LIKELIHOOD RATIO GRADIENT ESTIMATION FOR STOCHASTIC SYSTEMS

SYSTEMS
FOR STOCHASTIC
GRADIENT ESTIMATION
LIKELIHOOD RATIO

PETER W. GLYNN

Consider a computer system having a CPU that feeds jobs to two input/output (I/O) devices having different speeds. Let θ be the fraction of jobs routed to the first I/O device, so that $1 - \theta$ is the fraction routed to the second. Suppose that $\alpha = \alpha(\theta)$ is the steady-sate amount of time that a job spends in the system. Given that θ is a decision variable, a designer might wish to minimize $\alpha(\theta)$ over θ. Since $\alpha(\cdot)$ is typically difficult to evaluate analytically, Monte Carlo optimization is an attractive methodology. By analogy with deterministic mathematical programming, efficient Monte Carlo gradient estimation is an important ingredient of simulation-based optimization algorithms. As a consequence, gradient estimation has recently attracted considerable attention in the simulation community. It is our goal, in this article, to describe one efficient method for estimating gradients in the Monte Carlo setting, namely the *likelihood ratio method* (also known as the *efficient score method*). This technique has been previously described (in less general settings than those developed in this article) in [6, 16, 18, 21]. An alternative gradient estimation procedure is *infinitesimal perturbation analysis;* see [11, 12] for an introduction. While it is typically more difficult to apply to a given application than the likelihood ratio technique of interest here, it often turns out to be statistically more accurate.

In this article, we first describe two important problems which motivate our study of efficient gradient estimation algorithms. Next, we will present the likelihood ratio gradient estimator in a general setting in which the essential idea is most transparent. The section that follows then specializes the estimator to discrete-time stochastic processes. We derive likelihood-ratio-gradient estimators for both time-homogeneous and non-time homogeneous discrete-time Markov chains. Later, we discuss likelihood ratio gradient estimation in continuous time. As examples of our analysis, we present the gradient estimators for time-homogeneous continuous-time Markov chains; non-time homogeneous continuous-time Markov chains; semi-Markov processes; and generalized semi-Markov processes. (The analysis throughout these sections assumes the performance measure that defines $\alpha(\theta)$ corresponds to a terminating simulation.) Finally, we conclude the article with a brief discussion of the basic issues that arise in extending the likelihood ratio gradient estimator to steady-state performance measures.

Efficient Gradient Estimation: Motivating Applications

As we have indicated, one motivation for studying Monte Carlo gradient estimation is to be able to optimize complex stochastic systems. More precisely, consider a stochastic system depending on d decision variables $\theta_1, \theta_2, \ldots, \theta_d$. Let $\alpha(\theta)$ ($\theta = (\theta_1, \ldots, \theta_d)$) be the expected "cost" of running the system at parameter choice θ.

A powerful method for computing the value θ^* which minimizes $\alpha(\cdot)$ is the Robbins-Monro algorithm. This technique recognizes that, under suitable regularity on $\alpha(\cdot)$, θ^* must be a θ-root of the equation

$$\nabla\alpha(\theta) = 0, \tag{2.1}$$

where $\nabla\alpha(\theta)$ is the gradient of $\alpha(\cdot)$ evaluated at θ. The idea then is to construct a stochastic recursion which has the root θ^* as its limit point.

This approach is most clearly illustrated when $d = 1$. In this case, such a recursion is given by

$$\theta_{n+1} = \theta_n - \frac{a}{n}V_{n+1} \tag{2.2}$$

($a > 0$) where the V_n's mimic $\alpha'(\cdot)$ in expectation. More precisely, one is required to compute V_n's with the property that

$$E\{V_{n+1}|V_0, \theta_0, \ldots, V_n, \theta_n\} = \alpha'(\theta_n) \quad \text{a.s.} \tag{2.3}$$

(Throughout this article, a.s. is our shorthand for "almost surely," otherwise known as "with probability one"). Under appropriate additional hypotheses, it then follows that there exists a finite constant such that σ

$$\theta_n \to \theta^* \ a.s.$$
$$n^{1/2}(\theta_n - \theta^*) \Rightarrow \sigma N(0,1) \tag{2.4}$$

as $n \to \infty$, *where* $N(0,1)$ is a standard normal r.v. and $\Rightarrow$ denotes "weak convergence" (also known as "convergence in distribution"). The key result in (2.4) is the central limit theorem which asserts that θ_n converges to θ^* at rate $n^{-1/2}$, in the number n of V_i's generated. Since the convergence rate $n^{-1/2}$ is typically the best that one can expect of a Monte Carlo algorithm (because of central limit effects), this suggests that recursive algorithms of the form (2.2) should lead to reasonably efficient procedures for calculating θ^*. Of course, the critical component of such an algorithm is the sequence of gradient estimates (derivative estimates when $d = 1$) $\{V_n : n \geq 0\}$ appearing in (2.3). Thus, efficient stochastic optimization is one setting which requires gradient estimation.

A second problem context which leads naturally to gradient estimation is statistical estimation for complex stochastic systems. As an example, consider a single-server infinite capacity queue in which the inter-arrival distribution F_a and service distribution F_s are unknown. Suppose that one is given data $X_1, X_2, \ldots, X_n$ for the inter-arrival times and observations $Y_1, \ldots, Y_m$ for the service times, with the goal of estimating the steady-state queue-length α. The parameter α may then be regarded as a function of the inter-arrival and service time distributions, i.e., $\alpha = \alpha(F_a,F_s)$. If F_a^* and F_s^* are respectively the "true" inter-arrival and service time distributions, our goal here is to estimate $\alpha^* = \alpha(F_a^*,F_s^*)$ from the data.

Assume that F_a^*, F_s^* are elements of one-parameter families of distributions $\{F_a(\theta_1)\},\{F_s(\theta_2)\}$, respectively, such that $F_a^* = F_a(\theta_1^*), F_s^* = F_s(\theta_2)$. We can then reduce the problem of estimating α^* to that of determining $\tilde{\alpha}(\theta_1^*,\theta_2^*)$, where $\tilde{\alpha}(\theta_1,\theta_2) = \alpha(F_a(\theta_1), F_s(\theta_2))$. For example, if $F_a(\theta_1)$ and $F_s(\theta_2)$ are both exponential and the performance measure is steady-state mean queue-length, the resulting system is an $M/M/1$ queue with ($\tilde{\alpha}$ can be calculated analytically here):

$$\tilde{\alpha}(\theta_1,\theta_2) = \begin{cases} (\theta_1/\theta_2)(1-(\theta_1/\theta_2))^{-1}, & \theta_1 < \theta_2 \\ \infty, & \theta_1 \geq \theta_2. \end{cases}$$

On the other hand, if $F_a(\cdot)$ and $F_s(\cdot)$ are Weibull with scale parameters θ_1 and θ_2 respectively, and non-unit-shape parameter, $\tilde{\alpha}$ is not available in closed form. Monte Carlo evaluation may then be necessary.

The natural estimate for α^* is $\hat{\alpha} = \hat{\tilde{\alpha}}(\hat{\theta}_1,\hat{\theta}_2)$, where $\hat{\theta}_1$ is an estimate for θ_1^* calculated from $X_1, \ldots, X_n$ and $\hat{\theta}_2$ is an estimate for θ_2^* derived from $Y_1, \ldots, Y_m$; $\hat{\tilde{\alpha}}(\cdot)$ is a Monte Carlo estimate for $\tilde{\alpha}(\cdot)$. To calculate the error in $\hat{\alpha}$ as an estimate for α^*, note that

$$\hat{\alpha} - \alpha^* = [\hat{\tilde{\alpha}}(\hat{\theta}_1,\hat{\theta}_2) - \tilde{\alpha}(\hat{\theta}_1,\hat{\theta}_2)] + [\tilde{\alpha}(\hat{\theta}_1,\hat{\theta}_2) - \tilde{\alpha}(\theta_1^*,\theta_2^*)]. \quad (2.5)$$

The first term on the right-hand side of (2.5) is error incurred from the Monte Carlo estimation of $\tilde{\alpha}(\hat{\theta}_1,\hat{\theta}_2)$; the second term, which is (conditionally) independent of the first, reflects the intrinsic error in α^* due to uncertainty in the data sets. The error in the first term can be estimated from conventional output analysis procedures. For the second, note that if $\alpha(\cdot)$ is differentiable, then

$$\tilde{\alpha}(\hat{\theta}_1,\hat{\theta}_2) - \tilde{\alpha}(\theta_1^*,\theta_2^*) \approx \nabla\tilde{\alpha}(\theta^*)(\hat{\theta} - \theta^*).$$

Typically, the vector $\hat{\theta} - \theta^*$ will be a mean zero multivariate normal, with a covariance matrix that can be easily estimated from the data sets. (This occurs, for example, if the $\hat{\theta}_i$'s are maximum likelihood estimators for the θ_i^*'s; see, for example, [13].) To calculate the distribution of the second term, it therefore remains to compute $\nabla\tilde{\alpha}(\theta^*)$ or, more precisely, its estimator $\nabla\tilde{\alpha}(\hat{\theta})$. For analytically intractable models (such as the single-server infinite capacity queue with uniform interarrival and service time distributions), this entails calculating a gradient via Monte Carlo simulation.

The situation we have just described in the single-server queueing context is typical of many statistical problems that arise in the analysis of complex stochastic systems. To fully resolve the statistical error generally requires Monte Carlo estimation of an appropriate gradient. Of course, one approach to estimating a gradient is to use a finite-difference approximation. However, when this is implemented in a Monte Carlo setting, it tends to be quite inefficient; see [8] for details. Likelihood ratio gradient estimation offers a (much) more efficient alternative.

From a mathematical viewpoint, there is no loss of generality in specializing our gradient estimation discussion to the one-dimensional setting in which $d = 1$. We shall therefore make this simplification, in order to clarify the notation. The derivative formulas that appear in the remainder of this article can be easily translated into statements about the partial derivatives that comprise the components of the gradient vector.

Likelihood Ratio Derivative Estimation

Here, we provide a brief introduction to the basic ideas underlying likelihood ratio derivative estimation. To set the stage, consider a family of stochastic systems that is indexed by a scalar decision parameter θ. For example, in a queuing context, θ might correspond to the service rate at a particular station. Given the sample space Ω, let $X(\theta,\omega)$ be the sample performance measure observed at sample outcome ω and decision parameter θ; we permit $X(\theta,\omega)$ to depend explicitly on θ in order to encompass situations in which the "cost" of running the stochastic system (as measured through $X(\theta)$) depends on the parameter θ. (However, in many estimation settings, $X(\theta)$ is independent of θ and therefore depends only on ω.) For example, in the "load-balancing" problem mentioned earlier (involving a single CPU and two I/O units), the performance measure $X(\theta)$ described is precisely the long-run sample average of the job "waiting times" experienced by the system, so that $X(\theta,\omega)$ is a function solely of the sample trajectory, ω, i.e., $X(\theta,\omega) = X(\omega)$. On the other hand, suppose that if a large fraction of jobs is routed through one I/O unit, there is a high propensity for the unit to fail, thereby increasing maintenance costs. To force the system to spread the load around, one might consider minimizing the performance measure $X(\theta,\omega) = X(\omega) + [\theta(1-\theta)]^{-1}$, where $X(\omega)$ is the performance measure described above. In this case, $X(\theta,\omega)$ depends explicitly on θ.

In addition, the probability distribution P_θ on Ω typically depends on θ; P_θ then reflects the manner in which the random environment is affected by the decision parameter. The performance measure $\alpha(\theta)$ associated with parameter value θ is then defined as the expectation

$$\alpha(\theta) = \int_\Omega X(\theta,\omega)P_\theta(d\omega).$$

Our goal is to describe an estimation methodology for calculating $\alpha'(\theta_0)$.

The likelihood ratio method for derivative estimation is based on the following idea: Suppose there exists a measure μ (not necessarily a probability measure) such that $P_\theta(d\omega) = f(\theta,\omega)\mu(d\omega)$ i.e., $f(\theta,\cdot)$ is the density of P_θ with respect to μ. For example, suppose that $\Omega = \mathbb{R}$ and that $\mu(d\omega) = d\omega$. Then, the assumption that $P_\theta(d\omega) = f(\theta,\omega)\mu(d\omega)$ merely asserts that $f(\theta,\omega)$ is the (Lebesgue) density of the distribution P_θ. On the other hand, if μ is the distribution that assigns probability mass 2^{-k} to the point x_k, then requiring that $P_\theta(d\omega)$ be representable as $f(\theta,\omega)\mu(d\omega)$ asserts that P_θ assigns all its probability to some subset of the x_k's, and that the probability assigned to x_k is $f(\theta,x_k)2^k$.

Under our density assumption on P_θ,

$$\alpha(\theta) = \int_\Omega X(\theta,\omega)f(\theta,\omega)\mu(d\omega).$$

(We note that if μ is a continuous distribution, then the above expression is an integral, whereas it reduces to a summation when μ is a discrete distribution.) Assuming the derivative and integral can be interchanged, we obtain

$$\alpha'(\theta_0) = \int_\Omega X'(\theta_0,\omega)f(\theta_0,\omega)\mu(d\omega) + \int_\Omega X(\theta_0,\omega)f'(\theta_0,\omega)\mu(d\omega). \tag{3.1}$$

(For a generic r.v. $h(\theta,\omega)$ depending on both θ and ω, the r.v. $h'(\theta_0,\omega)$ denotes the derivative of $h(\theta,\omega)$ with respect to θ, evaluated at $\theta = \theta_0$.) We note that the first term on the right-hand side of (3.1) is just $E_{\theta_0}X'(\theta_0)$ (where $E_\theta(\cdot)$ denotes the expectation operator associated with P_θ). Since this term can be represented as the expectation of a r.v., standard Monte Carlo methods may be applied to estimate it. Specifically, suppose that one simulates i.i.d. replicates of $X'(\theta_0)$ under distribution P_{θ_0}; the sample mean of these observations then converges (at rate $n^{-1/2}$ in the number n of observations) to the first term.

To handle the second term using Monte Carlo methods, we need to represent it as the expectation of a r.v. To accomplish this, suppose that $g(\omega)$ is a non-negative function that

$$\int_\Omega g(\omega)\mu(d\omega) = 1. \tag{3.2}$$

Then, the measure $P(d\omega) = g(\omega)\mu(d\omega)$ is a probability distribution on Ω. If g has the additional property that

$$|X(\theta_0,\omega)f'(\theta_0,\omega)| > 0 \text{ implies that } g(\omega) > 0, \tag{3.3}$$

then we can represent the second term as

$$\int_\Omega X(\theta_0,\omega)\frac{f'(\theta_0,\omega)}{g(\omega)}g(\omega)\mu(d\omega) = EX(\theta_0)H(\theta_0) \tag{3.4}$$

where $H(\theta_0,\omega) = f'(\theta_0,\omega)/g(\omega)$ and $E(\cdot)$ denotes expectation relative to the probability P. (Note that (3.3) is required to avoid dividing by zero in (3.4).) Given the representation (3.4) of the second term as an expectation, we can now easily apply Monte Carlo methods to estimate it (in the same way as for the first term).

We now turn to the question of selecting the sampling density g. The theory of importance sampling asserts that the choice of g which minimizes the variance of the observations of $X(\theta_0)H(\theta_0)$ is

$$g^*(\omega) = \frac{|X(\theta_0,\omega)f'(\theta_0,\omega)|}{\int_\Omega|X(\theta_0,\omega)f'(\theta_0,\omega)|\mu(d\omega)}, \tag{3.5}$$

see [9,10], for further details. (In fact, if $X(\theta_0,\omega)f'(\theta_0,\omega) \geq 0$, using g^* results in an estimator having zero variance.) Unfortunately, the optimal sampling density g^* basically requires knowledge of the integral (appearing in the second term in (3.1)) that we are trying to estimate. Therefore, the choice of g^* as defined by (3.5) is typically impractical to implement.

We now describe a popular alternative to g^*. Suppose that the densities $f(\theta,\omega)$ are such that for θ in an open neighborhood of θ_0,

$$\Lambda(\theta) = \{\omega : f(\theta,\omega) > 0\} \text{ is independent of } \theta. \tag{3.6}$$

To gain an understanding of the condition, suppose that $\Omega = \mathbb{R}$ and that the distribution μ is supported on $x_1, x_2, \ldots$. Condition (3.6) states that the set of values $x_1, x_2, \ldots$ which have positive probability under P_θ must be independent of θ in some neighborhood of θ_0. Note that the support of the distribution P_θ cannot depend on θ, so that, in particular, a situation in which P_θ assigns positive probability to the point θ is disallowed by (3.6).

Then, $f(\theta_0,\omega) = 0$ implies that $f(\theta,\omega)$ vanishes in a neighborhood of θ, from which it follows that $f'(\theta_0,\omega) = 0$, so that $f'(\theta_0,\omega)X(\theta_0,\omega) = 0$. Thus, $g(\omega) = f(\theta_0,\omega)\mu(d\omega)$ satisfies both (3.2) and (3.3). In this case,

$$H(\theta_0,\omega) = \frac{f'(\theta_0,\omega)}{f(\theta_0,\omega)} \left(= \frac{d}{d\theta}\log f(\theta_0,\omega)\right); \tag{3.7}$$

the right-hand side of (3.7) is known as the **likelihood ratio derivative** (because $H(\theta_0,\omega) = \frac{d}{d\theta}\frac{f(\theta,\omega)}{f(\theta_0,\omega)}$ is the derivative of the quantity known in the statistics literature as the likelihood ratio of P_θ with respect to P_{θ_0}).

This choice of g has an important advantage. Note that if we sample outcomes ω according to $f(\theta_0,\omega)\mu(d\omega)$, we can use the r.v.'s $X(\theta_0)$, $X'(\theta_0)$, and $X(\theta_0)H(\theta_0)$ to estimate $\alpha(\theta_0)$ and both the terms appearing on the right-hand side of

(3.1) simultaneously. Thus, with this choice of g, we may estimate $\alpha(\theta_0)$ and $\alpha'(\theta_0)$ using the **original** sampling distribution associated with parameter θ_0. At the same time, it should be noted that there are important problem classes (e.g., rare event simulations) in which much better choices of g can be made (better in the sense of smaller variance). For example, in reliability systems that are modelled as continuous-time Markov chains, one needs to "failure bias" (i.e., choose a g which forces the system to fail more frequently) the estimator $X(\theta_0)H(\theta_0)$ in order to obtain reasonable statistical efficiency (see [14]).

We will conclude by recalling that to derive (3.1), an interchange of the differentiation and expectation operators was required. In virtually all practical examples, the interchange is valid under mild additional regularity assumptions on the problem (see, [4,], p. 485). Consequently, we shall disregard this interchange issue throughout the remainder of this article.

Likelihood Ratio Derivative Estimation in Discrete Time

Here, we specialize the previous discussion to the case where $X(\theta,\omega)$ is a sample performance measure associated with a discrete-time sequence $Y = (Y_n : n \geq 0)$ taking values in a discrete state space S. Specifically, we suppose that $\Omega = S \times S \times \ldots$ and that Y_n is the coordinate r.v. $Y_n(\omega) = \omega_n$ for $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$. We assume that $X(\theta)$ takes the form

$$X(\theta) = h(\theta, Y_0, Y_1, \ldots),$$

for some real-valued function h. Since S is discrete, there exist joint probability mass functions $p_0, p_1, \ldots$ such that

$$P_\theta\{Y_0 = y_0, \ldots, Y_n = y_n\} = p_n(\theta, \vec{y}_n) \tag{4.1}$$

where $\vec{y}_n = (y_0, \ldots, y_n)$. Letting

$$p_n(\theta, \vec{y}_{n-1}; y_n) = P_\theta\{Y_n = y_n | Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}\},$$

we can write (4.1) as the product

$$P_\theta\{Y_0 = y_0, \ldots, Y_n = y_n\} = p_0(\theta, y_0) \prod_{k=0}^{n-1} p_k(\theta, \vec{y}_k; y_{k+1}) \tag{4.2}$$

Suppose now that $X(\theta)$ is a function of Y up to some finite (deterministic) time horizon m, so that $X(\theta) = h(\theta, \vec{Y}_m)$ where $\vec{Y}_m = (Y_0, \ldots, Y_m)$. To apply the idea of Section 3, we need to obtain a representation $P_\theta(d\omega) = f(\theta,\omega)\mu(d\omega)$ for some measure μ. But observe that for $w \in \Omega_m$,

$$P_\theta(d\omega) = p_0(\theta; \omega_0) \prod_{k=0}^{m-1} p_k(\theta, \vec{\omega}_k; w_{k+1}) \cdot \mu_m(d\omega)$$

where $\vec{\omega}_k = (\omega_0, \ldots, \omega_k)$ and μ_m is counting measure on $\Omega_m = S \times S \times \cdots \times S (m + 1 \text{ times})$, i.e., μ_m assigns unit mass to each point in Ω_m. Hence, we may take

$$f(\theta,\omega) = p_0(\theta,\omega_0) \prod_{k=0}^{m-1} p_k(\theta, \vec{\omega}_k; \omega_{k+1}),$$

so that

$$f'(\theta_0,\omega) = p_0'(\theta_0,\omega_0) \prod_{k=0}^{m-1} p_k(\theta_0, \vec{\omega}_k; \omega_{k+1}) + p_0(\theta_0,\omega_0) \sum_{k=0}^{m-1} p_k'(\theta_0, \vec{\omega}_k; \omega_{k+1}) \cdot \prod_{j \neq k} p_j(\theta_0, \vec{\omega}_j; \omega_{j+1}). \tag{4.3}$$

We can simplify this formula somewhat. We claim that if $p_k'(\theta_0, \vec{\omega}_k, \omega_{k+1}) \neq 0$, it must follow that $p_k(\theta_0, \vec{\omega}_k, \omega_{k+1}) > 0$. For suppose that $p_k(\theta_0, \vec{\omega}_k; \omega_{k+1}) = 0$. Then it follows that

$$p_k(\theta_0 + h, \vec{\omega}_k; \omega_{k+1}) = p_k'(\theta_0, \vec{\omega}_k; \omega_{k+1})h + o(h)$$

as $h \downarrow 0$, from which it is evident that $p_k(\theta_0 + h, \vec{\omega}_k; \omega_{k+1}) < 0$ for some h. But $p_k(\theta, \vec{\omega}_k; \omega_{k+1})$ is a mass function and hence must be non-negative. This contradiction guarantees that $p_k(\theta_0, \vec{\omega}_k; \omega_{k+1}) > 0$. A similar argument shows that $p_0(\theta_0, \omega_0) > 0$ whenever $p_0'(\theta_0, \omega_0) \neq 0$. Hence, we may write (4.3) as

$$f'(\theta_0,\omega) = f(\theta_0,\omega)\left[\frac{p_0'(\theta_0,\omega_0)}{p_0(\theta_0,\omega_0)} + \sum_{k=0}^{m-1} \frac{p_k'(\theta_0, \vec{\omega}_k; \omega_{k+1})}{p_k(\theta_0, \vec{\omega}_k; \omega_{k+1})}\right]$$

Suppose we choose a g such that $\int_{\Omega_m} g(\omega)\mu_m(d\omega) = 1$ and $f(\theta_0,\omega) > 0$ implies that $g(\omega) > 0$; then (3.3) is automatically in force. (In particular, setting $g(\omega) = f(\theta_0,\omega)$ works.) Hence, we find that

$$\alpha'(\theta_0) = E_{\theta_0} X'(\theta_0) + E_g X(\theta_0) H(\theta_0) \tag{4.4}$$

where $E_g(\cdot)$ denotes the expectation operator associated with the probability $P_g(d\omega) = g(\omega)\mu_n(d\omega)$, $E_\theta(\cdot)$ denotes expectation relative to P_θ, and

$$H(\theta_0) = \frac{f(\theta_0, \vec{Y}_m)}{g(\vec{Y}_m)}\left[\frac{p_0'(\theta_0, Y_0)}{p_0(\theta_0, Y_0)} + \sum_{k=0}^{m-1} \frac{p_k'(\theta_0, \vec{Y}_k; Y_{k+1})}{p_k(\theta_0, \vec{Y}_k; Y_{k+1})}\right].$$

The same argument can be extended to a certain class of random time horizons. In particular, suppose that T is a *stopping time* with respect to Y i.e., for each $m \geq 0$, $I(T = m) = k_m(\vec{Y}_m)$ for some function k_m. For example, T is a stopping time if it can be represented as the first time that Y hits some specified subset; for further discussion, see [2]. We assume the performance measure $X(\theta)$ is a function of the path of Y up to the random time horizon T i.e., there exists a family of functions $h_0, h_1, \ldots$ such that

$$X(\theta) = \sum_{m=0}^{\infty} h_m(\theta, \vec{Y}_m) I(T = m) = h_T(\theta, \vec{Y}_T) I(T < \infty). \quad (4.5)$$

As in the derivation of (4.4), we need to represent P_θ as $P_\theta(d\omega) = f(\theta,\omega)\mu(d\omega)$. Let $\Omega_T = \cup_{m=0}^{\infty} \{\vec{\omega}_m \in \Omega_m : k_m(\vec{\omega}_m) = 1\}$ (basically, Ω_T is the restriction of the sample space for Y in which $T < \infty$) and note that for $\omega = (\omega_0, \omega_1, \ldots, \omega_T) \in \Omega_T$,

$$P_\theta(\alpha\omega) = p_0(\theta,\omega_0) \prod_{k=0}^{T-1} p_k(\theta, \vec{\omega}_k; \omega_{k+1}) \cdot \mu_T(d\omega) \quad (4.6)$$

where μ_T is counting measure on Ω_T. Suppose that g is chosen as a non-negative function on Ω_T having the property that $\int_\Omega g(\omega)\mu_T(d\omega) = 1$ and $p_0(\theta_0,\omega_0)\Pi_{k=0}^{T-1} p_k(\theta_0, \vec{\omega}_k; \omega_{k+1}) > 0$ implies that $g(\omega) > 0$ for $\omega \in \Omega_T$. By combining (4.5) and (4.6) and proceeding as in the derivation of (4.4), we obtain the following stopping time generalization of (4.4):

$$\alpha'(\theta_0) = E_{\theta_0} X'(\theta_0) + E_g X(\theta_0) H(\theta_0) \quad (4.7)$$

where

$$H(\theta_0) = \frac{p_0(\theta_0,Y_0)\Pi_{k=0}^{T-1} p_k(\theta_0,\vec{Y}_k;Y_{k+1})}{g(\vec{Y}_T)} \cdot \left[\frac{p_0'(\theta_0,Y_0)}{p_0(\theta_0,Y_0)} + \sum_{k=0}^{T-1} \frac{p_k'(\theta_0,\vec{Y}_k;Y_{k+1}}{p_k(\theta_0,\vec{Y}_k;Y_{k+1})}\right]$$

As in the case of (4.4), one possible choice of g is $f(\theta_0)$, in which event (4.7) simplifies to:

$$\alpha'(\theta_0) = E_{\theta_0}[X'(\theta_0) + X(\theta_0)H(\theta_0)] \quad (4.8)$$

where

$$H(\theta_0) = \frac{p_0'(\theta_0,Y_0)}{p_0(\theta_0,Y_0)} + \sum_{k=0}^{T-1} \frac{p_k'(\theta_0;\vec{Y}_k;Y_{k+1})}{p_k(\theta_0,\vec{Y}_k;Y_{k+1})}. \quad (4.9)$$

A few examples illustrate (4.7) and (4.8).

EXAMPLE. Suppose that under distribution P_θ, Y is a time-homogeneous Markov chain with initial distribution $\mu(\theta)$ and transition matrix $P(\theta)$. Assume that $X(\theta) = h_T(\vec{Y}_T)I(T < \infty)$ (with T a stopping time), so that $\alpha(\theta) = E_\theta\{h_T(\vec{Y}_T); T < \infty\}$. Then, (4.8) yields

$$\alpha'(\theta_0) = E_{\theta_0}\{h_T(\vec{Y}_T)H(\theta_0); T < \infty\}, \quad (4.10)$$

where $H(\theta_0) = \mu'(\theta_0,Y_0)/\mu(\theta_0,Y_0) + \Sigma_{k=0}^{T-1} P'(\theta_0,Y_k,Y_{k+1})/P(\theta_0,Y_k,Y_{k+1})$. In certain settings, the estimator suggested by (4.10) may have a large variance (e.g., rare event simulation). For such problems, suppose that we select g to satisfy the positivity conditions stated earlier. Then

$$\alpha'(\theta_0) = E_g\{h_T(\vec{Y}_T)H(\theta_0); T < \infty\}, \quad (4.11)$$

where

$$H(\theta_0) = \mu(\theta_0,Y_0)\Pi_{k=0}^{T-1}(P(\theta_0,Y_k,Y_{k+1})/g(\vec{Y}_T) \cdot [\mu'(\theta_0,Y_0)/\mu(\theta_0,Y_0) + \sum_{k=0}^{T-1} P'(\theta_0,Y_k,Y_{k+1})/P(\theta_0,Y_k,Y_{k+1})].$$

In a "rare event" setting, one would typically choose g so as to bias the system to force the occurence of more rare events.

EXAMPLE. In this example, we assume that under P_θ, Y is a Markov chain with non-stationary transition probabilities, so that $P_\theta\{Y_{k+1} = y_{k+1} | Y_k = y\} = P_k(\theta, y_k, y_{k+1})$. Then, if $\alpha(\theta) = E_\theta\{h_T(\vec{Y}_T); T < \infty\}$, (4.8) yields

$$\alpha'(\theta_0) = E_{\theta_0}\{h_T(\vec{Y}_T)H(\theta_0); T < \infty\}$$

where $H(\theta_0) = \mu'(\theta_0,Y_0)/\mu(\theta_0,Y_0) + \Sigma_{k=0}^{T-1} P_k'(\theta_0,Y_k,Y_{k+1})/P_k(\theta_0,Y_k,Y_{k+1})$; the obvious analog of (4.11) can also be written down.

Likelihood Ratio Derivative Estimation in Continuous Time

We will now generalize the ideas of the previous section to continuous-time discrete-event dynamical systems. We view $X(\theta,\omega)$ as a sample performance measure associated with a continuous-time process $(Y = Y(t): t \geq 0)$ taking values in a discrete state space S. The process Y is assumed to be piece-wise constant with jump times $S_1, S_2, \ldots$ ($S_n \to \infty$ as $n \to \infty$). Hence, if $S_0 = 0$ and $Y_n = Y(S_n)$, we may write

$$Y(t) = \sum_{n=0}^{\infty} Y_n I(S_n \leq t < S_{n+1}).$$

Let $\Delta_n = S_{n+1} - S_n$ and put $Z_n = (Y_n, \Delta_n)$. We suppose that $\Omega = \hat{S} \times \hat{S} \times \ldots$ where $\hat{S} = S \times [0,\infty)$ and that Z_n is the co-ordinate r.v. $Z_n(\omega) = \omega_n$ for $\omega = (w_0, w_1, \ldots) \in \Omega$.

In order to proceed in parallel with the development of Section 4, we shall require that the distributions P_θ on Ω have the property that there exist measures $\mu_0, \mu_1, \ldots$ such that

$$P_\theta\{Z_0 \in dz_0\} = p_0(\theta,z_0)\mu_0(dz_0)$$
$$P_\theta\{Z_{n+1} \in dz_{n+1} | \vec{Z}_n = \vec{z}_n\} = p_n(\theta, \vec{z}_n, z_{n+1})\mu_n(\vec{z}_n, dz_{n+1})$$

where $\vec{Z}_n = (Z_0, \ldots, Z_n)$ and $\vec{Z}_n = (z_0, \ldots, z_n) \in \hat{S} \times \ldots \times \hat{S} = \Omega_n$ ($(n + 1)$ times). Then, analogously to (4.2), we may write

$$P_\theta\{\vec{Z}_n \in d\vec{z}_n\} = p_0(\theta,z_0) \cdot \prod_{k=0}^{n-1} p_k(\theta,\vec{z}_k;z_{k+1})\mu_n(d\vec{z}_n) \quad (5.1)$$

where

$$\mu_n(d\vec{z}_n) = \mu_0(dz_0) \prod_{k=1}^{n-1} \mu_k(\vec{z}_k, dz_{k+1}).$$

Suppose now that we consider a performance measure $X(\theta)$ that is a function of the path up to horizon T; this obviously includes any performance measure that depends on Y up to time S_{T+1}. As in the previous section, we require that T be a stopping time with respect to $Z = (Z_n : n \geq 0)$ i.e., for each $m \geq 0$, $I(T = m) = k_m(\vec{Z}_m)$ for some function k_m. Then, the performance measure $X(\theta)$ may be written in the form

$$X(\theta) = \sum_{k=0}^{\infty} h_k(\theta,\vec{Z}_k) I(T = k) = h_T(\vec{Z}_T) I(T < \infty).$$

Let $\Omega_T = \cup_{m=0}^{\infty}\{\vec{z}_m \in \Omega_m : k_m(\vec{\omega}_m) = 1\}$ and note that for $\vec{z}_T = (z_0, \ldots, z_T) \in \Omega_T$, we may extend (5.1) to

$$P_\theta\{\vec{Z}_T \in d\vec{z}_T\} = p_0(\theta,z_0) \prod_{k=0}^{T-1} p_k(\theta,\vec{z}_k;z_{k+1}) \cdot \mu_T(d\vec{z}_T) \tag{5.2}$$

where

$$\mu_T(d\vec{z}_T) = \mu_0(dz_0) \prod_{k=0}^{T-1} \mu_k(\vec{z}_k, dz_{k+1}).$$

By arguing identically as in the previous section, we obtain the following continuous-time generalization of (4.7). Suppose g is chosen as a nonnegative function on Ω_T having the property that $\int_{\Omega_T} g(\vec{z}_T)\mu_T(d\vec{z}_T) = 1$ and $p_0(\theta_0,z_0)\prod_{k=0}^{T-1} p_k(\theta_0,\vec{z}_k,z_{k+1}) > 0$ implies that $g(\vec{z}_T) > 0$ for $\vec{z}_T \in \Omega_T$. Then, if $E_g(\cdot)$ is the expectation operator associated with $P(d\vec{z}_T) = g(\vec{z}_T)\mu(d\vec{z}_T)$, we obtain the derivative representation

$$\alpha'(\theta_0) = E_{\theta_0} X'(\theta_0) + E_g X(\theta_0) H(\theta_0)$$

for $\alpha(\theta) = E_\theta\{h(\theta,\vec{Z}_T); T < \infty\}$, where

$$H(\theta_0) = p_0(\theta_0,Z_0) \prod_{k=0}^{T-1} p_k(\theta_0,\vec{Z}_k;Z_{k+1}) / g(\vec{Z}_T) \cdot \left[\frac{p_0'(\theta_0,Z_0)}{p_0(\theta_0,Z_0)} + \sum_{k=0}^{T-1} \frac{p_k'(\theta_0,\vec{Z}_k;Z_{k+1})}{p_k(\theta_0,\vec{Z}_k;Z_{k+1})}\right].$$

As mentioned earlier, one possible choice for g is $g(\vec{z}_T) = p_0(\theta_0,z_0)\prod_{k=0}^{T-1} p_k(\theta_0,\vec{z}_k;z_{k+1})$, in which case P is identical to P_{θ_0}, yielding

$$\alpha'(\theta_0) = E_{\theta_0}[X'(\theta_0) + X(\theta_0)H(\theta_0)] \tag{5.3}$$

where

$$H(\theta_0) = \frac{p_0'(\theta_0,Z_0)}{p_0(\theta_0,Z_0)} + \sum_{k=0}^{T-1} \frac{p_k'(\theta_0,\vec{Z}_k;Z_{k+1})}{p_k(\theta_0,\vec{Z}_k;Z_{k+1})}.$$

These formulas are illustrated by the following examples:

EXAMPLE. Suppose that under P_θ, Y is a continuous-time Markov chain with initial distribution $\mu(\theta)$ and generator $Q(\theta)$. Assume that $X(\theta) = h(Y(s): 0 \leq s \leq t)$. Then, $X(\theta)$ can be represented as $X(\theta) = \hat{h}(Z_0, Z_1, \ldots, Z_T)$ where T is the stopping time $T = \inf\{n \geq 0 : \Sigma_{k=0}^{n}\Delta_k \geq t\}$. Set $z_n = (y_n,t_n)$(recall that $z_n \in \hat{S} = S \times [0,\infty)$). Then,

$$P_\theta\{Z_0 \in dz_0\} = p_0(\theta,z_0)\mu_0(dz_0)$$

where $p_0(\theta,y_0,t_0) = \mu(\theta,y_0)q(\theta,y_0) \cdot \exp(-q(\theta,y_0)t_0), q(\theta,y) = -Q(\theta,y,y)$, and $\mu_0(dz_0)$ is the product of counting measure and Lebesgue measure. We note that $p_0(\theta,y_0,t_0)$ is a product of two terms, the first being a contribution of $\mu(\theta,y_0)$ from the distribution of the initial state and the second being the exponential holding time density of the time spent in the initial state. Furthermore,

$$P_\theta\{Z_{n+1} \in dz_{n+1} | \vec{Z}_n = \vec{z}_n\} = p_n(\theta,\vec{z}_n;z_{n+1})\mu_n(\vec{z}_n,dz_{n+1})$$

where $p_n(\theta,\vec{z}_n;z_{n+1}) = Q(\theta,y_n,y_{n+1}) \cdot q(\theta,y_{n+1})\exp(-q(\theta,y_{n+1})t_{n+1})/q(\theta,y_n)$ and $\mu_n(\vec{z}_n,dz_{n+1})$ is again the product of counting measure and Lebesgue measure. Again, $p_n(\theta,\vec{z}_n;z_{n+1})$ is the product of a state transition term $(Q(\theta,y_n,y_{n+1})/q(\theta,y_n))$ and an exponential holding time density. Formula (5.3) now becomes

$$\alpha'(\theta_0) = E_{\theta_0}[h(Y(s): 0 \leq s \leq t)H(\theta_0)] \tag{5.4}$$

where

$$H(\theta_0) = \frac{\mu'(\theta_0,Y_0)}{\mu(\theta_0,Y_0)} + \sum_{k=0}^{T-1} \frac{Q'(\theta_0,Y_k,Y_{k+1})}{Q(\theta_0,Y_k,Y_{k+1})} + \frac{q'(\theta_0,Y_T)}{q(\theta_0,Y_T)} - \sum_{k=0}^{T} q'(\theta_0,Y_k)\Delta_k.$$

EXAMPLE. Suppose that under P_θ, Y is a semi-Markov process with initial distribution $\mu(\theta)$, jump matrix $R(\theta)$, and holding-time distributions $(F(\theta,x,dt): x \in S)$. Suppose that for each x, $F(\theta,x,dt) = f(\theta,x,t)\mu(x,dt)$ for some measure μ. Assuming that $X(\theta) = h(Y(s): 0 \leq s \leq t)$, we again put $T = \inf\{n \geq 0 : \Sigma_{k=0}^{n}\Delta_k \geq t\}$. Formula (5.3) becomes

$$\alpha'(\theta_0) = E_{\theta_0}[h(Y(s): 0 \leq s \leq t)H(\theta_0)] \tag{5.5}$$

where

$$H(\theta_0) = \frac{\mu'(\theta_0,Y_0)}{\mu(\theta_0,Y_0)} + \sum_{k=0}^{T-1} \frac{R'(\theta_0,Y_k,Y_{k+1})}{R(\theta_0,Y_k,Y_{k+1})} + \sum_{k=0}^{T} \frac{f'(\theta_0,Y_k,\Delta_k)}{f(\theta_0,Y_k,\Delta_k)}.$$

EXAMPLE. In this example, we show that (5.3) easily handles the case where the process is time-inhomogeneous. In particular, suppose that under P_θ, Y is a time-homogeneous continuous-time Markov chain with initial distribution $\mu(\theta)$ and time-dependent generator $Q(\theta,t)$. Then,

$$P_\theta\{Y_{n+1} = y, \Delta_{n+1} \in dt | \vec{Z}_n\} = \frac{Q(\theta, S_{n+1}, Y_n, y)}{q(\theta, S_{n+1} Y_n)} q(\theta, S_{n+1} + t, y) \cdot \exp\left(-\int_0^t q(\theta, S_{n+1} + u, y) du\right) dt \tag{5.6}$$

where $q(\theta,t,y) = -Q(\theta,t,y,y)$. Suppose that $X(\theta) = h(Y(s): 0 \leq s \leq t)$. If we put $T = \inf\{n \geq 0 : \Sigma_{k=0}^{n} \Delta_k \geq t\}$, then (5.3) takes the form

$$a'(\theta_0) = E_{\theta_0}[h(Y(s): 0 \leq s \leq t) H(\theta_0)] \tag{5.7}$$

where

$$H(\theta_0) = \frac{\mu'(\theta_0, Y_0)}{\mu(\theta_0, Y_0)} + \sum_{k=0}^{T-1} \frac{Q'(\theta_0, S_{k+1}, Y_k, Y_{k+1})}{Q(\theta_0, S_{k+1}, Y_k, Y_{k+1})} + \frac{q'(\theta_0, S_{T-1}, Y_T)}{q(\theta_0, S_{T+1}, Y_T)} - \sum_{k=0}^{T} \int_{S_k}^{S_{k+1}} q'(\theta_0, t, Y_k) dt.$$

EXAMPLE. We now suppose that Y is a generalized semi-Markov process (GSMP) under P_θ. A GSMP is a mathematical description of a very general class of discrete-event stochastic systems. Roughly speaking, any discrete event simulation can be viewed as a GSMP; see [5] for further details on GSMPs. In what follows, we describe a GSMP by characterizing both the stochastic behavior of the sequence of discrete states visited and that of certain clocks that govern the amount of time spent in each successive state. To be specific, let E be the event set of the GSMP. The initial state of the GSMP is chosen according to the distribution $\mu(\theta)$, whereas the initial clock readings are chosen from the distributions $F(\theta,e,dt)$, for $e \in E$. When clock e initiates a transition from state y, the next state is chosen from the mass function $p(\theta;y,e)$. Typically, when the GSMP enters a new state, certain clocks need to be stochastically reset. We assume that the distribution used to reset clock e' in state y' when a transition just occurred from state y with clock e as triggering event is given by $F(\theta,e',y',e,y,dt)$. We require that there exist measures $\mu(e,dt)$, $\mu(e',y',e,y,dt)$ such that

$$F(\theta,e',y',e,y,dt) = f(\theta,e',y',e,y,t), \mu(e',y',e,y,dt) \tag{5.8}$$
$$F(\theta,e,dt) = f(\theta,e,t)\mu(e,dt).$$

In a strict sense, the analysis of this section does not apply to GSMPs, since the appropriate state descriptor for a GSMP includes the value of all the clock readings. Such a state descriptor cannot typically be encoded as an element of $\hat{S} = S \times [0,\infty)$. However, a close examination of the analysis given earlier shows that the essential feature was that (Y_n, Δ_n) be representable as a simple function of the process z_n; z_n need have no structure beyond (5.1). In particular, z_n need not be an element of $\hat{S}$. In the GSMP setting, the natural candidate for Z is the tuple $z_n = (Y_n, C_n)$, where C_n is the vector that describes the residual amount of time left on each of the clocks that are active in state Y_n. Clearly, Δ_n is a simple function of z_n (in a GSMP with unit speeds, Δ_n is just the minimal element in C_n); furthermore, under (5.8), the distribution P_θ for $\vec{Z}_n$ can be written in the form (5.1).

Let $N(y';y,e)$ be the set of clocks active in y' that need to be stochastically re-set when a transition from y just occurred with event e as the trigger. We further define $e^*(c)$ to be the index of the triggering event associated with clock vector c; we assume e^* is uniquely defined for each c. Suppose $X(\theta) = h(Y(s): 0 \leq s \leq t)$. If we put $T = \inf\{n \geq 0 : \Sigma_{k=0}^{n} \Delta_k \geq t\}$, it is easily verified that (5.3) takes the form

$$\alpha'(\theta_0) = E_{\theta_0}[h(Y(s): 0 \leq s \leq T) H(\theta_0)]$$

where

$$H(\theta_0) = \frac{\mu'(\theta_0, Y_0)}{\mu(\theta_0, Y_0)} + \sum_{k=0}^{T-1} \frac{p'(\theta_0, Y_{k+1}; Y_k, e^*(C_k))}{p(\theta_0, Y_{k+1}; Y_k, e^*(C_k))} + \sum_{e} \frac{f'(\theta_0, e, C_{0_e})}{f(\theta_0, e, C_{0_e})} + \sum_{k=1}^{T} \sum_{e \in N(Y_k; Y_{k-1}, e^*(C_{k-1}))} \frac{f'(\theta_0, e, Y_k, e^*(C_{k-1}), Y_{k-1}, C_{ke})}{f(\theta_0, e, Y_k, e^*(C_{k-1}), Y_{k-1}, C_{ke})}.$$

The first sum over k is the contribution to the likelihood ratio gradient from the sequence of discrete states visited, whereas the second sum over k is the contribution from the successive clocks that are set (the remaining sums are contributed by randomness in the initial condition).

These examples serve to illustrate the great variety of stochastic processes to which likelihood ratio derivative estimation may be applied.

Steady-State Gradient Estimation

The discussion of the previous sections of this article basically pertains to the case in which the performance measure depends on the process up to some (finite-valued) stopping time T (see, for example, (3.6)). Hence, the material described thus far is principally motivated by terminating simulations. Here, we will briefly discuss some of the basic issues that arise in dealing with steady-state performance measures.

Most of the technical difficulties associated with the likelihood ratio approach to steady-state gradient estimation stem from an assumption made in the section on Likelihood Ratio Derivatives Estimation that asserts existence of a smooth likelihood ratio—that there exists a measure μ and a function $f(\theta,\omega)$ (smooth in θ) such that $P_\theta(d\omega) = f(\theta,\omega)\mu(d\omega)$.

Now, a steady-state performance measures X (for a discrete-time process Y) typically can be expressed in the form

$$X = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(Y_k)$$

for some real-valued f, so that X depends on the infinite history of Y. Now, if Y is ergodic under P_θ, it follows that there exists a (deterministic) constant $\alpha(\theta)$ such that

$$\frac{1}{n}\sum_{k=0}^{n-1} f(Y_k) \to \alpha(\theta) \quad P_\theta \ a.s. \tag{6.1}$$

Hence, $P_\theta(A(\theta)) = 1$, where $A(\theta) = \{\omega : X(\omega) = \alpha(\theta)\}$. Therefore, unless $\alpha(\theta) = \alpha(\theta_0)$, $P_\theta(A(\theta_0)) = 0$ for θ arbitrarily close to θ_0, whereas $P_{\theta_0}(A(\theta_0)) = 1$. It follows that $f(\theta,\cdot)$ has a support that is disjoint from that of $f(\theta_0,\cdot)$, regardless of how close θ is to θ_0. This, of course, is incompatible with $f(\theta,\omega)$ being differentiable in θ at the point θ_0. Thus, f cannot be smooth in θ.

Since smooth likelihood ratios do not exist over infinite time horizons, this suggests one ought to try to reduce the infinite horizon steady-state estimation problem to a finite horizon problem before applying likelihood ratio techniques. (Recall that finite horizon problems can be treated by earlier techniques.)

One class of stochastic systems that is perfectly suited to such a reduction is the class of regenerative stochastic processes (see [2] for a definition). Assume that Y is a nondelayed regenerative process with regeneration times $T_0 = 0 < T_1 < T_2 < \cdots$. (For example, in a Markov process setting, the regeneration times would typically correspond to successive hitting times of some fixed state.) Then, under suitable moment conditions, it can be shown that

$$\frac{1}{n}\sum_{k=0}^{n-1} f(Y_k) \to u(\theta)/\ell(\theta) \quad P_\theta \ a.s. \tag{6.2}$$

as $n \to \infty$, where

$$u(\theta) = E_\theta \sum_{k=0}^{T_1-1} f(Y_k)$$

$$l(\theta) = E_\theta T_1.$$

Hence, the steady-state of a regenerative process can be expressed as a ratio of two expectations, in which both expected values correspond to performance measures that depend on Y only up to a stopping time, namely T_1. Intuitively, in a regenerative process, the infinite horizon steady-state behavior is faithfully represented by the behavior of the process over the "regenerative cycle" $[0, T_1)$ (which, of course, constitutes a finite horizon).

The relation (6.2) asserts that the steady-state mean $\alpha(\theta)$ can be expressed as $\alpha(\theta) = u(\theta)/\ell(\theta)$. Thus, via the quotient rule of differential calculus (i.e., $\alpha'(\theta) = l(\theta)^{-2}[u'(\theta)\ell(\theta) - l'(\theta)u(\theta)]$), it follows that derivative estimation for $\alpha(\cdot)$ can be reduced to that for $u(\cdot)$ and $l(\cdot)$. But the derivatives of the performance measures $u(\cdot)$ and $l(\cdot)$ can be estimated via the likelihood ratio methods described in Sections 2 through 5. For details on this approach, see [6, 16].

In the non-regenerative setting, likelihood ratio gradient estimation of steady-state performance measures is more problematic. One approach that holds some promise is to observe that if the process Y is ergodic (i.e., (6.1) holds), then it typically follows that

$$E_\theta\left[\frac{1}{n}\sum_{k=0}^{n-1} f(Y_k)\right] \to \alpha(\theta) \tag{6.3}$$

as $n \to \infty$. Since the left-hand side depends on Y only up to the (finite) time horizon n, likelihood ratio gradient estimation techniques can be applied. In [7], this approach is discussed in more detail. Related techniques are described in [1, 3]. One difficulty is that since (6.3) is only an approximation for any finite n, the corresponding gradient estimator will be biased. This induces statistical inefficiencies in the estimator (i.e., a slower rate of convergence than that obtained in the regenerative setting). The development of more efficient likelihood ratio gradient estimators for steady-state nonregenerative performance measures continues to be an active research area.

Acknowledgment.
Much of the material in this article is adapted from previous articles by the author. These appeared in both the 1987 and 1989 Proceedings of the Winter Simulation Conference. The author also wishes to thank the referee of this article for extensive and useful comments.

References

1. Fuerverger, A., McLeish, D. L., Kreimer, J., and Rubinstein, R. Y. Sensitivity analysis and the "what if" problem in simulation analysis. *Math. Comput. Modelling 12*, (1989), 193–219.
2. Çinlar, E. *Introduction to Stochastic Processes.* Prentice-Hall, Englewood Cliffs, New Jersey, 1975.
3. Fuerverger, A., McLeish, D. L., and Rubinstein, R. Y. A cross-spectral method for sensitivity analysis of computer simulation models. *Comp. Rend. Math. Rep. Acad. Sci. Roy. Can. 8*, (1986), 335–339.
4. Fulks, W. *Advanced Calculus.* John Wiley, New York, 1969.
5. Glynn, P. W. On the role of generalized semi-Markov processes in simulation output analysis. In *Proceedings of the 1983 Winter Simulation Conference*, (1983), pp. 39–42.
6. Glynn, P. W. Stochastic approximation for Monte Carlo optimization. In *Proceedings of the 1986 Winter Simulation Conference*, (1986), pp. 356–365.
7. Glynn, P. W. Likelihood ratio gradient estimation: an overview. In *Proceedings of the 1987 Winter Simulation Conference*, (1987), pp. 366–375.
8. Glynn, P. W. Optimization of stochastic systems via simulation. In *Proceedings of the 1989 Winter Simulation Conference*, (1989), pp. 90–105.

9. Glynn, P. W. and Iglehart, D. L. Importance sampling for stochastic simulations. *Manage. Sci.* (1989).
10. Hammersley, J. M. and Handscomb, D. C. *Monte Carlo Methods.* Chapman and Hall, New York, (1965).
11. Ho, Y. C. Performance evaluation and perturbation analysis of discrete event dynamic systems. *IEEE Trans. on Auto. Control AC-32,* (1987), 563–572.
12. Ho, Y. C. and Cassandras, C. A new approach to the analysis of discrete event dynamic systems. *Automatica 19,* (1983), 149–167.
13. Ibragimov, I. A. and Has'minskii, R. Z. *Statistical Estimation: Asymptotic Theory.* Springer-Verlag, New York, (1981).
14. Nakayama, M. K., Goyal, A., and Glynn, P. W. Likelihood ratio sensitivity analysis for Markovian models of highly dependable systems. Tech. Rep. 54, Dept. of Operations Research, Stanford Univ., Stanford, CA, 1989.
15. Reiman, M. I. and Weiss, A. Sensitivity analysis via likelihood ratios. In *Proceedings of the 1986 Winter Simulation Conference,* (1986), 285–289.
16. Reiman, M. I. and Weiss, A. Sensitivity analysis for simulations via likelihood ratios. *Oper. Res. 37,* (1989), 830–844.
17. Rubinstein, R. Y. The score function approach for sensitivity analysis of computer simulation models. *Math. and Comput. in Simulation 28,* (1986), 351–379.
18. Rubinstein, R. Y. Sensitivity analysis and performance extrapolation for computer simulation models. *Oper. Res. 37,* (1989), 72–81.

CR Categories and Subject Descriptors: G.3 [**Mathematics of Computing**]: Probability and Statistics—Probabilistic Algorithms including Monte Carlo

General Terms: Algorithm, Performance

Additional Key Words and Phrases: Gradients, likelihood ratios, optimization, sensitivity analysis, simulation.

About the Author:
PETER W. GLYNN is an associate professor in the Department of Operations Research at Stanford University. His research interests include discrete-event dynamical systems, queuing theory, simulation, and computational probability.

Author's Present Mailing Address: Department of Operations Research, Stanford University, Stanford, CA 94305.