P_{i,j}$, $j = 1, \dots, m$. Consequently, if it is reported to her that $X(-2) = i$, then she will report that $X(-1) = N_{-2}(i)$. Combining this with the early generation of person -1 shows