Probability in the Engineering and Informational Sciences, 11, 1997, 267–271. Printed in the U.S.A.

ESTIMATING THE MEAN COVER TIME OF A SEMI-MARKOV PROCESS VIA SIMULATION

EROL PEKÖZ AND SHELDON M. ROSS
Department of Industrial Engineering and Operations Research
University of California, Berkeley
Berkeley, California 94720

An efficient simulation estimator of the expected time until all states of a finite state semi-Markov process have been visited is determined.

1. INTRODUCTION

Consider a semi-Markov process that, after entering state i, next goes to state j with probability $P_{i,j}$, and given that the next state is j, the time until the transition from i to j occurs is a random variable with distribution $F_{i,j}$ having mean $m(i,j)$. Starting in state 0, suppose we are interested in estimating $\mu = E[T]$, where T, called the cover time, is the time until all of the states $1,2,\ldots,m$ have been visited. Let $\mu(i,j)$ denote the expected time, given that the process has just entered state i, until it enters state j, and suppose that we are able to compute all of the values of $\mu(i,j)$ for the pairs i,j of interest. In Section 2, we show how μ can be efficiently estimated by a simulation of the embedded Markov chain with transition probabilities $P_{i,j}$. We then consider the problem of using simulation to estimate $E[T_n]$, the mean time until n of the states $1,\ldots,m$ have been visited, where $1 < n < m$. In Section 3, we present an estimator of $E[T_n]$ that is recommended when n is not too small. A different simulation estimator, which involves a conditional expectation and uses random hazards as control variates, and which is preferable to the estimator of Section 3 when n is small, is presented in Section 4.

© 1997 Cambridge University Press 0269-9648/97 $11.00 + .10

2. THE SIMULATION ESTIMATOR

Suppose that the semi-Markov process has been simulated up to the point that all the states $1,\dots,m$ have been visited. Let $i_1,\dots,i_m$ be any permutation of $1,\dots,m$. Let A_1 denote the time at which the process first enters state i_1; let A_2 denote the additional time after A_1 until the process has visited both i_1 and i_2; and, in general, let A_j denote the additional time after the process has visited $i_1,\dots,i_{j-1}$ until it has also visited i_j. (Thus, if i_j is not the last of $i_1,\dots,i_j$ to be visited, then $A_j = 0$.) With these definitions, we have $T = \sum_{j=1}^m A_j$ and so

$$\mu = \sum_{j=1}^{m} E[A_j].$$

Letting $L(i_1,\dots,i_k)$ denote the last one of the states $i_1,\dots,i_k$ to be visited, and letting $T(i,j)$ denote the time it takes to go from state i to state j, we have that

$$A_j = I\{L(i_1,\dots,i_j) = i_j\}T(L(i_1,\dots,i_{j-1}),i_j).$$

Hence,

$$E[A_j \mid L(i_1,\dots,i_{j-1}),L(i_1,\dots,i_j)] = I\{L(i_1,\dots,i_j) = i_j\}\mu(L(i_1,\dots,i_{j-1}),i_j),$$

and so $\sum_{j=1}^m I\{L(i_1,\dots,i_j) = i_j\}\mu(L(i_1,\dots,i_{j-1}),i_j)$ is an unbiased estimator of μ.

As the preceding is true for all permutations, it follows that

$$\hat{\mu} = \frac{1}{m!}\sum\sum_{j=1}^{m} I\{L(i_1,\dots,i_j) = i_j\}\mu(L(i_1,\dots,i_{j-1}),i_j)$$

is an unbiased estimator of μ, where the leftmost summation is over all the $m!$ permutations.

Now, let V_k denote the kth new state to be visited, $k = 1,\dots,m$, and consider the coefficient of $\mu(V_r, V_k)$ in the expression $\hat{\mu}$. As it is 0 for $r > k$, suppose that $r < k$. The coefficient of $\mu(V_r, V_k)$ will be the number of permutations of the form $i_1,\dots,i_{j-1},V_k,i_{j+1},\dots,i_m$ such that $L(i_1,\dots,i_{j-1},V_k) = V_k$ and $L(i_1,\dots,i_{j-1}) = V_r$. Because for this to be true $i_1,\dots,i_{j-1}$ must include V_r and must all be among the first r states visited, it follows that, for fixed j, there are $(j-1)(r-1)\cdots(r-j+2)(m-j)! = (m-1)!\binom{r-1}{j-2}\big/\binom{m-1}{j-1}$ such permutations. Hence, including the coefficients of $\mu(0, V_k)$ [all equal to $(m-1)!$] we obtain that

$$\hat{\mu} = \frac{1}{m}\sum_{i=1}^{m}\mu(0,i) + \frac{1}{m}\sum_{k=2}^{m}\sum_{r=1}^{k-1}\mu(V_r, V_k)\sum_{j=2}^{m}\binom{r-1}{j-2}\bigg/\binom{m-1}{j-1},$$

where we use the convention that $\binom{n}{i} = 0$ when $n < i$.

Because the only quantity needed from each simulation run is the ordering in which the states are visited, it is only necessary to simulate the embedded

Markov chain to obtain this quantity. Once all the states $1, \ldots, m$ have been visited, the data V_k, $k = 1, \ldots, m$, should be collected and $\hat{\mu}$ computed. This should then be repeated over many simulation runs and the average of the values of $\hat{\mu}$ is the estimator of μ. Because the terms of the form $\sum_{j=2}^{m} \binom{r-1}{j-2} / \binom{m-1}{j-1}$ only need be computed once and saved, the computation of $\hat{\mu}$ only involves a double summation.

Let

$$\mu_1 = \min_{i,j} \mu(i,j), \qquad \mu_2 = \max_{i,j} \mu(i,j).$$

It is known (see Matthews [1]) that

$$\frac{1}{m} \sum_{i=1}^{m} \mu(0,i) + \mu_1 \sum_{i=2}^{m} 1/i \le \mu \le \frac{1}{m} \sum_{i=1}^{m} \mu(0,i) + \mu_2 \sum_{i=2}^{m} 1/i.$$

We now show that the estimator $\hat{\mu}$ also falls within these bounds.

THEOREM 1:

$$\frac{1}{m} \sum_{i=1}^{m} \mu(0,i) + \mu_1 \sum_{i=2}^{m} 1/i \le \hat{\mu} \le \frac{1}{m} \sum_{i=1}^{m} \mu(0,i) + \mu_2 \sum_{i=2}^{m} 1/i.$$

The theorem follows immediately from the following lemma.

LEMMA 1:

$$\frac{1}{m} \sum_{k=2}^{m} \sum_{r=1}^{k-1} \sum_{j=2}^{m} \binom{r-1}{j-2} \Bigg/ \binom{m-1}{j-1} = \sum_{i=2}^{m} 1/i.$$

PROOF: Consider the Markov chain for which $P_{i,j} = 1/m$, $i,j = 1, \ldots, m$. In this case, $\mu(i,j) = m$ and so $\hat{\mu}$ is a constant. Because $\hat{\mu}$ is an unbiased estimator, it must thus equal $\mu = \sum_{i=1}^{m} m/i$ (for this is just the coupon collector's problem with equal probabilities). ∎

3. ESTIMATING THE MEAN TIME UNTIL n DISTINCT STATES HAVE BEEN VISITED

For $n \le m$, let T_n denote the time until n of the states $1, \ldots, m$ have been visited (and so T_m is equal to T of the previous section), and suppose that we are interested in using simulation to estimate $E[T_n]$. For n small, we propose using the raw simulation estimator in conjunction with the n hazards as controls. Now suppose n is not small. In this case, we propose estimating $E[T_n]$ by using $\hat{\mu}$, the estimator of $E[T_m]$ of Section 2, minus an estimator of $E[T_m - T_n]$.

To estimate $E[T_m - T_n]$, let $i_1, \ldots, i_m$ be any permutation of $1, \ldots, m$. Now, suppose that the process has been simulated until all the states $1, \ldots, m$ have been visited. Let, as before, V_i be the ith state visited, and let R_i be such that state i is the R_ith state visited. That is, if $V_i = j$, then $R_j = i$.

Let S_0 denote the first time that n distinct states have been visited, and for $j = 1, \ldots, m$ let S_j denote the first time at which the process has visited at least n states including all of the states $i_1, \ldots, i_j$. Let L_j be the state that is entered at time S_j, $j = 0, \ldots, m$. Then, we may write

$$T_m - T_n = \sum_{j=1}^{m} (S_j - S_{j-1})^+.$$

Now, for $j = 1, \ldots, m$.

$$E[(S_j - S_{j-1})^+ \,|\, R_{i_1}, \ldots, R_{i_j}, L_{j-1}]$$
$$= I\{R_{i_j} > n\} I\{R_{i_j} = \max(R_{i_1}, \ldots, R_{i_j})\} \mu(L_{j-1}, i_j).$$

Hence, $\sum_{j=1}^{m} I\{R_{i_j} > n\} I\{R_{i_j} = \max(R_{i_1}, \ldots, R_{i_j})\} \mu(L_{j-1}, i_j)$ is an unbiased estimator of $E[T_m - T_n]$. As this is true for all the $m!$ permutations, we see that

$$\hat{\mu}_{n,m} = \frac{1}{m!} \sum \sum_{j=1}^{m} I\{R_{i_j} > n\} I\{R_{i_j} = \max(R_{i_1}, \ldots, R_{i_j})\} \mu(L_{j-1}, i_j)$$

is also an unbiased estimator, where the first sum is over all $m!$ permutations.

To simplify the preceding, consider the coefficient of $\mu(V_r, V_k)$, which will only be positive when $n \le r < k$. The coefficient of $\mu(V_n, V_k)$ is the number of permutations of the form $i_1, \ldots, i_{j-1}, V_k, i_{j+1}, \ldots, i_m$ for which $i_1, \ldots, i_{j-1}$ are all among the first n states visited. As such a permutation can either have or not have V_n as one of its first $j-1$ components, we see that there are

$$(j-1)(n-1) \cdots (n-j+2)(m-j)! + (n-1) \cdots (n-j+1)(m-j)!$$
$$= \frac{n!\,(m-j)!}{(n-j+1)!}$$

such permutations. For $r > n$, the coefficient of $\mu(V_r, V_k)$ is the number of permutations of the form $i_1, \ldots, i_{j-1}, V_k, i_{j+1}, \ldots, i_m$ for which V_r is one of $i_1, \ldots, i_{j-1}$ and $i_1, \ldots, i_{j-1}$ are all among the first r states visited. Hence, there are

$$(j-1)(r-1) \cdots (r-j+2)(m-j)! = (m-1)! \binom{r-1}{j-2} \Big/ \binom{m-1}{j-1}$$

such permutations. Hence,

$$\hat{\mu}_{n,m} = \frac{1}{m!} \sum_{k=n+1}^{m} \mu(V_n, V_k) \sum_{j=1}^{m} n!\,(m-j)!/(n-j+1)!$$
$$+ \frac{1}{m} \sum_{k=n+2}^{m} \sum_{r=n+1}^{k-1} \mu(V_r, V_k) \sum_{j=2}^{m} \binom{r-1}{j-2} \Big/ \binom{m-1}{j-1}$$

is an unbiased estimator of $E[T_m - T_n]$. Hence, when n is not small, we recommend simulating the embedded Markov chain until all the states $1, \ldots, m$

have been visited and then evaluating $\hat{\mu} - \hat{\mu}_{n,m}$. Its average over many simulation runs is the simulation estimator of $E[T_n]$.

4. ESTIMATING THE MEAN TIME UNTIL n DISTINCT STATES HAVE BEEN VISITED: n SMALL

When n is small, we recommend simulating the embedded Markov chain until n of the states $1, \ldots, m$ have been visited. Let N denote the number of transitions needed, and let $0 = X_0, X_1, \ldots, X_N$ be the sequence of states visited by the embedded chain. Then,

$$\hat{\mu}_n = E[T_n \mid N, X_1, \ldots, X_N] = \sum_{i=1}^{N} m(X_{i-1}, X_i)$$

is an unbiased estimator of $E[T_n]$, where $m(i,j)$ is the conditional expected time it takes the process to make a transition from state i to state j given that it has just entered i and will next go to j. This estimator can, however, be improved by the use of control variates (see Ross [2]). Let $I_1, \ldots, I_n$ denote the n distinct states visited, in the order of visit; let $T_1, \ldots, T_n$ denote the transition numbers of the first time these states are visited (so, e.g., $X_{T_j} = I_j$). Also, let $D_0 = \{0\}$ and $D_j = \{0, I_1, \ldots, I_j\}$. The random hazards H_j, $j = 1, \ldots, n$, are defined by

$$H_j = \sum_{i=T_{j-1}}^{T_{j-1}} \sum_{k \notin D_{j-1}} P_{X_i,k}.$$

These random hazards all have mean 1, and as they would appear to be negatively correlated with $\hat{\mu}_n$ an estimator of the form

$$\hat{\mu}_n + \sum_{j=1}^{n} C_j(H_j - 1)$$

should be quite efficient. The values of the control variate constants C_j should be obtained from the simulation by standard techniques.

References

1. Matthews, P. (1988). Covering problems for Markov chains. *Annals of Probability* 16: 1215–1228.
2. Ross, S.M. (1997). *Simulation*, 2nd ed. New York: Academic Press.