Asymptotic Formulas for Markov Processes with Applications to Simulation

Ward Whitt

Operations Research, Volume 40, Issue 2 (Mar. - Apr., 1992), 279-291.

Stable URL:
http://links.jstor.org/sici?sici=0030-364X%28199203%2F04%2940%3A2%3C279%3AAFFMPW%3E2.0.CO%3B2-L

ASYMPTOTIC FORMULAS FOR MARKOV PROCESSES WITH APPLICATIONS TO SIMULATION

WARD WHITT

AT&T Bell Laboratories, Murray Hill, New Jersey

(Received July 1989; revision received October 1990; accepted January 1991)

The simulation run length required to achieve desired statistical precision for a sample mean in a steady-state stochastic simulation experiment is largely determined by the asymptotic variance of the sample mean and, to a lesser extent, by the second-order asymptotics of the variance and the asymptotic bias. The asymptotic variance, the second-order asymptotics of the variance, and the asymptotic bias of the sample mean of a function of an ergodic Markov process can be expressed in terms of solutions of Poisson's equation, as indicated by positive recurrent potential theory. We review this positive recurrent potential theory, giving special attention to continuous-time Markov chains. We provide explicit formulas for birth-and-death processes and diffusion processes, and recursive computational procedures for skip-free chains. These results can be used to help design simulation experiments after approximating the stochastic process of interest by one of the elementary Markov processes considered here.

This paper was motivated by our desire to better understand how to plan simulation experiments. We have in mind stochastic simulations in which we estimate steady-state quantities by sample means. To design the experiment, we need to have some idea about the simulation run lengths required to achieve the desired statistical precision. Of course, since we intend to simulate, we probably cannot directly analyze the model of interest prior to performing the simulation. Thus, we suggest approximating the model of interest by a more elementary model that can be analyzed analytically. In Whitt (1989) we show how reflecting Brownian motion (RBM) can be used to approximate various queueing models for this purpose. Earlier work along these lines was done by Blomqvist (1969), Moeller and Kobayashi (1974) and Woodside, Pagurek and Newell (1980). Further work has been done by Asmussen (1989, 1992).

We can also use other elementary Markov models in essentially the same way. For example, the Ornstein-Uhlenbeck diffusion process is a natural approximation for infinite-server queues; see Whitt (1982), Glynn and Whitt (1991) and the references cited there. Here we derive basic formulas and computational procedures for a large class of Markov processes. The main quantity is the asymptotic variance of the sample mean because it determines the size of confidence intervals for large samples. While we were primarily motivated by simulation, the asymptotic variance is also important for estimation from direct system measurements.

There is substantial supporting theory for calculating the asymptotic variance, the second-order asymptotics of the variance, and the asymptotic bias of a sample mean of a function of an irreducible positive recurrent Markov process, in particular, the positive recurrent potential theory for Markov processes. The main idea is that these asymptotic quantities can be expressed in terms of solutions of Poisson's equation; e.g., see Proposition 10 and its Corollaries 3-5. For finite-state, discrete-time Markov chains (DTMCs), this is the familiar theory associated with the fundamental matrix Z in Chapters 4 and 5 of Kemeny and Snell (1960); the asymptotic variance of the sample mean is given in Corollary 4.6.2 there. With minor modification, the positive recurrent potential theory for DTMCs carries over to continuous-time Markov chains (CTMCs) too, as was pointed out by Kemeny and Snell (1961). Potential theory for MCs has many applications, e.g., it also plays an important role in Markov decision processes (MDPs), as was first shown by Howard (1960). Relevant discussions appear in Miller and Veinott (1969), Veinott (1969) and Denardo (1972). The potential theory for CTMCs is also discussed in Chapters 2 and 7 of Keilson (1979). Extensions to denumerable DTMCs are contained in Chapter 9 of Kemeny, Snell and Knapp (1966). Extensions to Harris's recurrent DTMCs

Subject classifications: Probability, diffusion: the asymptotic variance of the sample mean. Probability, Markov processes: potential theory and Poisson's equation. Simulation, design of experiments: required run lengths.
Area of review: SIMULATION.

0030-364X/92/4002-0279 $01.25

with nondenumerable state spaces are contained in Glynn (1989a) and the references cited there. The recurrent potential theory for the M/G/1 queue is contained in Glynn (1989b). A key early reference for DTMCs is Doeblin (1938); see pp. 90 and 111 of Chung (1967) for the history.

Expressions for the asymptotic variance and the asymptotic bias of the sample mean of DTMCs and CTMCs are contained in Kemeny and Snell (1960, 1961) and Chung. Additional results and further discussion are contained in Keilson and Rao (1970), Hordijk, Iglehart and Schassberger (1976), Hazen and Pritsker (1980), Burman (1980), Grassmann (1982, 1987a) and Glynn (1984). We contribute primarily by providing a survey, i.e., by making clear the connections to Kemeny and Snell (1960, 1961) and the positive recurrent potential theory. We also provide explicit formulas for birth-and-death (BD) processes and diffusion processes. The explicit formulas facilitate deriving analytical expressions in special cases (see Examples 1–3), determining bounds (see Propositions 3 and 5) and doing calculations.

We begin in Section 1 by deriving an expression for the asymptotic variance of the sample mean of a BD process (see (6)). In Section 2 we calculate the fundamental matrix of a positive recurrent BD process and related quantities of interest, such as the asymptotic bias and the second-order asymptotics of the variance of the sample mean. We obtain continuous-state analogs of the BD formulas for diffusion processes in Section 3. These are of special interest because diffusion processes are natural candidates to approximate other stochastic processes. In Section 4 we review the positive recurrent potential theory for CTMCs; i.e., we show that the fundamental matrix of a CTMC is related to its generator via Poisson's equation. We also discuss recursive procedures for calculating solutions of Poisson's equations for BD processes and skip-free CTMCs. This discussion provides insight into the recursive procedure developed by Grassmann (1987a) for calculating the asymptotic variance of a sample mean of a BD process. Finally, we make some concluding remarks in Section 5.

Another approach to the asymptotic variance of the sample mean and related quantities of interest is the spectral representation. As discussed by Keilson, pp. 33 and 109, when a CTMC is reversible, there is a complete set of orthonormal eigenvectors associated with real eigenvalues, so that the transition function can be represented as a linear combination of exponentials. If we calculate the eigenvalues and eigenvectors, we can thus describe the full transient behavior, not just the asymptotic values. Halfin (1978, 1984) developed an efficient algorithm for calculating the eigenvalues and eigenvectors of a BD process (which is reversible), exploiting the special structure derived by Ledermann and Reuter (1954); see also Abate and Whitt (1989).

1. THE ASYMPTOTIC VARIANCE OF A SAMPLE MEAN FOR A BD PROCESS

Given a real-valued, continuous-time stochastic process $Y(t)$, the sample mean is

$$\bar{Y}(t) = t^{-1} \int_0^t Y(s)\, ds, \quad t > 0, \tag{1}$$

and its asymptotic variance is

$$\bar{\sigma}^2 = \lim_{t\to\infty} t \operatorname{Var}(\bar{Y}(t)). \tag{2}$$

If $Y(t)$ is a stationary process with $E[Y(t)^2] < \infty$, then it has a finite (auto)covariance function

$$\begin{aligned} R(t) &= \operatorname{Cov}[Y(0), Y(t)] \\ &= E[Y(0)Y(t)] - [EY(0)]^2, \quad t \geq 0. \end{aligned} \tag{3}$$

If R is integrable, then

$$\hat{\sigma}^2 = 2 \int_0^\infty R(t)\, dt; \tag{4}$$

e.g., see Chapter 3 of Parzen (1962). In many cases $Y(t)$ is not stationary, but is asymptotically stationary; then formula (4) for the stationary version typically provides the correct value. (A proof for the case of irreducible finite MCs appears in (3.10) of Glynn (1984).) The asymptotic variance is especially important because, together with the asymptotic mean $\bar{\nu} = \lim_{t\to\infty} E\bar{Y}(t)$, it typically determines the full asymptotic distribution of $\bar{Y}(t)$. Typically, $t^{1/2}(\bar{Y}(t) - \bar{\nu})$ is asymptotically normally distributed with mean 0 and variance $\bar{\sigma}^2$ as $t \to \infty$; e.g., see Theorem 20.1 of Billingsley (1968). (This covers functions of MCs as a special case, but the central limit for functions of MCs has a long history; see I.14-16 of Chung (1967) and Glynn (1989a). These results are for discrete-time processes, but corresponding results hold for continuous-time processes by the same argument.)

From (4), it is evident that we can calculate $\bar{\sigma}^2$ if we can calculate the covariance function $R(t)$, but the covariance function is often difficult to work with directly. (For positive results, see Glynn (1989c), Abate and Whitt (1988) and the references cited there.) Fortunately, it is often not necessary to calculate $\operatorname{Var}(\bar{Y}(t))$ or $R(t)$ explicitly in order to find the asymptotic value $\bar{\sigma}^2$ because simplification occurs in the limit. This simplification is well illustrated by considering the case of an ergodic BD process $X(t)$ on the subset of integers $\{0, 1, \ldots, n\}$ with positive birth rates λ_i, death rates μ_i and stationary probabilities $\pi_i = \pi_0(\lambda_0\lambda_1 \ldots \lambda_{i-1})/(\mu_1 \ldots \mu_i)$. (The process is reflecting at 0 and n; i.e., $\lambda_n = \mu_0 = 0$.) Let $Y(t) = f(X(t))$, where f is a real-valued function on the state space of the BD process $X(t)$. Let $f_i = f(i)$ and let $\bar{f}$ be the steady-state mean, i.e.,

$$\bar{f} = \sum_{i=0}^{n} \pi_i f_i. \tag{5}$$

The following convenient formula was derived by Burman (1980).

Proposition 1. *For a function of a BD process,*

$$\bar{\sigma}^2 = 2 \sum_{j=0}^{n-1} \frac{1}{\lambda_j \pi_j} \left[\sum_{i=0}^{j} (f_i - \bar{f})\pi_i \right]^2. \tag{6}$$

Example 1. Assuming that (6) remains valid for infinite-state BD processes under appropriate regularity conditions, which it does (see Remark 6), it is easy to apply it to obtain $\bar{\sigma}^2$ for the process recording the number in system in an $M/M/1$ queue with service rate 1 and arrival rate (and traffic intensity) $\rho < 1$; if $f_i = i$, then $\lambda_j = \rho$, $\pi_j = (1 - \rho)\rho^j$ and

$$\bar{\sigma}^2 = \frac{2\rho(1 + \rho)}{(1 - \rho)^4}, \tag{7}$$

whereas the full covariance function $R(t)$ is relatively complicated; see Morse (1955), p. 168 of Grassmann (1987a), Abate and Whitt (1988) and Whitt (1989).

We first give a relatively simple direct proof of (6). (Our proof is very different from Burman's (1980), which followed the lines of Burman (1979, 1981).) Indeed, (6) is a special case of the following formula for irreducible finite-state CTMCs, which seems to be of some independent interest. Let the state space again be $\{0, 1, \ldots, n\}$. Let T_{ij} be the first passage time from i to j (with $T_{ii} = 0$); again let π be the stationary probability vector and let $\bar{f}$ be as in (5).

Proposition 2. *For a function of a CTMC,*

$$\bar{\sigma}^2 = -2 \sum_{j=0}^{n} \sum_{i=0}^{n} (f_i - \bar{f})\pi_i ET_{ij}(f_j - \bar{f})\pi_j$$

$$= -2 \sum_{j=1}^{n} \sum_{i=0}^{j-1} (f_i - \bar{f})\pi_i(ET_{ij} + ET_{ji})(f_j - \bar{f})\pi_j. \tag{8}$$

Formula (6) follows immediately from (8) and the BD formula

$$ET_{ji} + ET_{ij} = \sum_{k=i}^{j-1} \frac{1}{\lambda_k \pi_k}, \quad i < j. \tag{9}$$

(Change the order of summation in (8) and use the fact that $\sum_{i=0}^{n} (f_i - \bar{f})\pi_i = 0$.) Formula (9) is an immediate consequence of the well known formula

$$ET_{j,j+1} = (1/\lambda_j\pi_j) \sum_{i=0}^{j} \pi_i; \tag{10}$$

see p. 93 Heyman and Sobel (1982). By changing the role of the birth-and-death rates, from (10) we obtain

$$ET_{j+1,j} = (1/\lambda_j\pi_j) \sum_{i=j+1}^{n} \pi_i. \tag{11}$$

Of course, (9) follows easily from (10) and (11).

Formula (8) in turn follows relatively directly from the expression for $\bar{\sigma}^2$ in terms of the fundamental matrix of a CTMC, i.e.,

$$\bar{\sigma}^2 = 2 \sum_i \sum_j f_i \pi_i Z_{ij} f_j, \tag{12}$$

where Z the fundamental matrix, defined by

$$Z_{ij} = \int_0^\infty [P_{ij}(t) - \pi_j] \, dt, \tag{13}$$

with $P_{ij}(t)$ being the transition function; see Kemeny and Snell (1961), (16) of Keilson and Rao (1970), Chapter 7 of Keilson (1979) and Glynn (1984). (The terminology is not consistent; we follow Keilson.) We discuss (13) further in Section 4; (12) follows easily from (3), (4) and (13). The DTMC analog of (12) is Corollary 4.6.2 of Kemeny and Snell (1960); it would look exactly like (12) if we worked with $Z_{ij} - \pi_j$ instead of Z_{ij}, which is a common convention in recurrent potential theory; see Chapter 9 of Kemeny, Snell and Knapp, and p. 482 of Denardo.

The key connection between (8) and (12) is provided by the CTMC formula

$$Z_{ij} = Z_{jj} - \pi_j ET_{ij}, \tag{14}$$

which is easy to derive from (13) (see the Appendix); (14) is the CTMC analog of Theorem 4.4.7 of Kemeny and Snell (1960), which appears on p. 102 of Kemeny and Snell (1961). (As noted before (9) there, (3) there holds when the asterisk is removed.) We obtain (8) from (12) and (14) by rewriting (12) as

$$\bar{\sigma}^2=\sum_{j=0}^{n}\sum_{i=0}^{n}(f_i-\bar{f})\pi_i Z_{ij}(f_j-\bar{f})$$

$$=2\sum_{j=0}^{n}\sum_{i=0}^{n}(f_i-\bar{f})\pi_i(Z_{jj}-\pi_j ET_{ij})(f_j-\bar{f}) \qquad (15)$$

$$=-2\sum_{j=0}^{n}\sum_{i=0}^{n}(f_i-\bar{f})\pi_i ET_{ij}(f_j-\bar{f})\pi_j.$$

In the first line of (15) we use the fact that $\sum_i \pi_i Z_{ij} = 0$ and $\sum_j Z_{ij} = 0$. (Other expressions for $\bar{\sigma}^2$ will be given later; see Corollary 3 to Propositions 10 and 12.)

Hence, we have the desired derivation of (6) and the generalization to CTMCs in (8). However, it is not evident that (6) is the most effective way to compute $\bar{\sigma}^2$ for a BD process. An attractive alternative is Grassmann's (1987a) recursive algorithm, which is a special case of the recursive solution to Poisson's equation for a BD process; see Remarks 1 and 2 below. A naive direct application of (6) can be unsuitable for calculating $\bar{\sigma}^2$ when some of the stationary probabilities π_j are very small. To avoid exponent overflow, it is natural to move the factor $1/\pi_j$ in (6) into the inner sum and exploit the fact that

$$\sum_{i=0}^{j}(f_i-\bar{f})\pi_i=-\sum_{j+1}^{n}(f_i-\bar{f})\pi_i. \qquad (16)$$

We thus rewrite (6) as

$$\bar{\sigma}^2=2\sum_{j=0}^{j^*}\frac{1}{\lambda_j}\left[\sum_{i=0}^{j}(f_{j-i}-\bar{f})\sqrt{\pi_{j-i}}\sqrt{\frac{\pi_{j-i}}{\pi_j}}\right]^2$$

$$+2\sum_{j=j^*+1}^{n-1}\frac{1}{\lambda_j}\left[\sum_{i=j}^{n}(f_i-\bar{f})\sqrt{\pi_i}\sqrt{\frac{\pi_i}{\pi_j}}\right]^2. \qquad (17)$$

The idea is to let j^* in (17) be such that π_{j^*} is nearly the maximum probability. If π is unimodal and j^* is the mode (or one of the modes), then π_{j-i}/π_j is nonincreasing in i for $j \leq j^*$ and π_i/π_j is nonincreasing in i for $j > j^*$. A sufficient condition for π to be unimodal is for π to be log-concave or log-convex, which occurs if λ_j/μ_{j+1} is decreasing or increasing; see pp. 70–73 of Keilson. Of course, unimodality is not essential to use (17).

We conclude this section with some applications of (6) and (17) to treat special cases and obtain bounds. We first show that the unimodality of π discussed above enables us to bound $\bar{\sigma}^2$ above, at least crudely, independent of π. For this purpose, let

$$K_f=\max\{|f_i-\bar{f}|:0\leq i\leq n\}$$

$$\leq\max\{|f_i|:0\leq i\leq n\} \qquad (18)$$

and

$$K_\lambda = 1/\min\{\lambda_i\colon 0 \leq i \leq n-1\}. \qquad (19)$$

The following bound depends on K_f, K_λ and n, but otherwise not on π.

Proposition 3. *If the stationary probability vector π of a BD process is unimodal, then*

$$\bar{\sigma}^2 \leq 2n^2K_f^2 K_\lambda.$$

Proof. Apply (17) to get

$$\bar{\sigma}^2\leq 2K_f^2K_\lambda\left(\sum_{j=0}^{j^*}\frac{1}{\pi_j}\left[\sum_{i=0}^{j}\pi_i\right]^2+\sum_{j=j^*+1}^{n-1}\frac{1}{\pi_j}\left[\sum_{i=j}^{n}\pi_j\right]^2\right).$$

Noting that

$$\sum_{i=0}^{j}\pi_i\leq(j+1)\pi_j \quad \text{for } j\leq j^*$$

and

$$\sum_{i=j}^{n}\pi_i\leq(n-j+1)\pi_j \quad \text{for } j\geq j^*,$$

we obtain

$$\bar{\sigma}^2\leq 2K_f^2K_\lambda\left(\sum_{j=0}^{j^*}(j+1)^2\pi_j+\sum_{j=j^*+1}^{n-1}(n-j+1)^2\pi_j\right)$$

$$\leq 2n^2K_f^2K_\lambda.$$

We can also apply Proposition 2 to obtain a bound on $\bar{\sigma}^2$ in terms of the mean first passage times for general CTMCs. For this purpose, let

$$K_T = \max\{ET_{ij}\colon 0 \leq i, j \leq n\}. \qquad (20)$$

Note that $K_T = \max\{ET_{0n}, ET_{n0}\}$ for a BD process.

Proposition 4. *For a CTMC,*

$$\bar{\sigma}^2 \leq 2K_T\left(\sum_{i=0}^{n}|f_i-\bar{f}|\,\pi_i\right)^2.$$

Proposition 3 is interesting, because if π is not unimodal, then $\bar{\sigma}^2$ can be arbitrarily large for a BD process with bounded n, as we show below. From Proposition 4, we see that the mean first passage times necessarily can be arbitrarily large too.

Example 2. To see that $\bar{\sigma}^2$ can be arbitrarily large for a BD process with bounded n, we consider a nearly decomposable BD process. In particular, suppose that $n = 2$, so that there are only three states. Consider the symmetric case with $\lambda_0 = \mu_2 = x$ and $\lambda_1 = \mu_1 = 1$; then $\pi_0 = \pi_2 = 1/(2+x)$ and $\pi_1 = x/(2+x)$. Let $f_i = i$, so that $\bar{f} = 1$ and $\bar{\sigma} = 4/x(2+x) \to \infty$ as $x \to 0$. Also,

$ET_{02} = ET_{20} = (2 + x)/x \to \infty$ as $x \to 0$. This example has a high $\bar{\sigma}^2$ and high mean first passage times for small x, because for small x the model is bistable, tending to remain in the states 0 and 2 a long time; i.e., $ET_{01} = ET_{21} = 1/x$.

To simplify the expressions further, let P_j be the cumulative distribution function associated with the stationary probability vector π, i.e.,

$$P_j = \sum_{i=0}^{j} \pi_i, \; 0 \leq j \leq n. \tag{21}$$

Example 3. To see the possible computational difficulties with (6), and for its intrinsic interest, consider an $M/M/c/0$ loss system with c servers, no extra waiting space, an individual service rate 1 and an arrival rate λ, so that the number in service is a BD process with $\lambda_j = \lambda$, $0 \leq j \leq c - 1$, and $\mu_j = j$, $1 \leq j \leq c$. A special case is $M/M/\infty$. These models, like $M/M/1$ in Example 1, serve as simple approximations for a large class of models of interest.

With $f_i = i$, it is well known that $\bar{f} = \lambda(1 - \pi_c)$ and the steady-state variance (variance of π) is $\sigma_f^2 = \bar{f} - \lambda\pi_c(c - \bar{f})$. From (6) we obtain

$$\begin{aligned}
\bar{\sigma}^2 &= 2\sum_{j=0}^{c-1} \frac{1}{\lambda\pi_j}\left[\sum_{i=0}^{j} (i - \lambda(1-\pi_c))\pi_i\right]^2 \\
&= 2\sum_{j=0}^{c-1} \frac{1}{\lambda\pi_j}\left[\sum_{i=0}^{j} \lambda\pi_{i-1} - \lambda\pi_i + \lambda\pi_i\pi_c\right]^2 \\
&= 2\lambda \sum_{j=0}^{c-1} \frac{1}{\pi_j}[\pi_c P_j - \pi_j]^2 \\
&= 2\lambda \sum_{j=0}^{c-1} \frac{1}{\pi_j}[\pi_c^2 P_j^2 - 2\pi_c\pi_j P_j + \pi_j^2] \\
&= 2\lambda\left[1 - 2c\pi_c + 2\pi_c \sum_{j=0}^{c-1}(1 - P_j) + \pi_c^2 \sum_{j=0}^{c-1} \frac{P_j^2}{\pi_j}\right] \\
&= 2\lambda\left[1 - 2c\pi_c + 2\pi_c\lambda(1-\pi_c) + \pi_c^2 \sum_{j=0}^{c-1} \frac{P_j^2}{\pi_j}\right], \\
&= 2\lambda\left[1 - 2\pi_c(c - \bar{f}) + \pi_c^2 \sum_{j=0}^{c-1} \frac{P_j^2}{\pi_j}\right],
\end{aligned} \tag{22}$$

where P_j is given by (21), which is an alternative to the formula in Section 6 of Beneš (1961), which was derived via the spectral representation. Formula (22) evidently dose not simplify as much as (7), but for the special case in which $c = \infty$, $\pi_c = 0$ and we obtain the well known $M/M/\infty$ result

$$\bar{\sigma}^2 = 2\sum_{j=0}^{\infty} \frac{1}{\lambda\pi_j}\left[\sum_{i=0}^{j} (\lambda\pi_{i-1} - \lambda\pi_i)\right]^2 = 2\lambda. \tag{23}$$

When $c < \infty$, we can compute with (22), but a direct computation with either (6) or (22) can be difficult because π_j in the denominator can be very small, e.g., $\pi_0 = 1/\sum_{i=0}^{c} (\lambda^i/i!)$ approaches $e^{-\lambda}$ as $c \to \infty$. To illustrate, let $c = 1{,}000$ and $\lambda = 400$; then the $M/M/c/0$ system behaves like an $M/M/\infty$ system with the same parameters. The steady-state mean and variance are approximately 400, so that π_j is negligible for j outside the interval [300, 500] and $\pi_0 \approx e^{-400}$. From (23), we know that $\bar{\sigma}^2 \approx 800$, but we have difficulty computing directly from (6). Similar problems occur when λ is closer to c.

However, π is unimodal with mode $[\lambda]$ or $[\lambda] + 1$, where $[x]$ is the greatest integer less than x, so that with parameters $c = 1{,}000$ and $\lambda = 400$ we set $j^* = 400$ in (17) and the computation is easy to carry out. (To see that the mode is indeed as claimed, recall that the property is well known for the Poisson distribution associated with the $M/M/\infty$ model, and for the $M/M/c$ model the $M/M/\infty$ stationary probability is just truncated and renormalized.)

Formula (22) is convenient for developing bounds and approximations. For example, using the crude bound $P_j^2 \geq P_j\pi_j$, we apply (22) to obtain a lower bound.

Proposition 5. *For the $M/M/c/0$ model,*

$$\begin{aligned}
\bar{\sigma}^2 &\geq \bar{\sigma}_L^2 \equiv 2\lambda[1 - 2\pi_c(c - \bar{f}) + \pi_c^2(c - \bar{f})] \\
&\geq \bar{\sigma}_{LL}^2 \equiv 2\lambda[1 - 2\pi_c(c - \bar{f})].
\end{aligned}$$

On the other hand, Beneš showed for this model that

$$0 \leq R(t) \leq \sigma_f^2 e^{-r_1 t} \quad \text{and} \quad 1 < r_1 < \frac{\bar{f}}{\sigma_f^2}, \tag{24}$$

from which follow the upper bounds

$$\bar{\sigma}^2 \leq \bar{\sigma}_U^2 \equiv 2\,\sigma_f^2 \leq 2\bar{f} \leq 2\lambda \tag{25}$$

and the approximation

$$\begin{aligned}
\bar{\sigma}^2 \approx \bar{\sigma}_{AP}^2 &\equiv \frac{2(\sigma_f^2)^2}{\bar{f}} \\
&= 2\lambda\left[1 - 2\pi_c(c - \bar{f}) - \pi_c + \frac{\lambda\pi_c^2(c - \bar{f})^2}{\bar{f}}\right] \\
&= 2\lambda\left[1 - 2\pi_c(c - \bar{f}) - \pi_c + \frac{\pi_c^2(c - \bar{f})^2}{1 - \pi_c}\right].
\end{aligned} \tag{26}$$

From (22), we see that the lower bounds evidently are good approximations when π_c is small. For example, consider the case $c = 100$ with $\pi_c = 0.01$, so that $\lambda = 84.06$ and $\bar{f} = 83.22$. Then $\bar{\sigma}_{LL}^2 = (0.664)\ 2\lambda$, $\bar{\sigma}_L^2 = (0.666)\ 2\lambda$, $\bar{\sigma}_{AP}^2 = (0.683)\ 2\lambda$, $\bar{\sigma}^2 = (0.706)\ 2\lambda$ and $\bar{\sigma}_\mu^2 = (0.822)\ 2\lambda$. On the other hand, suppose that we consider higher blocking with $\pi_c = 0.10$, so that $\lambda = 104.11$ and $\bar{f} = 93.7$ with $c = 100$. Then $\sigma_{LL}^2 = (-0.26)\ 2\lambda$, $\bar{\sigma}_L^2 = (-0.20)\ 2\lambda$, $\bar{\sigma}_{AP}^2 = (0.081)\ 2\lambda$, $\bar{\sigma}^2 = (0.103)\ 2\lambda$ and $\bar{\sigma}_U^2 = (0.270)\ 2\lambda$. In both cases, the Beneš approximation (26) performs reasonably well.

2. OTHER ASYMPTOTIC QUANTITIES FOR A BD PROCESS

Another way to derive (6) as well as other useful formulas for BD processes is to directly calculate the fundamental matrix Z in (13) for a BD process. Since $\sum_i \pi_i Z_{ij} = 0$ for all j for general CTMCs, from (14) we see that

$$Z_{jj} = \pi_j ET_{ej} = \pi_j \sum_{i=0}^{n} \pi_i ET_{ij} \tag{28}$$

and

$$Z_{ij} = \pi_j[ET_{ej} - ET_{ij}],\ i \neq j, \tag{29}$$

for general CTMCs, where T_{ej} is first passage time to j from equilibrium. (Formula (28) is the CTMC analog of Theorem 4.4.9 of Kemeny and Snell (1960).) From (10), (11), (14), (21) and (28), we obtain the following formula for BD processes.

Proposition 6. *For a BD process,*

$$\pi_j^{-1} Z_{jj} = ET_{ej} = \sum_{k=0}^{j-1} \frac{P_k^2}{\lambda_k \pi_k} + \sum_{k=j}^{n-1} \frac{(1-P_k)^2}{\lambda_k \pi_k} \tag{30}$$

and

$$Z_{ij} = \begin{cases} Z_{jj} - \pi_j \sum_{k=i}^{j-1} \frac{P_k}{\lambda_k \pi_k},\ i<j, \\ Z_{jj} - \pi_j \sum_{k=j}^{i-1} \frac{1-P_k}{\lambda_k \pi_k},\ i>j. \end{cases} \tag{31}$$

Example 1 Revisited. For the $M/M/1$ queue, $P_k = 1 - \rho^{k+1}$,

$$\pi_j^{-1} Z_{jj} = ET_{ej} = \frac{\rho^{-j}}{(1-\rho)^2} - \frac{(2j+1)}{(1-\rho)}$$

and

$$\frac{Z_{ij} - Z_{jj}}{\pi_j} = ET_{ij} = \begin{cases} \frac{\rho^{-j} - \rho^{-i}}{(1-\rho)^2} - \frac{(j-1)}{1-\rho},\ i<j \\ \frac{(i-j)}{(1-\rho)},\ i>j. \end{cases}$$

Of course, ET_{ij} is just $(i - j)$ times the mean busy period for $i > j$. In agreement with Corollaries 3.1.3 and 3.2.2 of Abate and Whitt (1987), $ET_{e0} = \rho/(1-\rho)^2$.

Next consider the *asymptotic bias* of the sample mean (1) for a function of a CTMC with an initial probability vector α, defined to be

$$\bar{\beta}_\alpha = \lim_{t\to\infty} t[E_\alpha(\bar{Y}(t)) - \bar{f}], \tag{32}$$

where E_α is the expectation starting with α. (In MDPs f is a reward function, $\bar{f}$ is the long-run average reward or gain rate and $\bar{\beta}_\alpha$ is the bias.) For a CTMC,

$$\begin{aligned} \bar{\beta}_\alpha &= \lim_{t\to\infty}\left[\int_0^t \sum_{i=0}^{n}\sum_{j=0}^{n} \alpha_i P_{ij}(s) f_j\, ds - t\bar{f}\right] \\ &= \lim_{t\to\infty} \sum_{i=0}^{n} \alpha_i \sum_{j=0}^{n} f_j \int_0^t [P_{ij}(s) - \pi_j]\, ds \\ &= \sum_{i=0}^{n}\sum_{j=0}^{n} \alpha_i Z_{ij} f_j. \end{aligned} \tag{33}$$

From (33) we see that it is easy to apply Proposition 6 to compute $\bar{\beta}_\alpha$. Alternatively, we can apply the argument in (15) to obtain an analog of (6).

Proposition 7. *For a function of a BD process,*

$$\begin{aligned} \bar{\beta}_\alpha &= \sum_{i=0}^{n}\sum_{j=0}^{n} (\alpha_i - \pi_i) Z_{ij} (f_j - \bar{f}) \\ &= -\sum_{i=0}^{n}\sum_{j=0}^{n} (\alpha_i - \pi_i) ET_{ij} (f_j - \bar{f}) \pi_j \\ &= \sum_{j=0}^{n-1} \frac{1}{\lambda_j \pi_j} \sum_{i=0}^{j} (f_i - \bar{f}) \pi_i \sum_{i=0}^{j} (\alpha_i - \pi_i). \end{aligned} \tag{34}$$

The variance of the sample mean of a function of a CTMC under stationary initial conditions can be expressed as

$$\begin{aligned} \mathrm{Var}(\bar{Y}(t)) &= 2t^{-2} \int_0^t (t-s) R(s)\, ds \\ &= \frac{\bar{\sigma}^2}{t} - \frac{\bar{\gamma}}{t^2} + O(e^{-\eta t}), \end{aligned} \tag{35}$$

where η is some positive constant and

$$\bar{\gamma}=2\sum_{i=0}^{n}\sum_{j=0}^{n} f_i\pi_i(Z^2)_{ij}f_j; \tag{36}$$

see (3.5) of Glynn (1984), noting that further simplification occurs there (write $P(t) = \Pi + (P(t) - \Pi)$ and $F = Z + \Pi$). We can thus apply Proposition 6 to compute the second-order term $\bar{\gamma}$ as well as the first-order term $\bar{\sigma}^2$. Alternatively, we can apply (14), (15) and (28) to rewrite (36) as

$$\begin{aligned}\bar{\gamma}&=2\sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=0}^{n}(f_i-\bar{f})\\&\quad\cdot\pi_i[\pi_k ET_{ik}(ET_{ej}-ET_{kj})](f_j-\bar{f})\pi_j\\&=2\left(\sum_{i=0}^{n}\sum_{j=0}^{n}(f_i-\bar{f})\pi_i\pi_j ET_{ij}\right)\left(\sum_{i=0}^{n}\sum_{j=0}^{n}(f_i-\bar{f})\pi_i\pi_j ET_{ji}\right)\\&\quad-2\sum_{i=0}^{n}\sum_{j=0}^{n}\sum_{k=0}^{n}(f_i-\bar{f})\pi_i[\pi_k ET_{ik}ET_{kj}](f_j-\bar{f})\pi_j,\end{aligned} \tag{37}$$

which can aid calculation because ET_{ij} can be calculated recursively. However, (37) has the same problem for computation as (6). In Section 4 we will show how to calculate $\bar{\beta}_\alpha$, $\bar{\sigma}^2$ and $\bar{\gamma}$ recursively.

The appearance of Z^2 in (36) can be explained by considering the associated matrices

$$Z_{ij}^{(k)}\equiv\int_0^\infty t^k[P_{ij}(t)-\pi_j]\,dt, \tag{38}$$

which satisfy $Z_{ij}^{(k)} = k!\,(Z^k)_{ij}$; see p. 108 of Keilson.

3. DIFFUSION PROCESSES

Since diffusion processes on the real line are continuous-state analogs of BD processes, we should expect that continuous-state analogs of the BD formulas hold for diffusion processes, and indeed they do. In fact, the analog of (6) is given on p. 94 of Mandl (1968). This formula can also be derived easily using the diffusion analogs of (8)–(15) here, as we will show. For background on diffusion processes, see Mandl, and Chapter 15 of Karlin and Taylor (1981).

Let the diffusion process have drift coefficient $\mu(x)$ and strictly positive diffusion coefficient $\sigma^2(x)$, which we assume are piecewise continuous on the interval $[a, b]$. (These conditions cover almost all applications and, with them, it is easy to verify the results.) Paralleling our treatment of BD processes, we assume that the diffusion is reflecting at the boundary points a and b. If we want to let $a = -\infty$ or $b = +\infty$, then we assume that the boundary point is inaccessible (it cannot be reached in finite time). We assume that the transition function $p(t, x, y)$ converges to a proper stationary density $\pi(y)$. It is well known that

$$\pi(y) = \frac{m(y)}{M(b)},\ a \leq y \leq b, \tag{39}$$

where

$$m(y) = \frac{2}{\sigma^2(y)s(y)} \tag{40}$$

is the *speed density*,

$$s(y) = \exp\left\{-\int_a^y \frac{2\mu(x)}{\sigma^2(x)}\,dx\right\} \tag{41}$$

is the *scale density*, and

$$M(y) = \int_a^y m(x)\,dx, \quad a \leq y \leq b, \tag{42}$$

provided that the integrals in (41) and (42) are finite, which we assume; see Sections 15.3 and 15.5 of Karlin and Taylor or pp. 13 and 90 of Mandl. Moreover, if $T(x, y)$ is the first passage time from x to y, then

$$ET(x,y)=\begin{cases}\displaystyle\int_x^y M(u)s(u)\,du, & x<y\\ \displaystyle\int_x^y [M(b)-M(u)]s(u)\,du, & x\geq y,\end{cases} \tag{43}$$

see problem 15, p. 385, of Karlin and Taylor or p. 91 of Mandl, so that

$$ET(x, y) + ET(y, x) = M(b)\int_x^y s(u)\,du. \tag{44}$$

Paralleling (13), let

$$Z(x, y) = \int_0^\infty [p(t, x, y) = \pi(y)]\,dt. \tag{45}$$

Then, paralleling (14) and (28),

$$Z(x, y) = Z(y, y) - \pi(y)ET(x, y) \tag{46}$$

and

$$\begin{aligned}Z(y,y)&=\pi(y)ET(e,y)\\&\equiv\pi(y)\int_a^b \pi(x)ET(x,y)\,dx\end{aligned} \tag{47}$$

from which it is easy to establish the analogs of (6), (34) and (37), assuming that all integrals converge. Let $\bar{f} = \int_a^b f(x)\pi(x)\,dx$. Here are the analogs of (6) and (34).

Proposition 8. *For a function of a diffusion process,*

a. $$\bar{\sigma}^2 = 2\int_a^b \int_a^b f(x)\pi(x)Z(x,y)f(y)\,dx\,dy$$
$$= -2\int_a^b \int_a^b (f(x)-\bar{f})\pi(x)ET(x,y)$$
$$(f(y)-\bar{f})\pi(y)\,dx\,dy$$
$$= 2M(b)\int_a^b s(y)\left[\int_a^y (f(x)-\bar{f})\pi(x)\,dx\right]^2 dy$$
$$= 2\int_a^b \frac{2}{\sigma^2(y)\pi(y)}\left[\int_a^y (f(x)-\bar{f})\pi(x)\,dx\right]^2 dy. \quad (48)$$

b. $$\bar{\beta}_\alpha = \int_a^b \int_a^b \alpha(x)Z(x,y)f(y)\,dx\,dy$$
$$= -\int_a^b \int_a^b (\alpha(x)-\pi(x))ET(x,y)$$
$$\cdot (f(y)-\bar{f})\pi(y)\,dx\,dy$$
$$= \int_a^b \frac{2}{\sigma^2(y)\pi(y)}\int_a^y (f(x)-\bar{f})\pi(x)\,dx$$
$$\int_a^y (\alpha(x)-\pi(x))\,dx\,dy. \quad (49)$$

Example 4. Let $X(t)$ be reflected Brownian motion (RBM) on $[0, \infty)$ with drift coefficient -1 and diffusion coefficient 1. Then $\pi(x) = 2e^{-2x}$ and, if $f(x) = x$, then (48) gives $\bar{\sigma}^2 = 1/2$, which agrees with Corollary 9 to Theorem 1 of Abate and Whitt (1988).

Example 5. Let $X(t)$ be the Ornstein-Uhlenbeck (*O-U*) diffusion process on $(-\infty, \infty)$ with $\sigma^2(x) = \sigma^2$ and $\mu(x) = -\mu x$, where $\mu > 0$. Then $\pi(x) = n(x; \sigma^2/2\mu)$, where $n(x; c)$ is the normal density with mean 0 and variance c. Using the identity

$$\int_{-\infty}^y xn(x;1)\,dx = -n(y;1), \quad (50)$$

we obtain $\bar{\sigma}^2 = \sigma^2/\mu^2$ from (48), which is consistent with the well known formula $R(t) = (\sigma^2/2\mu)e^{-\mu t}$, $t \geq 0$.

To see that (48) is the continuous-state analog of (6), we can define a sequence of BD processes approximating the diffusion process; see Stone (1963). Let the nth BD process be defined on the finite set $\{a + k/n\colon 0 \leq k \leq [bn]\}$ and let the birth-and-death rates $\lambda_n(x)$ and $\mu_n(x)$ in state x be defined in terms of $\sigma^2(x)$ and $\mu(x)$ by

$$\lambda_n(a+k/n) = \frac{n^2\sigma^2(a+k/n) + n\mu(a+k/n)}{2} \quad (51)$$
$$\mu_n(a+k/n) = \frac{n^2\sigma^2(a+k/n) - n\mu(a+k/n)}{2},$$

so that the infinitesimal means and variances are

$$\frac{\lambda_n(a+k/n) - \mu_n(a+k/n)}{n} = \mu(a+k/n)$$
$$\frac{\lambda_n(a+k/n) + \mu_n(a+k/n)}{n^2} = \sigma^2(a+k/n), \quad (52)$$

which approach $\mu(x)$ and $\sigma^2(x)$ at all points of continuity as $n \to \infty$. By Stone, this implies weak convergence in function space (see Billingsley) of the BD processes to the diffusion process as $n \to \infty$.

The diffusion formulas in this section can now be obtained as the limit of the BD formulas as $n \to \infty$. For example, assuming that $a + x$ is a continuity point of $\sigma^2(x)$:

$$\prod_{k=1}^{[xn]} \frac{\lambda_n(a+(k-1)/n)}{\mu_n(a+k/n)}$$
$$= \frac{\lambda_n(a)}{\mu_n(a+[xn])}\exp\left\{\sum_{k=1}^{[xn]-1} \log\frac{\lambda_n(a+k/n)}{\mu_n(a+k/n)}\right\}$$
$$\to \frac{\sigma^2(a)}{\sigma^2(a+x)}\exp\left\{\int_a^{a+x} \frac{2\mu(y)}{\sigma^2(y)}\,dy\right\} \text{ as } n \to \infty. \quad (53)$$

Similary, (6) approaches (48).

While this limiting approach yields the correct answer, there is a technical gap in the argument. If $\{\{Y_n(t)\colon t \geq 0\}\colon n \geq 1\}$ is a sequence of stochastic processes converging weakly to another process $\{Y(t)\colon t \geq 0\}$ as $n \to \infty$, by the continuous mapping theorem it follows that the sample averages $\bar{Y}_n(t)$ converge weakly to $\bar{Y}(t)$ as $n \to \infty$ for each t. Moreover, if there is appropriate uniform integrability, which always holds if the state space is compact as assumed here, then $E\bar{Y}_n(t) \to E\bar{Y}(t)$ and Var $\bar{Y}_n(t) \to$ Var $\bar{Y}(t)$ as $n \to \infty$ for each t. These limits certainly provide a good basis for using the limiting processes as approximations, but the complete argument requires interchanging the limits $n \to \infty$ and $t \to \infty$.

4. FUNDAMENTAL MATRICES AND THE GENERATOR: POISSON'S EQUATION

So far, we have not mentioned the generator of the CTMC, but, of course, the fundamental matrix Z and the associated asymptotic quantities are determined by the generator. To discuss the generator of an irreducible finite-state CTMC, we will use matrix notation. Matrices will be denoted by capital letters and row vectors by lower case letters. Let x^t be the column vector obtained as the transpose of the row vector x, e the vector of all 1's, θ the vector of all 0's and e_i the unit vector having a 1 in the ith place and 0's elsewhere.

Let $P(t) \equiv (P_{ij}(t))$ be the transition function of an irreducible finite-state CTMC and let

$$Q = \lim_{t \downarrow 0} [P(t) - I] \tag{54}$$

be its generator. It is well known that $P_{ij}(t) \to \pi_j$ as $t \to \infty$ for all i and j; equivalently, we write $P(t) \to \Pi$, where Π is a matrix with all rows π. Moreover, π is the unique solution to the system of linear equations $xQ = 0$ with $xe^t = 1$; see p. 264 of Çinlar (1975). In this section, we discuss basic properties of Z for CTMCs. These properties are not new, but they have important consequences.

Our first result is an inverse matrix expression for the fundamental matrix Z due to Kemeny and Snell (1961). A nice proof is given on pp. 102–103 of Glynn (1984).

Proposition 9. *The integrals in* (13) *converge absolutely for all i and j and*

$$Z = (\Pi - Q)^{-1} - \Pi \tag{55}$$

with the inverse always existing.

From Proposition 9 perhaps it is not obvious how to exploit special structure in Q, such as occurs with a BD process, in order to calculate $(\Pi - Q)^{-1}$. Hence, we give another representation for Z, which shows how to calculate both π and Z with a single matrix inversion. In particular, we express Z as a solution to Poisson's equation; see Miller and Veinott (1969), Veinott (1969) and Denardo (1972). This representation shows how to calculate quantities related to Z recursively when the CTMC is skip-free. (By skip-free, we mean that either $Q_{ij} = 0$ for all $j > i + 1$ and all i or $Q_{ij} = 0$ for all $j < i - 1$ and all i.) Recall that e^t is the column vector of 1's.

Proposition 10. a. *Poisson's equation*

$$xQ = y \tag{56}$$

has a solution if and only if $ye^t = 0$. *All solutions are of the form*

$$x = -yZ + (xe^t)\pi. \tag{57}$$

b. *The alternate form of Poisson's equation*

$$Qx^t = y^t \tag{58}$$

has a solution if and only if $\pi y^t = 0$. *All solutions are of the form*

$$x^t = -Zy^t + (\pi x^t)e^t. \tag{59}$$

Proof. We only prove a because the proof of b is similar. Recall that $\theta = (0, 0, \ldots, 0)$. Since $P(t)e^t = e^t$, $Qe^t = \theta$ and $(xQ)e^t = 0$, so that $ye^t = 0$ is necessary. Henceforth, suppose that $ye^t = 0$. Note that (56) is equivalent to $x(\Pi - Q) = (xe^t)\pi - y$, which by Proposition 9 has the unique solution

$$x = ((xe^t)\pi - y)(Z + \Pi) = -yZ + (xe^t)\pi;$$

we have used $\pi Z = y\Pi = \theta$ and $\pi\Pi = \pi$.

Remark 1. For a BD process, (56) reduces to

$$x_{j-1}\lambda_{j-1} - x_j(\lambda_j + \mu_j) + x_{j+1}\mu_{j+1} = y_j, \quad j \geq 0, \tag{60}$$

where $x_{-1} = x_{n+1} = 0$, which in turn is equivalent to (add the first $j + 1$ equations)

$$x_{j+1} = (\lambda_j x_j + s_j)/\mu_{j+1}, \tag{61}$$

where $s_j = \sum_{i=0}^{j} y_i$, which has solution

$$x_j = \pi_j(t_j - \bar{t}), \tag{62}$$

where

$$t_j = \sum_{i=0}^{j-1} s_i/\lambda_i\pi_i, \; j \geq 1,$$

and $t_0 = 0$, provided that $\sum_{i=0}^{n} y_i = 0$.

Corollary 1. a. *Z is the unique solution of* $XQ = \Pi - I$ *and* $Xe^t = \theta^t$.

b. *Z is also the unique solution of* $QX = \Pi - I$ *and* $\pi X = \theta$.

Proof. Again we only prove a. Note that $(\Pi - I)e^t = \theta^t$, so that the rows of $\Pi - I$ are suitable y in (56). Then, from (57), $X = -(\Pi - I)Z = Z$.

Corollary 2. *The* $(n + 1) \times (n + 1)$ *matrix Q has rank n and the* $(n + 1) \times (n + 1)$ *matrix* Q_* *obtained by replacing the first column of Q with* e^t *has rank* $n + 1$. *Moreover, the first row of* Q_*^{-1} *is* π, *the jth row of* Q_*^{-1} *is the jth row of* $-Z$ *for* $j \geq 1$, *and the sum of the last* $n - 1$ *rows of* Q_*^{-1} *is the first row of Z.*

Proof. Since $Qe^t = 0$, Q has rank at most n. Proposition 10 implies that in fact Q has rank n and Q_* has rank $n + 1$. Solving $xQ_* = y_*$ where y_* is y in

(56) with the first element replaced by c is equivalent to solving (56) itself; the missing first equation $\sum_{i=0}^{n} x_i Q_{i0} = y_0$ can be obtained by adding the last n equations:

$$\sum_{i=0}^{n} x_i \sum_{j=1}^{n} Q_{ij} = \sum_{j=1}^{n} y_j.$$

Proposition 10 provides the solution for the many descriptive quantities of interest associated with CTMCs that can be characterized as the solution of the equations in (56) or (58). Several examples are given by Denardo, who establishes the generalization of Proposition 10 to semi-Markov processes. On the other hand, if a descriptive quantity is given in the form (57) or (59), then Proposition 10 shows that we can calculate it by solving (56) or (58). This is especially useful if we have a BD process or a skip-free CTMC, because then the equations can be solved recursively, as indicated for BD processes in Remark 1.

Corollary 3. a. *The asymptotic variance* $\bar{\sigma}^2$ *in* (12) *can be expressed as* $2xf^t$, *where* x *is the unique solution to*

$xQ = -y$ *with* $y_i = (f_i - \bar{f})\pi_i$ *and* $xe^t = 0$.

b. $\bar{\sigma}^2$ *can also be expressed as* $2 \sum_i f_i \pi_i x_i$, *where* x^t *is the unique solution of* $Qx^t = -f^t + (\pi f^t)e^t$ *with* $\pi x^t = 0$.

Proof. Apply Proposition 10 with (12).

Remark 2. Grassmann's (1987a) recursive algorithm for $\bar{\sigma}^2$ is essentially Corollary 3a plus (61).

Remark 3. Corollary 3 and Remark 1 provide an alternate proof of (6). Since $xe^t = 0$, $\bar{\sigma}^2$ can also be expressed as $2x(f^t - \bar{f}e^t)$. Then

$$\bar{\sigma}^2 = 2 \sum_{j=0}^{n} (f_j - \bar{f})\pi_j(t_j - \bar{t})$$

$$= 2 \sum_{j=0}^{n} (f_j - \bar{f})\pi_j \sum_{i=0}^{j-1} \frac{1}{\lambda_i \pi_i} \sum_{k=0}^{i} (f_k - \bar{f})\pi_k \qquad (63)$$

which becomes (6) after changing the order of summation.

Corollary 4. a. *The asymptotic bias* $\bar{\beta}_\alpha$ *in* (32) *can be expressed as* xf^t, *where* x *is the unique solution to* $xQ = -\alpha + \pi$ *with* $xe^t = 0$.

b. $\bar{\beta}_\alpha$ *can also be expressed as* αx^t, *where* x^t *is the unique solution of* $Qx^t = -f^t + \bar{f}e^t$ *with* $\pi x^t = 0$.

Proof. Apply Proposition 10 with (33).

Corollary 5. a. *The second-order term* $\bar{\gamma}$ *in* (35) *and* (36) *can be expressed as* $2xf^t$, *where* x *is the unique solution to* $xQ = -w + (we^t)\pi$ *with* $xe^t = 0$ *and* w *is the unique solution to* $wQ = -y$ *with* $y_i = (f_i - \bar{f})\pi_i$ *and* $we^t = 0$.

b. $\bar{\gamma}$ *can also be expressed as* $2 \sum_i f_i \pi_i x_i$, *where* x^t *is the unique solution to* $Qx^t = -w^t + (\pi w^t)e^t$ *with* $\pi x^t = 0$ *and* w^t *is the unique solution to* $Qw^t = -f^t + \bar{f}e^t$ *with* $\pi w^t = 0$.

Proof. Apply Proposition 10 twice with (36).

5. CONCLUDING REMARKS

Remark 4. For a DTMC with transition matrix P, the analog of the fundamental matrix (55) is $Z = (I - P + \Pi)^{-1}$, where Π is again the matrix all of whose rows are the stationary probability vector, and the analog of Poisson's equation (56) is $x(I - P) = y$; see Lemma 1 on p. 482 on Denardo. Equation (56) can be solved by transforming it to a DTMC using uniformization; i.e., for sufficiently small δ, $I + \delta Q \equiv \hat{P}$ is the transition function of a DTMC such that $x(I - \hat{P}) = -\delta y$ if and only if $xQ = y$; see p. 113 of Howard, and p. 102 of Kemeny and Snell (1961). Gross and Miller (1984) suggested calculating $E\bar{Y}(t)$ using uniformization. For other uses of uniformization, see Grassmann (1987b) and the references there.

Remark 5. When the state space is large and Q is sparse, we can exploit sparse matrix methods in storing the matrix and iteratively solving Poisson's equation (56). We can also enhance convergence by doing Gauss-Seidel iterations with a judicious ordering of the states; see van der Wal and Schweitzer (1987), Mitra and Tsoucas (1988), and Greenberg and Vanderbei (1990).

Remark 6. We have assumed that the state space of the CTMC is finite, but the results extend to positive recurrent CTMCs with infinite state spaces (denumerable and nondenumerable) too under appropriate regularity conditions; see Chapter 9 of Kemeny, Snell and Knapp (1966) and Glynn (1989a) for associated DTMC theory. We can apply the familiar regenerative argument used for DTMCs in I.14-16 of Chung (1967). With that approach it suffices to assume that

$$EC_{00}^2 < \infty \quad \text{and} \quad E_i\left[\int_0^{C_{00}} |f(X)(t))|\, dt\right]^2 < \infty,$$

where C_{00} is the regenerative cycle associated with state 0, i.e., the time required starting from state 0 to leave and return again. Then

$$\bar{\sigma}^2 = \frac{1}{EC_{00}} \operatorname{Var}\left[\int_0^{C_{00}} f(X(t))\, dt - \bar{f} C_{00}\right] \tag{64}$$

where $X(0)) = 0$. The associated method for computing $\bar{\sigma}^2$ in (64) and related quantities is via the absorbing chain theory using the 0-avoiding transition probabilities; e.g., see p. 121 of Kemeny and Snell (1960), and Hordijk, Iglehart and Schassberger (1976). Mandl (1968) uses a related regenerative argument to obtain (48).

We conclude by giving another expression for $\bar{\sigma}^2$ in terms of zero-avoiding probabilities. (Similar expressions exist for $\bar{\beta}_\alpha$.) Let ${}^kN_{ij}$ be the number of transitions from i to j before first hitting k. We apply the following continuous-time analog of Proposition 9-58 of Kemeny, Snell and Knapp to our finite state-space situation.

Proposition 11. *In a CTMC,*

a. $ET_{ik} + ET_{kj} - ET_{ij} = \dfrac{E^kN_{ij}}{\pi_j}$;

b. $ET_{ij} + ET_{ji} = \dfrac{E^iN_{jj}}{\pi_j}$.

Proof. Uniformize and then apply the DTMC result.

The following complements Proposition 1.

Proposition 12. *For a function of a CTMC,*

$$\bar{\sigma}^2 = -2 \sum_{j=0}^{n} \sum_{i=0}^{n} (f_i - \bar{f})\pi_i E^iN_{jj}(f_j - \bar{f})$$

$$= 2 \sum_{i=1}^{n} \sum_{j=1}^{n} (f_i - \bar{f})\pi_i (E^0N_{ij}(f_j - \bar{f}).$$

Proof. For the first line, use Proposition 11b and the second line of (8). For the second line, use Proposition 11a and the first line of (8). Alternatively, by Theorem 6.2.5 of Kemeny and Snell (1960),

$$Z_{ij} = \pi_j + E^0N_{ij} - \sum_{k=1}^{n} \pi_k E^0N_{kj}$$

$$- \pi_j \sum_{k=1}^{n} E^0N_{ik} + \pi_j \sum_{k=1}^{n} \sum_{j=1}^{n} \pi_k E^0N_{kj}.$$

Substituting into the first line of (15) completes the proof.

Proposition 12 is closer to the expressions derived by Hordijk, Iglehart and Schassberger.

APPENDIX

Proof of (14)

Let 1_A be the indicator function of the set A, let $x \wedge y = min\{x, y\}$, and let E_i be the expectation operator starting in state i. From (13),

$$Z_{ij} = \lim_{t\to\infty} \int_0^t [P_{ij}(s) - \pi_j]\, ds$$

$$= \lim_{t\to\infty} E_i \int_0^t [1_{\{j\}}(X(s)) - \pi_j]\, ds$$

$$= \lim_{t\to\infty} E\left(E_i\left[\int_0^t [1_{\{j\}}(X(s)) - \pi_j\, ds] \mid T_{ij}\right]\right)$$

$$= \lim_{t\to\infty} E\Bigg(-\pi_j(T_{ij} \wedge t)$$

$$+ E_j\left[\int_0^{t-(T_{ij}\wedge t)} [1_{\{j\}}(X(s))\, ds - \pi_j]\, ds \mid T_{ij}\right]\Bigg)$$

$$= \lim_{t\to\infty} \Bigg(-\pi_j E(T_{ij} \wedge t)$$

$$+ E\int_0^{t-(T_{ij}\wedge t)} [P_{jj}(s) - \pi_j(s)]\, ds\Bigg)$$

$$= -\pi_j ET_{ij} + \int_0^\infty [P_{jj}(s) - \pi_j]\, ds = -\pi_j ET_{ij} + Z_{jj}.$$

ACKNOWLEDGMENT

I thank David Burman, Richard Gibbens, Peter Glynn, Shlomo Halfin, Philip Heidelberger, Marty Reiman, Bob Vanderbei, and the referees for assistance.

REFERENCES

ABATE, J., AND W. WHITT. 1987. Transient Behavior of the M/M/1 Queue: Starting at the Origin. *Queueing Sys.* **2,** 41–65.

ABATE, J., AND W. WHITT. 1988. The Correlation Functions of RBM and M/M/1. *Stochastic Models* **4,** 315–359.

ABATE, J., AND W. WHITT. 1989. Spectral Theory for Skip-Free Markov Chains. *Prob. Engr. Inform. Sci.* **3,** 77–88.

ASMUSSEN, S. 1989. Validating the Heavy Traffic Performance of Regenerative Simulation. *Stochastic Models* **5,** 617–628.

ASMUSSEN, S. 1992. Queueing Simulation in Heavy Traffic. *Math. Opns. Res.* (to appear).

BENEŠ, V. E. 1961. The Covariance Function of a Simple

Trunk Group, With Applications to Traffic Measurement. *Bell Sys. Tech. J.* **40,** 117–148.

BILLINGSLEY, P. 1968. *Convergence of Probability Measures.* John Wiley, New York.

BLOMQVIST, N. 1969. Estimation of Waiting-Time Parameters in the GI/G/1 Queueing System, Part II: Heavy-Traffic Approximations. *Skand. Akt.* 125–136.

BURMAN, D. Y. 1979. An Analytical Approach to Diffusion Approximations in Queueing. Ph.D. Thesis, Department of Applied Mathematics, New York University, New York.

BURMAN, D. Y. 1980. A Functional Central Limit Theorem for Birth and Death Processes. Talk at the ORSA/TIMS Conference in Washington, D.C., and personal communication.

BURMAN, D. Y. 1981. Insensitivity in Queueing Systems. *Adv. Appl. Prob.* **13,** 846–859.

CHUNG, K. L. 1967. *Markov Chains,* 2nd ed. Springer-Verlag, New York.

ÇINLAR, E. 1975. *Introduction to Stochastic Processes.* Prentice-Hall, Englewood Cliffs, N.J.

DENARDO, E. V. 1972. Markov Renewal Programming With Small Interest Rates. *Ann. Math. Statist.* **42,** 477–496.

DOEBLIN, W. 1938. On Two Problems of Kolmogorov Concerning Countable Chains. *Bull. Soc. Math. France* **66,** 210–220 (in French).

GLYNN, P. W. 1984. Some Asymptotic Formulas for Markov Chains With Applications to Simulation. *J. Statist. Comput. Simul.* **19,** 97–112.

GLYNN, P. W. 1989a. A Lyapunov Bound for Solutions of Poisson's Equation. Department of Operations Research, Stanford University, Stanford, Calif.

GLYNN, P. W. 1989b. Poisson's Equation for the Recurrent M/G/1 Queue. Department of Operations Research, Stanford University, Stanford, Calif.

GLYNN, P. W. 1989c. The Covariance Function of a Regenerative Process. Department of Operations Research, Stanford University, Stanford, Calif.

GLYNN, P. W., AND W. WHITT. 1991. A New View of the Heavy-Traffic Limit Theorem for Infinite-Server Queues. *Adv. Appl. Prob.* **23,** 188–209.

GRASSMANN, W. K. 1982. Initial Bias and Estimation Error in Discrete Event Simulation. In *Proceedings 1982 Winter Simulation Conference,* 377–382.

GRASSMANN, W. K. 1987a. The Asymptotic Variance of a Time Average in a Birth-Death Process. *Ann. Opns. Res.* **8,** 165–174.

GRASSMANN, W. K. 1987b. Means and Variances of Time Averages in Markovian Environments. *Eur. J. Opnl. Res.* **31,** 132–139.

GREENBERG, A. G., AND R. J. VANDERBEI. 1990. Quicker Convergence for Iterative Numerical Solutions to Stochastic Problems: Probabilistic Interpretation, Ordering Heuristics, and Parallel Processing. *Prob. Engr. Infor. Sci.* **4,** 493–521.

GROSS, D., AND D. R. MILLER. 1984. The Randomization Technique as a Modeling Tool and Solution Procedure for Transient Markov Processes. *Opns. Res.* **30,** 343–361.

HALFIN, S. 1978. A Computer Routine to Compute the Spectral Decomposition of a Birth and Death Process With Finite State Space. Unpublished Report, AT&T Bell Laboratories, Holmdel, N.J.

HALFIN, S. 1984. The Backlog of Data in Buffers With Variable Input and Output Rates. In *Performance of Computer-Communications Systems,* H. Rudin and W. Bux (eds.). North-Holland, Amsterdam, 307–319.

HAZEN, G. B., AND A. A. B. PRITSKER. 1980. Formulas for the Variance of the Sample Mean in Finite State Markov Processes. *J. Statist. Comput. Simul.* **12,** 25–40.

HEYMAN, D. P., AND M. J. SOBEL. 1982. *Stochastic Models in Operations Research.* McGraw-Hill, New York.

HORDIJK, A., D. L. IGLEHART AND R. SCHASSBERGER. 1976. Discrete Time Methods for Simulating Continuous Time Markov Chains. *Adv. Appl. Prob.* **8,** 772–788.

HOWARD, R. A. 1960. *Dynamic Programming and Markov Processes.* MIT Press, Cambridge, Mass.

KARLIN, S., AND H. M. TAYLOR. 1981. *A Second Course in Stochastic Processes.* Academic Press, New York.

KEILSON, J. 1979. *Markov Chain Models—Rarity and Exponentiality.* Springer-Verlag, New York.

KEILSON, J., AND S. S. RAO. 1970. A Process With Chain Dependent Growth Rate. *J. Appl. Prob.* **7,** 699–711.

KEMENY, J. G., AND J. L. SNELL. 1960. *Finite Markov Chains.* Van-Nostrand, Princeton, N.J.

KEMENY, J. G., AND J. L. SNELL. 1961. Finite Continuous Time Markov Chains. *Theor. Prob. Appl.* **6,** 101–105.

KEMENY, J. G., J. L. SNELL AND A. W. KNAPP. 1966. *Denumerable Markov Chains.* Van Nostrand, Princeton, N.J.

LEDERMANN, W., AND G. E. H. REUTER. 1954. Spectral Theory for the Differential Equations of Simple Birth-and-Death Processes. *Phil. Trans. Roy. Soc. London Ser. A* **246,** 321–369.

MANDL, P. 1968. *Analytical Treatment of One-Dimensional Markov Processes.* Springer-Verlag, New York.

MILLER, B. L., AND A. F. VEINOTT, JR. 1969. Discrete Dynamic Programming With a Small Interest Rate. *Ann. Math. Statist.* **40,** 366–370.

MITRA, D., AND P. TSOUCAS. 1988. Relaxations for the Numerical Solutions of Some Stochastic Problems. *Stochastic Models* **4,** 387–419.

MOELLER, T., AND H. KOBAYASHI. 1974. Use of Diffusion Approximation to Estimate Run Length in Simulation Experiments. COMPSTAT 1974, In *Proceedings of Computational Statistics,* G. Bruckmann,

F. Fershel and L. Schmetterer (eds.). Physica-Verlag, Vienna, 363–372.

MORSE, P. M. 1955. Stochastic Properties of Waiting Lines. *Opns. Res.* **3,** 255–261.

PARZEN, E. 1962. *Stochastic Processes.* Holden-Day, San Francisco.

STONE, C. 1963. Limit Theorems for Random Walks, Birth and Death Processes, and Diffusion Processes. *Ill. J. Math.* **7,** 638–660.

VAN DER WAL, J., AND P. J. SCHWEITZER. 1987. Iterative Bounds on the Equilibrium Distribution of a Finite Markov Chain. *Prob. Engr. Inform. Sci.* **1,** 117–131.

VEINOTT, A. F., JR. 1969. Discrete Dynamic Programming With Sensitive Discount Optimality Criteria. *Ann. Math. Statist.* **40,** 1635–1660.

WHITT, W. 1982. On the Heavy-Traffic Limit Theorem for GI/G/∞ Queues. *Adv. Appl. Prob.* **14,** 171–190.

WHITT, W. 1989. Planning Queueing Simulations. *Mgmt. Sci.* **35,** 1341–1366.

WOODSIDE, C. M., B. PAGUREK AND G. F. NEWELL. 1980. A Diffusion Approximation for Correlation in Queues. *J. Appl. Prob.* **17,** 1033–1047.