Probability in the Engineering and Informational Sciences, **21**, 2007, 217–234. Printed in the U.S.A.
DOI: 10.1017/S0269964807070143

ASYMPTOTIC VARIANCE OF PASSAGE TIME ESTIMATORS IN MARKOV CHAINS

Michael A. Zazanis
Department of Statistics
Athens University of Economics and Business
Athens 104 34, Greece
E-mail: zazanis@aueb.gr

We consider the problem of estimating passage times in stochastic simulations of Markov chains. Two types of estimator are considered for this purpose: the "simple" and the "overlapping" estimator; they are compared in terms of their asymptotic variance. The analysis is based on the regenerative structure of the process and it is shown that when estimating the *mean* passage time, the simple estimator is always asymptotically superior. However, when the object is to estimate the expectation of a nonlinear function of the passage time, such as the probability that the *passage time exceeds a given threshold*, then it is shown that the overlapping estimator can be superior in some cases. Related results in the Reinforcement Learning literature are discussed.

1. INTRODUCTION

This article deals with passage time estimation in discrete-time Markov chains. Starting from a given state, i, of an irreducible, positive recurrent Markov chain $\{X_n; n \in \mathbb{N}\}$ with countable state space $\mathbb{S}$, we want to estimate the mean time required for the process to reach another state, j. More generally, if $T^j = \inf\{n > 0: X_n = j\}$ denotes the hitting time of state j, then we are interested in estimating $E_i f(T^j)$, where the function $f: \mathbb{S} \mapsto \mathbb{R}$ is such that the expectation exists. Of particular importance is the probability that the passage time exceeds a given threshold, $\nu \in \mathbb{N}$, obtained if we choose $f(x) = \mathbf{1}(x > \nu)$.

 0269-9648/07 $25.00

For Markov chains with a large state space and lack of structure, simulation might be the only practical option; thus, the problem we discuss is relevant in applications. Passage-time-estimation problems arise naturally in many situations. For instance, in queuing networks, one is often interested in the statistics of the time until a certain critical buffer overflows. Harrison and Knottenbelt [7] gave a modern survey of passage time problems in queuing applications. Such problems are of great importance in the analysis of manufacturing systems and communications networks. Another important application for passage time estimation comes from the connection between the mean passage time, $E_i T_j$, and the potential (or fundamental) matrix of the Markov chain, (Z_{ij}), defined via $Z_{ij} = \sum_{n=0}^{\infty}(P_{ij}^n - \pi_j)$. This is related in turn to bias properties, rapidity of convergence to steady state, and the sensitivity of the stationary distribution with respect to perturbations of the transition probability matrix. For this last application, which is becoming increasingly important, we refer the reader to Meyer [11], Cao and Chen [3], Cho and Meyer [4], and Heidergott and Hordijk [8].

A number of approaches have been proposed for improving the naïve Monte Carlo estimate of the passage time. In particular, we mention Ross and Schechner [14], who proposed a method using hazard rates (and which is essentially based on conditioning) in order to reduce the variance of the estimator. An alternative idea, which becomes particularly advantageous when the target state corresponds to a rare event, uses change of measure arguments. We refer the reader to Glynn and Iglehart [6], Nakayama, Goyal, and Glynn [12], and Shahabuddin [15].

However, even the naïve estimate for passage times presents some intricacies. If starting from a given state, i, we happen to return to i several times before we eventually visit j, are we to consider that we have multiple, correlated measurements of the passage time or a single one? Depending on how we answer this question, we obtain two different estimators. The *simple estimator* derives a single estimate of the passage time from each such realization. The *overlapping estimator*, on the other hand, derives one additional measurement for each return to state i before the visit to j. However, these multiple measurements obtained by the overlapping estimator are, of course, not independent. As it is shown in Section 3, both estimators are strongly consistent, so we need a criterion to decide which one is preferable. We answer this question using an *asymptotic variance comparison* carried out via elementary regenerative arguments in Section 4. Despite the messy analysis, the question has a rewarding answer. For the problem of estimating *mean passage times*, the simple estimator, which ignores repeated visits to the initial state i, is always superior compared to the overlapping estimator, and the ratio of the two asymptotic variances has a simple form. On the other hand, when estimating the *probability that the passage time exceeds a given threshold*, the situation is more complicated and, in some cases, the overlapping estimator might be asymptotically superior to the simple estimator.

Another important application of the estimation problem that we discuss in this article stems from its connection with Reinforcement Learning. This is a simulation-based stochastic approximation framework that has been successfully used for the computation of near-optimal policies in large-scale Markov decision processes (see Bertsekas and Tsitsiklis [1]).

In the analysis of such algorithms there appear a "first visit" and an "every visit" estimator that correspond to the simple and overlapping estimators, respectively, introduced in this article. In Singh and Sutton [16] some of the statistical properties of these estimators have been analyzed in the context of Reinforcement Learning algorithms and, in particular, it has been shown that the "every visit" (or overlapping) estimator is asymptotically inferior to the "first visit" (or simple) estimator in terms of the mean squared error (MSE) criterion. In this article, (1) we provide a rigorous analysis of the statistical properties of the estimators, (2) by using standard results on asymptotic statistical properties of ratio estimators we are able to shift our focus from the MSE to the asymptotic variance ratio of the two estimators and thus provide much simpler asymptotic results, and (3) we examine not only mean passage times but also expectations of nonlinear functions of the passage time as well, such as the probability that the passage time exceeds a given threshold. In that case, there are situations where the overlapping estimator might be superior to the simple estimator in terms of asymptotic variance.

Finally, we should point out the conceptual affinity between the estimators in this article and the ones proposed in Calvin and Nakayama [2], who consider regenerative simulations with two sequences of regeneration points and use a permutation scheme that results in variance reduction.

Preliminary results on this problem, together with an analysis of a similar problem involving superposition of two point processes, can also be found in Karamichalakou and Zazanis [9]. For analytic results regarding passage times in Markov chains we refer the reader to Kemperman [10], and Syski [17].

2. ESTIMATORS FOR THE MEAN PASSAGE TIME IN MARKOV CHAINS

We consider a discrete-time Markov chain, $\mathbb{X} = \{X_n; n = 0, 1, 2, \ldots\}$, on a countable state space $\mathbb{S}$. We will denote its transition probability matrix by (P_{ij}), and we will assume that it is irreducible and positive recurrent. We are interested in estimating the mean passage time from state i to state j using data obtained from the sample path of the chain. Although we restrict our attention here to discrete-time processes, the situation in continuous time is quite analogous. Let $S_1^i = \min\{n \geq 0 : X_n = i\}$ denote the time of the first visit of $\mathbb{X}$ to state i and $S_1^j = \min\{n > S_1^i : X_n = j\}$ denote the time of the *first visit to j after the first visit to i*. We define recursively $S_k^i = \min\{n > S_{k-1}^j : X_n = i\}$ and $S_k^j = \min\{n > S_k^i : X_n = j\}$ for $k = 2, 3, \ldots$. The $\{S_k^i; k \geq 1\}$ are a (possibly delayed) sequence of regeneration points for $\mathbb{X}$ and, with the possible exception of an initial segment until the first visit to state i, the sample path of $\mathbb{X}$ splits into cycles starting with a visit to i and ending with the first subsequent return to *i after visiting j*.

Consider a sample path consisting of n such cycles and let us take a closer look at the kth cycle. Two estimators for the mean passage time from i to j naturally suggest themselves:

- The *simple estimator*

$$\Phi_n := \frac{1}{n}\sum_{k=1}^{n} W_{1,k},$$

where n is the number of cycles and $W_{1,k} := S_k^j - S_k^i$ is the first passage time from state i to state j in the kth cycle (see Fig. 1).
- The *overlapping estimator*

$$\Psi_n := \frac{\sum_{k=1}^{n}\sum_{l=1}^{Q_k} W_{l,k}}{\sum_{k=1}^{n} Q_k},$$

where n is again the number of cycles, Q_k is the total number of visits to state i in the kth cycle, and $W_{l,k}$ is the passage time from state i to state j for the lth visit of state i in the kth cycle (see Fig. 2 and the discussion that follows).

In order to clarify the structure of the overlapping estimator, note that each cycle begins with a visit to i and consists of a number of excursions from i during which the process returns to i without having visited j plus a final excursion starting from i and reaching j without returning to i. The last segment of each cycle is the return from j to i. During this last segment, the state j can appear several times, but the state i appears only at the end of the segment. The appearance of i signals the beginning of the next cycle. If we denote by Q the number of times that state i occurs during a typical cycle (note that we have dropped the subscript that refers to the cycle for simplicity), then we have $Q-1$ excursions starting from i and ending at i that do not include state j plus a trip from i to j that does not return to i. Let us denote by $U_1, U_2, \ldots, U_{Q-1}$ the lengths of the $Q-1$ excursions that return to i and by V the length of the final trip. Then, during this cycle, the overlapping estimator

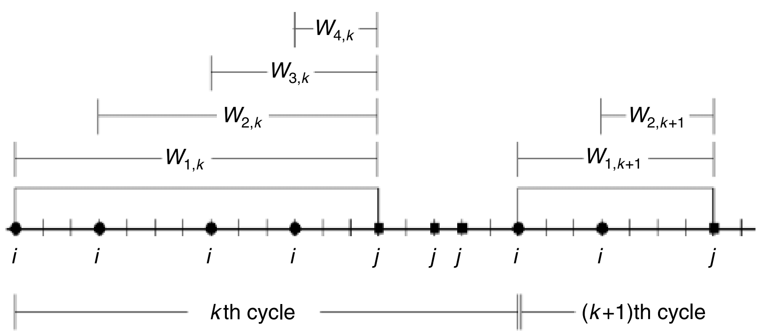

FIGURE 1. Overlapping passage time estimator in a Markov chain.

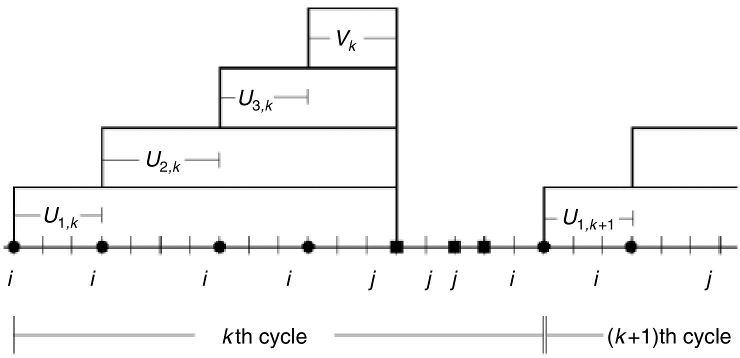

FIGURE 2. Structure of the overlapping estimator.

counts Q passage times from i to j, namely $W_1 = U_1 + U_2 + \cdots + U_{Q-1} + V$ and $W_2 = U_2 + \cdots + U_{Q-1} + V, \ldots, W_Q = V$. For instance, in Figure 2, $Q_k = 4$, $W_{1,k} = U_{1,k} + U_{2,k} + U_{3,k} + V_k$, and $W_{3,k} = U_{3,k} + V_k$.

Clearly, because of the regenerative structure of this process, the resulting cycles are independent and identically distributed (i.i.d.) objects, and, in particular, the random variables Q_k, $k = 1, 2, \ldots$, are i.i.d. Let

$$p := P_i(T^j < T^i) \tag{1}$$

denote the probability that, starting from state i, we visit state j before we return to i, and let

$$q := 1 - p = P_i(T^i < T^j) \tag{2}$$

denote the probability of the complementary event. Then it is clear by the strong Markov property that Q is a geometric random variable with probability of success p. Furthermore, also as a consequence of the strong Markov property, it follows that the successive excursion times from i to i without visiting j within a cycle, U_l, $l = 1, 2, \ldots, Q-1$, are conditionally independent, given Q, with common distribution

$$g_{ii}(n) := P_i(T^i = n | T^i < T^j), \qquad n = 1, 2, \ldots,$$

which *does not depend on Q*. Finally, V is independent of $\{Q; U_1, \ldots, U_{Q-1}\}$ with distribution

$$g_{ij}(n) := P_i(T^j = n | T^j < T^i), \qquad n = 1, 2, \ldots.$$

From the above we obtain the following equivalent description of the structure of our regenerative cycles. We can assume that we have a sequence of i.i.d. random variables with distribution g_{ii} which we will designate by $\{U_l; l = 1, 2, \ldots\}$, a random variable V with distribution g_{ij}, and a geometric random variable (r.v.) Q with probability of success p. Furthermore, let $\{U_l\}$, V, and Q be independent. Then the collection of r.v.'s $\{Q; U_l, l = 1, 2, \ldots, Q-1, V\}$ describes the cycle completely.

We thus have $W = \sum_{l=1}^{Q-1} U_l + V$ and, taking expectations,

$$EW = E\left[\sum_{l=1}^{Q-1} U_l + V\right] = EUE[Q-1] + EV.$$

Since Q is geometric, $EQ = 1/p$ and $E[Q-1] = q/p$, from which we obtain

$$EW = \frac{q}{p} EU + EV. \tag{3}$$

We close the above discussion by noting that the overlapping estimator can also be written as

$$\Psi_n = \frac{\sum_{k=1}^{n}\left(V_k + \sum_{l=1}^{Q_k-1} lU_{l,k}\right)}{\sum_{k=1}^{n} Q_k}, \tag{4}$$

a fact that will be used in the sequel.

3. STRONG CONSISTENCY OF THE TWO ESTIMATORS FOR THE MEAN PASSAGE TIME IN MARKOV CHAINS

Recall that an estimator is strongly consistent if it converges with probability 1 to the quantity to be estimated as the number of observations goes to infinity.

PROPOSITION 1: *Both the simple and the overlapping estimator defined earlier are strongly consistent; that is, as $n \to \infty$,*

$$\Phi_n = \frac{1}{n}\sum_{k=1}^{n} W_{1,k} \to EW \quad w.p.\ 1 \tag{5}$$

and

$$\Psi_n = \frac{\sum_{k=1}^{n} V_k + \sum_{l=1}^{Q_k} lU_{l,k}}{\sum_{k=1}^{n} Q_k} \to \frac{E\left[V + \sum_{l=1}^{Q-1} lU_l\right]}{EQ} \quad w.p.\ 1, \tag{6}$$

with

$$\frac{E\left[V + \sum_{l=1}^{Q-1} lU_l\right]}{EQ} = EW. \tag{7}$$

PROOF: Equation (5) is an immediate consequence of the Strong Law of Large Numbers. (The irreducibility and positive recurrence of the Markov chain guarantee the finiteness of EW.) Thus, the simple estimator is clearly strongly consistent. The limit (6) is immediate if we divide both the numerator and denominator by the number of cycles n and appeal again to the Strong Law of Large Numbers. Thus, to establish the strong consistency of the overlapping estimator, it suffices to show (7). Indeed, the numerator on the left-hand side of (7) is

$$E\left[QV + \sum_{l=1}^{Q-1} lU_l\right] = EQEV + EUE\left[\sum_{l=1}^{Q-1} l\right]$$

(where we have used the independence of $\{U_l\}$, $V = U_Q$, and Q). Hence, the left-hand side of (7) becomes

$$EV + EU\,\frac{E[Q(Q-1)]}{2EQ}. \tag{8}$$

Since the *descending factorial moment of order m* of the geometric distribution is given by

$$E[Q(Q-1)(Q-2)\ldots(Q-m+1)] = \frac{m!\,q^{m-1}}{p^m}, \tag{9}$$

we easily see from (8) and (9) that

$$\frac{E\left[QV + \sum_{l=1}^{Q-1} lU_l\right]}{EQ} = EV + \frac{q}{p}\,EU. \tag{10}$$

The above, together with (3), establishes (7) and, hence, the strong consistency of the overlapping estimator. ■

Proposition 1 is established in Singh and Sutton [16] by means of a heuristic argument and in Bertsekas and Tsitsiklis [1]. It is included here for the sake of completeness of the discussion.

4. ASYMPTOTIC VARIANCE COMPARISON OF THE TWO ESTIMATORS

We now proceed to compare the two estimators in terms of their respective asymptotic variances. For the simple estimator, we have $\mathrm{Var}(\Phi_n) = (1/n)\mathrm{Var}(W)$ (due to the independence of cycles) and, hence,

$$\sigma_\Phi^2 := \lim_{n\to\infty} n\,\mathrm{Var}(\Phi_n) = \mathrm{Var}(W) = \mathrm{Var}\left(V + \sum_{l=1}^{Q-1} U_l\right).$$

Using the independence of $\{U_l\}$, V, and Q, we have

$$\mathrm{Var}\left(V + \sum_{l=1}^{Q-1} U_l\right) = \mathrm{Var}(V) + \mathrm{Var}(U)E(Q-1) + (EU)^2\,\mathrm{Var}(Q),$$

and since Q is geometrically distributed,

$$\sigma_\Phi^2 = \mathrm{Var}(V) + \frac{q}{p}\,\mathrm{Var}(U) + \frac{q}{p^2}\,(EU)^2. \tag{11}$$

Thus, we conclude that the simple estimator is unbiased and its asymptotic variance constant is given by the above expression.

We now turn our attention to the overlapping estimator, which is a *ratio estimator*. We saw in Section 3 that it is strongly consistent. The following theorem (which is an adaptation of a classic result regarding asymptotic statistics; e.g., see Cramér [5, p. 353] and Prakasa Rao [13, p. 146]) shows that it is also asymptotically unbiased and gives its asymptotic variance.

THEOREM 1 (Ratio estimator): *Let (ξ_k, ζ_k), $k = 1, 2, \ldots$, be i.i.d. vectors such that, for some $\delta > 0$, $P(|\zeta| > \delta) = 1$ and $E\zeta \neq 0$. We assume that both ξ and ζ have finite second moments. Then, as $n \to \infty$, the ratio estimator*

$$R_n = \frac{\sum_{k=1}^{n} \xi_k}{\sum_{k=1}^{n} \zeta_k}$$

converges with probability 1 *to*

$$r := \frac{E\xi}{E\zeta}.$$

Furthermore, R_n is asymptotically unbiased and its bias is $O(n^{-1})$; in fact,

$$\lim_{n\to\infty} n(ER_n - r) = r\left(\frac{\mathrm{Var}(\zeta)}{(E\zeta)^2} - \frac{\mathrm{Cov}(\xi,\zeta)}{E\xi E\zeta}\right). \tag{12}$$

Finally, the limit

$$\lim_{n\to\infty} n\,\mathrm{Var}(R_n) =: \sigma^2 \tag{13}$$

exists and is given by the expression

$$\sigma^2 = \frac{\mathrm{Var}(\xi) + r^2\,\mathrm{Var}(\zeta) - 2r\,\mathrm{Cov}(\xi,\zeta)}{(E\zeta)^2}$$

$$= \frac{E[\xi^2] + r^2E[\zeta^2] - 2rE[\xi\zeta]}{(E\zeta)^2}. \tag{14}$$

The limit (13) is the *asymptotic variance constant* and it characterizes the behavior of the estimator when the number of cycles is large, since, in that case, $\mathrm{Var}(R_n) = (\sigma^2/n) + o(1/n)$. Based on Theorem 1, and in view of the fact that MSE = (bias)^2 + variance, we realize that the contribution of the bias term to the MSE is asymptotically negligible compared to that of the variance. Thus, we can ignore the bias term in our comparison of the two estimators.

We now turn our attention to the asymptotic variance of the two estimators.

THEOREM 2: *In any irreducible, positive recurrent Markov chain with countable state space, the simple estimator of the mean passage time for any pair of states i and j is superior to the overlapping estimator in terms of asymptotic variance and in fact it holds that the asymptotic variance ratio (AVR) is given by*

$$\mathrm{AVR} := \frac{\sigma_\Psi^2}{\sigma_\Phi^2} = 1 + q, \tag{15}$$

where $\sigma_\Psi^2 := \lim_{n\to\infty} n\,\mathrm{Var}\,\Psi_n$ *and q is given by (2).*

PROOF: Applying Theorem 1 (which guarantees the existence of σ_Ψ) and taking into account Proposition 1, we obtain

$$\sigma_\Psi^2 = \frac{E\left[\left(V + \sum_{l=1}^{Q-1} lU_l\right)^2\right] + r^2E[Q^2] - 2rE\left[QV + Q\sum_{l=1}^{Q-1} lU_l\right]}{(EQ)^2}, \tag{16}$$

with

$$r = EV + \frac{q}{p}EU. \tag{17}$$

In order to be able to compare directly the asymptotic variance constant of the overlapping estimator, σ_Ψ^2, with that of the simple one, given in (11), we need to express (16) in terms of p and the moments of U and V. Since Q is geometric, we have

$$E[Q^2] = \frac{1+q}{p^2}. \tag{18}$$

The "covariance" term in (16) can also be easily computed taking into account the independence of Q, V, and $\{U_l\}$ as follows:

$$E\left[QV+\sum_{l=1}^{Q-1} lU_l\right]=E[Q^2V]+E\left[Q\sum_{l=1}^{Q-1} lU_l\right]$$
$$=E[Q^2]EV+EUE\left[Q\frac{1}{2}Q(Q-1)\right].$$

Writing $\frac{1}{2}E[Q^2(Q-1)]=\frac{1}{2}E[Q(Q-1)(Q-2)]+E[Q(Q-1)]$ and using (9), we find

$$E\left[QV+Q\sum_{l=1}^{Q-1} lU_l\right]=\frac{1+q}{p^2}EV+\frac{q(2+q)}{p^3}EU. \tag{19}$$

Finally,

$$E\left(QV+\sum_{l=1}^{Q-1} lU_l\right)^2=E(QV)^2+2E\left[QV\sum_{l=1}^{Q-1} lU_l\right]+E\left(\sum_{l=1}^{Q-1} lU_l\right)^2. \tag{20}$$

Let us examine each of the three terms on the right-hand side of (20). For the first, we have

$$E(QV)^2=E[V^2]E[Q^2]=E[V^2]\frac{1+q}{p^2}; \tag{21}$$

for the second,

$$E\left[QV\sum_{l=1}^{Q-1} lU_l\right]=(EV)(EU)E\left[Q\sum_{l=1}^{Q-1} l\right]$$
$$=(EV)(EU)E\left[\frac{1}{2}Q^2(Q-1)\right]$$
$$=(EV)(EU)\frac{q(2+q)}{p^3}, \tag{22}$$

and, finally, for the third,

$$E\left[\left(\sum_{l=1}^{Q-1} lU_l\right)^2\right]=E\left[\sum_{l=1}^{Q-1} l^2U_l^2\right]+2E\left[\sum_{l=2}^{Q-1}\sum_{m=1}^{l-1} lmU_lU_m\right]. \tag{23}$$

The first term on the right-hand side of (23) is equal to

$$\begin{aligned} E[U^2]E\left[\sum_{l=1}^{Q-1} l^2\right] &= E[U^2]E\left[\frac{1}{3}Q(Q-1)(Q-2) + \frac{1}{2}Q(Q-1)\right] \\ &= E[U^2]\frac{q(1+q)}{p^3}, \end{aligned} \tag{24}$$

where we have used, once more, the expression for the descending factorial moments of the geometric distribution (9). The second term on the right-hand side of (23) is

$$\begin{aligned} 2(EU)^2E\left[\sum_{l=2}^{Q-1} l \sum_{m=1}^{l-1} m\right] &= (EU)^2E\left[\sum_{l=2}^{Q-1} l(l-1)(l-2) + 2l(l-1)\right] \\ &= (EU)^2\left(\frac{1}{4}E[Q(Q-1)(Q-2)(Q-3)]\right. \\ &\qquad \left. + \frac{2}{3}E[Q(Q-1)(Q-2)]\right) \\ &= (EU)^2\left(\frac{4!q^3}{4p^4} + 2\frac{3!q^2}{3p^3}\right) \\ &= (EU)^2\frac{2q^2(2+q)}{p^4} \end{aligned} \tag{25}$$

Hence, from (20)–(25), we have

$$\begin{aligned} E\left[QV + \left(\sum_{l=1}^{Q-1} lU_l\right)^2\right] &= E[V^2]\frac{1+q}{p^2} + 2EVEU\frac{q(2+q)}{p^3} \\ &\quad + E[U^2]\frac{q(1+q)}{p^3} + (EU)^2\frac{2q^2(2+q)}{p^4}. \end{aligned} \tag{26}$$

Thus, using (16)–(26), after carrying out the necessary simplifications we obtain the following expression for the asymptotic variance:

$$\sigma_\Psi^2 = \mathrm{Var}(U)\frac{(1+q)q}{p} + \mathrm{Var}(V)(1+q) + (EU)^2\frac{q(1+q)}{p^2}. \tag{27}$$

Comparing the above expression for the asymptotic variance constant of the overlapping estimator with that of the simple estimator, given in (11), we obtain (15). ■

Thus, for the mean passage time estimation, the simple estimator is always superior to the overlapping one in terms of asymptotic variance, despite the fact that it might appear more wasteful. Although the overlapping estimator obtains more measurements over any given sample path than the simple estimator, these measurements are obviously strongly correlated. When q is close to zero (i.e., when returning to i before visiting j is very unlikely), the two estimators are essentially the same and their asymptotic variances are nearly equal. When, on the other hand, q is close to 1, then the overlapping estimator has asymptotic variance that is nearly twice that of the simple estimator.

We should point out that the above analysis is asymptotic and might not provide the complete picture if one is interested in the small sample behavior of the two estimators. In that case, one could use the mean squared error as a comparison criterion. For the simple estimator, which is unbiased, this is equal to $\mathrm{Var}(\Phi_n)$. However, for the overlapping estimator, which is only asymptotically unbiased and whose variance was obtained asymptotically using Theorem 1, the situation is more complicated. Some results on this are provided in Singh and Sutton [16] and Bertsekas and Tsitsiklis [1, p. 186].

5. ESTIMATING THE DISTRIBUTION OF THE PASSAGE TIME

Consider now the same problem where the object is to estimate the probability that the passage time exceeds a given threshold, $\nu \in \mathbb{N}$ [i.e., the probability $P_i(T^j > \nu)$]. Using the simple estimator, this is estimated by

$$\Phi_n = \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}(W_{1,k} > \nu),$$

whereas the overlapping estimator is given by

$$\Psi_n = \frac{\sum_{k=1}^{n} \sum_{l=1}^{Q_k} \mathbf{1}(W_{l,k} > \nu)}{\sum_{k=1}^{n} Q_k}.$$

It will be convenient to use the notation

$$\xi_{l,k} = \mathbf{1}(W_{l,k} > \nu). \tag{28}$$

Also, as in the previous sections, in order to simplify the notation we will drop the subscript designating the cycle when not necessary. Thus, for instance, ξ_l can be used instead of $\xi_{l,k}$. Clearly, from the Strong Law of Large Numbers, Φ_n is a strongly consistent estimator of

$$r := E\xi_1 = P(W_1 > \nu). \tag{29}$$

Also, let

$$\beta_m := E[\xi_1 | Q = m] = E\left[\mathbf{1}\left(\sum_{l=1}^{m-1} U_l + V > \nu\right)\right], \qquad m = 1, 2, \ldots. \tag{30}$$

Then $E[\xi_l | Q = m] = \beta_{m+1-l}$, where $l = 1, 2, \ldots, m$. Note in particular that

$$r = \sum_{m=1}^{\infty} pq^{m-1}\beta_m. \tag{31}$$

Finally, define the generating function $B(x) := \sum_{m=1}^{\infty} x^m \beta_m$. Note in particular that $B(q) = qp^{-1}\sum_{m=1}^{\infty} pq^{m-1}\beta_m = qp^{-1}r$.

THEOREM 3: *With the above definition, we have the following:*

(i) Ψ_n *is also a strongly consistent estimator of r.*
(ii) The AVR of the two estimators is given by

$$\text{AVR} := \frac{\sigma_\Psi^2}{\sigma_\Phi^2} = \frac{1 + q + pr - 2p^2 B'(q)}{1 - r}. \tag{32}$$

(iii) An alternative expression for the AVR is

$$\text{AVR} = p + 2q\,\frac{1 - \tilde{r}}{1 - r}, \tag{33}$$

where $\tilde{r} := \sum_{m=2}^{\infty}(m-1)p^2 q^{m-2}\beta_m$.

PROOF: By virtue of the Strong Law of Large Numbers, we have, as $n \to \infty$,

$$\Psi_n \to \frac{E\sum_{l=1}^{Q}\xi_l}{EQ} \quad \text{w.p. } 1. \tag{34}$$

The numerator of the right-hand side of (34) can be computed by conditioning on Q as follows:

$$\begin{aligned}
\sum_{m=1}^{\infty} pq^{m-1} \sum_{l=1}^{m} E[\xi_l | Q = m] &= \sum_{m=1}^{\infty} pq^{m-1} \sum_{l=1}^{m} \beta_{m-l+1} \\
&= \sum_{l=1}^{\infty} pq^{l-1} \sum_{m=l}^{\infty} q^{m-l}\beta_{m-l+1} \\
&= p^{-1} \sum_{m=1}^{\infty} \beta_m pq^{m-1} \\
&= p^{-1} r,
\end{aligned}$$

the last equation following as a result of definitions (30) and (29). This, together with the fact that $EQ = p^{-1}$, establishes the strong consistency of the overlapping estimator. Regarding the asymptotic variances, we first note that

$$n\,\text{Var}(\Phi_n) = r(1 - r). \tag{35}$$

On the other hand, the asymptotic variance of the overlapping estimator can be computed as in Section 4 by first computing EY^2 and EYQ, where $Y = \sum_{l=1}^{Q} \xi_l$. This is accomplished by first conditioning on Q, as follows:

$$
\begin{aligned}
E[Y^2 | Q = m] &= E\left[\sum_{l=1}^{m} \xi_l \sum_{l'=1}^{m} \xi_{l'} | Q = m \right] \\
&= E\left[\sum_{l=1}^{m} \xi_l^2 | Q = m \right] + 2E\left[\sum_{l=2}^{m} \sum_{l'=1}^{l-1} \xi_l \xi_{l'} | Q = m \right] \\
&= E\left[\sum_{l=1}^{m} \xi_l | Q = m \right] + 2E\left[\sum_{l=2}^{m} \sum_{l'=1}^{l-1} \xi_l | Q = m \right] \\
&= E\left[\sum_{l=1}^{m} (2l-1) \xi_l | Q = m \right] \\
&= \sum_{l=1}^{m} (2l-1) E[\xi_l | Q = m] \\
&= \sum_{l=1}^{m} (2l-1) \beta_{m-l+1}. \qquad \textbf{(36)}
\end{aligned}
$$

In the above string of equalities, the third follows from the fact that ξ_l takes only the values zero and one and also from the fact that for $l' < l$, $\{W_l > \nu\} \subset \{W_{l'} > \nu\}$ and, thus, $\xi_l \xi_{l'} = \xi_l$. Therefore, taking expectation with respect to Q in (36), we have

$$
\begin{aligned}
EY^2 &= \sum_{m=1}^{\infty} pq^{m-1} \sum_{l=1}^{m} (2l-1) \beta_{m-l+1} \\
&= \sum_{l=1}^{\infty} (2l-1) pq^{l-1} \sum_{m=l}^{\infty} \beta_{m-l+1} q^{m-l} \\
&= \left(\frac{2}{p} - 1 \right) p^{-1} \sum_{n=1}^{\infty} \beta_n pq^{n-1} \\
&= \frac{2-p}{p^2} r. \qquad \textbf{(37)}
\end{aligned}
$$

Also,

$$\begin{aligned}
E[QY] &= \sum_{m=1}^{\infty} pq^{m-1} m \sum_{l=1}^{m} E[\xi_l | Q = m] \\
&= \sum_{m=1}^{\infty} pq^{m-1} m \sum_{l=1}^{m} \beta_{m-l+1} \\
&= \sum_{l=1}^{\infty} pq^{l-1} \sum_{m=l}^{\infty} q^{m-l} m\beta_{m-l+1} \\
&= \sum_{l=1}^{\infty} pq^{l-1} \sum_{m=l}^{\infty} q^{m-l}(m-l+1)\beta_{m-l+1} + \sum_{l=1}^{\infty} (l-1)pq^{l-1} \sum_{m=l}^{\infty} q^{m-l}\beta_{m-l+1} \\
&= B'(q) + E[Q-1]p^{-1} \sum_{n=1}^{\infty} pq^{n-1}\beta_n \\
&= B'(q) + r\frac{1-p}{p^2}.
\end{aligned}$$

From Theorem 1 we have

$$\lim_{n\to\infty} n \operatorname{Var}(\Psi_n) = \frac{1}{(EQ)^2}(EY^2 + r^2 EQ^2 - 2rEYQ), \tag{38}$$

where $r = EY/EQ$, $EQ = 1/p$, and $EQ^2 = (1+q)/p^2$. Substituting the above into (38), we obtain

$$\lim_{n\to\infty} n \operatorname{Var}(\Psi_n) = r(1+q) + pr^2 - 2rp^2 B'(q). \tag{39}$$

Thus, dividing the above expression by $r(1-r)$, we establish (32).

The alternative expression for AVR is obtained by straightforward algebraic manipulations, which we omit. ■

In closing this section we should point out the significance of $\tilde{r}$ in (33). It is obtained by averaging the β_m's not with the geometric distribution with probability of success p (which is the distribution of Q) as in (31), but with a *negative binomial distribution*, which expresses the sum of *two* such geometric random variables. Since β_m is increasing in m, as can be seen by its definition, it follows from a simple stochastic ordering argument that $\tilde{r} > r$ and, hence, that $(1-\tilde{r})/(1-r) < 1$. Thus, we have

$$\text{AVR} < 1 + q.$$

This should be compared with the corresponding equality (3) of Theorem 2. As we will see in Section 6 by means of an example, there are situations where AVR < 1, thus making the overlapping estimator more efficient than the simple estimator for large samples, unlike the situation where one tries to estimate the mean passage time.

Finally, we point out that the above results can readily be extended to continuous-time Markov chains and to Markov-renewal processes. Also, the target state j could easily be replaced by a set of states, J.

6. A TWO-STATE MARKOV CHAIN

To obtain more insight into this situation, consider a caricature of the problem we have examined, namely a discrete-time Markov chain with state space $\mathbb{S} = \{0, 1\}$ and transition probability matrix

$$P = \begin{bmatrix} q & p \\ 1 & 0 \end{bmatrix}.$$

We want to estimate the mean passage time from zero to one, which, in this case, is a geometric random variable with probability of success p. Therefore, the variance constant for the simple estimator is

$$\sigma_\Phi^2 = \frac{q}{p^2}. \tag{40}$$

For the overlapping estimator we have

$$r = \frac{E[Q(Q+1)]}{2EQ} = \frac{1}{p}$$

and the asymptotic variance constant is given by

$$\sigma_\Psi^2 = \frac{\frac{1}{4} E[Q^2(Q+1)^2] + \frac{1}{p^2} E[Q^2] - 2 \frac{1}{2p} E[Q^2(Q+1)]}{(EQ)^2}.$$

Using again the expressions for the moments of the geometric distribution, we obtain, after some algebraic manipulations,

$$\sigma_\Psi^2 = \frac{q(1+q)}{p^2}. \tag{41}$$

We see that the ratio of (40) to (41) satisfies (15).

However, it should be pointed out that although the overlapping estimator for the mean passage time is inefficient, the situation is different if the object is to estimate the distribution of the passage time. For instance, suppose that we are inter-

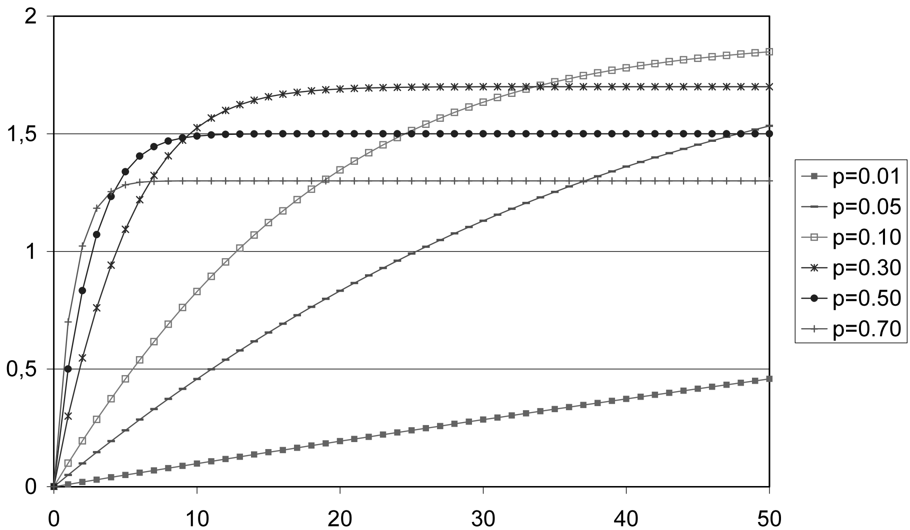

FIGURE 3. Ratio of asymptotic variances for the two-state Markov chain as a function of ν for various values of p.

ested in the probability that the time required to make the transition from zero to one is strictly greater than ν. In this case, $W_1 = Q$ and

$$r = P(W_1 > \nu) = P(Q > \nu) = q^\nu.$$

Also, $\beta_m = P(W_1 > \nu | Q = m) = \mathbf{1}(m \geq \nu + 1)$ and $B(x) = (1 - x)^{-1} x^{\nu+1}$. In that case straightforward algebraic manipulations give the following expression for the value of the AVR:

$$\text{AVR} = 1 + q - \frac{2\nu p q^\nu}{1 - q^\nu}. \tag{42}$$

Figure 3 shows plots of the ratio as a function of ν for various values of p. As we can see, for small values of ν and small values of p, the value of the ratio in (42) is less than one and thus the overlapping estimator has a smaller asymptotic variance than the simple estimator.

References

1. Bertsekas, D.P. & Tsitsiklis, J.N. (1996). *Neuro-dynamic programming*. Nashua, NH: Athena Press.
2. Calvin, J.M. & Nakayama, M.K. (1998). Using permutations in regenerative simulations to reduce variance. *ACM Transactions on Modelling and Computer Simulation* 8(2): 153–193.
3. Cao, X.R. & Chen, H.F. (1997). Perturbation realization, potentials, and sensitivity analysis of Markov processes. *IEEE Transactions on Automatic Control* 42(10): 1382–1393.

4. Cho, G.E. & Meyer, C.D. (2000). Markov chain sensitivity measured by mean first passage times. *Linear Algebra and its Applications* 316: 21–28.
5. Cramér, H. (1946). *Mathematical methods of statistics*. Princeton, NJ: Princeton University Press.
6. Glynn, P.W. & Iglehart, D.L. (1989). Importance sampling for stochastic simulations. *Management Science* 35(11): 1367–1392.
7. Harrison, P.G. & Knottenbelt, W.J. (2002). Passage time distributions in large Markov chains. In *Proceedings ACM SIGMETRICS*, June 2002, Marina Del Rey, CA, pp. 77–85.
8. Heidergott, B. & Hordijk, A. (2003). Taylor series expansions for stationary Markov chains. *Advances in Applied Probability* 35: 1046–1070.
9. Karamichalakou, C. & Zazanis, M.A. (2000). Asymptotic variance properties of passage time estimators for Markov chains. Technical report, Athens University of Economics and Business, Athens, Greece.
10. Kemperman, J.H.B. (1961). *The passage problem for a stationary Markov chain*. Chicago: University of Chicago Press.
11. Meyer, C.D. (1994). Sensitivity of the stationary distribution of a Markov chain. *SIAM Journal of Matrix Analysis and Applications* 15: 715–728.
12. Nakayama, M.K., Goyal, A., & Glynn, P.W. (1994). Likelihood ratio sensitivity analysis for Markovian models of highly dependable systems. *Operations Research* 42: 137–157.
13. Prakasa Rao, B.L.S. (1987). *Asymptotic theory of statistical inference*. New York: Wiley.
14. Ross, S.M. & Schechner, Z. (1985). Using simulation to estimate first passage distribution. *Management Science* 31(2): 224–234.
15. Shahabuddin, P. (1994). Importance sampling for highly reliable Markovian systems. *Management Science* 40(3): 333–352.
16. Singh, S.P., & Sutton, R.S. (1996). Reinforcement learning with replacing eligibility traces. *Machine Learning* 22: 123–158.
17. Syski, R. (1992). *Passage times for Markov chains*. Amsterdam: IOS Press.