MANAGEMENT SCIENCE
Vol. 38, No. 9, September 1992
Printed in U.S.A.

NOTES

PATHWISE CONVEXITY AND ITS RELATION TO CONVERGENCE OF TIME-AVERAGE DERIVATIVES*

PETER W. GLYNN
Department of Operations Research, Stanford University, Stanford, California 94305–4022

In this note, we further develop the pathwise convexity approach introduced by Hu (1991) to prove consistency of infinitesimal perturbation analysis for the derivative of the steady-state waiting time of the $G/G/1$ queue. In addition to generalizing the argument, we illustrate the technique with applications to stochastic storage theory and networks of queues.
(CONVEXITY; INFINITESIMAL PERTURBATION ANALYSIS; DERIVATIVES; STEADY-STATE)

1. Introduction

Given a stochastic system depending on a real-valued decision parameter θ, it is often of interest to calculate the derivative of steady-state performance measures with respect to θ. These derivatives play an important role in the sensitivity analysis and optimization of such systems (see Glynn 1990 for further discussion of these applications).

Typically, derivatives of steady-state performance measures cannot be calculated in analytical closed-form. As a consequence, it is desirable to develop numerically-based algorithms for obtaining such derivatives. Given the inherent flexibility and power of simulation as a numerical tool for the study of stochastic systems, considerable attention has recently been focussed on the question of how to construct efficient simulation-based estimators for derivatives of steady-state performance measures. Two general approaches to the problem have been suggested in the literature: likelihood ratio gradient estimation techniques (see Glynn 1986, 1987, 1990; Reiman and Weiss 1989; and Rubinstein 1986) and infinitesimal perturbation analysis (see, for example, Suri 1987; Cao 1988; and Glasserman 1988).

This note discusses infinitesimal perturbation analysis (IPA), and can be viewed largely as an elaboration of the basic ideas presented in Hu (1991). A major theoretical concern with IPA is the question of consistency. In our current setting, IPA is said to be consistent if the IPA derivative estimator converges in probability to the derivative of the steady-state performance measure. It is known, however, that IPA need not be consistent when applied to certain types of systems, such as queueing networks having multiple customer-types (see Heidelberger et al. 1988). As a consequence, IPA consistency continues to be a major theoretical issue.

Recently, Hu (1991) developed an elegant approach, based on convexity, for proving consistency of the IPA derivative estimator in the context of the $G/G/1$ queue. In this note, our principal objective is to show that Hu's convexity technique is a powerful tool that can be useful in a significantly broader setting than that of the single-server queue. §2 briefly discusses the approach followed by Hu (1991) to obtain consistency of the IPA derivative estimator. §3 contains the main contribution of the paper, namely a set of

* All Notes are refereed.

Accepted by James R. Wilson; received August 1, 1990. This Note has been with the author 8 months for 1 revision.

0025-1909/92/3809/1360$01.25

examples in which the convexity method is applied to obtain convergence. Of particular interest is Example 2. It is shown there that the IPA derivative estimator is consistent for a class of systems that is provably nonregenerative. This shows clearly that IPA consistency does not rely upon regeneration. The main proof is deferred to the Appendix.

2. The Convexity Approach for Obtaining IPA Consistency

Consider a real-valued discrete-time sequence $X = (X_n : n \geq 0)$ representing the output of a simulation. Suppose that the probability P governing the distribution of X depends on a real-valued decision parameter θ. (We specialize to real-valued decision parameters and discrete-time only in order to simplify our exposition; the same ideas easily extend to vector decision parameters and/or continuous-time processes.) To denote the dependence of P on θ, we write it as P_θ.

We assume that for each θ in some open interval Λ, the sequence X has a well-behaved steady-state. More precisely, we assume that for each $\theta \in \Lambda$, there exists a deterministic constant $\alpha(\theta)$ such that

$$\frac{1}{n}\sum_{k=0}^{n-1} X_k \to \alpha(\theta) \qquad P_\theta \text{ a.s.} \tag{1}$$

as $n \to \infty$. This constant $\alpha(\theta)$ then represents the steady-state mean of X under P_θ. The literature abounds with various mathematical techniques for establishing the strong law of large numbers (1) (e.g., Markov process techniques, stationary process theory, regenerative analysis).

The basic idea underlying IPA is the construction of a single probability space $(\Omega, \mathcal{F}, \tilde{P})$ and a collection of random variables $X(\theta) = \{X_n(\theta) : n \geq 0, \theta \in \Lambda\}$ such that:

$$\text{For each } \theta \in \Lambda, \qquad \tilde{P}\{X(\theta) \in B\} = P_\theta\{X \in B\}, \qquad \text{for each event } B \in \mathcal{F}. \tag{2}$$

Assumption (2) asserts that the distribution of the sequence $X(\theta)$ under $\tilde{P}$ is identical to that of X under P_θ. One (standard) way to construct a probability space satisfying (2) is to use the method of common random numbers to drive each of the processes $X(\theta)$, $\theta \in \Lambda$.

In addition, IPA demands that the construction of $(\Omega, \mathcal{F}, \tilde{P})$ be carried out in such a way that the behavior of $\tilde{X}_n = (X_n(\theta) : \theta \in \Lambda)$ is suitably smooth. In this note, we shall employ a *pathwise convexity* assumption, namely:

$$\text{For each } n \geq 0, \qquad \tilde{P}\{X_n(\theta) \text{ is a convex function of } \theta\} = 1. \tag{3}$$

Assumptions (1), (2), and (3) guarantee that the deterministic function $\alpha(\cdot)$ can be approximated well, in some uniform sense.

PROPOSITION 1. *Suppose $a, b \in \Lambda$, where $a < b$. Under* (1), (2), *and* (3),

$$\tilde{P}\{\bar{X}(\cdot) \text{ converges uniformly to } \alpha(\cdot) \text{ on } [a, b] \text{ as } n \to \infty\} = 1, \qquad \text{where}$$

$$\bar{X}_n(\theta) = n^{-1}\sum_{k=0}^{n-1} X_k(\theta).$$

PROOF. Let $A_\theta = \{\omega : \bar{X}_n(\theta, \omega) \to \alpha(\theta) \text{ as } n \to \infty\}$, and $\tilde{Q}$ be the set of rational numbers contained in Ω. We first show that $\tilde{P}(A) = 1$, where $A = \bigcap_{\theta \in \tilde{Q}} A_\theta$.

By assumption (2), $\tilde{P}(A_\theta) = P_\theta(B_\theta)$, where $B_\theta = \{\omega : n^{-1}\sum_{k=0}^{n-1} X_k(\omega) \to \alpha(\theta) \text{ as } n \to \infty\}$. But (1) guarantees that $P_\theta(B_\theta) = 1$. Since $\tilde{Q}$ is countable, it follows that $\tilde{P}(A) = 1$.

Let $C = \{\omega : \bar{X}_n(\cdot, \omega)$ converges uniformly to $\alpha(\cdot)$ on $[a, b]$ as $n \to \infty\}$. By (3), it is evident that $\bar{X}_n(\cdot, \omega)$ is convex in θ for each $n \geq 0$ and $\omega \in \Omega$. Hence, we may apply Theorem 10.8 of Rockafellar (1970) to conclude that $A \subseteq C$. It follows that $\tilde{P}(C) = 1$, proving the proposition. □

A consequence of Proposition 1 is that there exists an event having $\tilde{P}$-probability one whose occurrence ensures that $\bar{X}_n(\theta)$ converges uniformly for $\theta \in [a, b]$. As Hu pointed out, this permits one to apply Theorem 25.7 of Rockafellar (1970). (Note that any $\theta_0 \in \Lambda$ can be embedded in some closed interval $[a, b] \subseteq \Lambda$.) The following theorem is a slight generalization of the consistency result found in Hu (1991). It proves that the *time-average derivative* converges a.s. to the steady-state derivative at almost every point $\theta_0 \in \Lambda$. The proof follows easily from Proposition 1 above and Lemma 1 of Hu (1991), and is therefore omitted.

THEOREM 1. *Assume* (1), (2), *and* (3) *and let* θ_0 *be a point at which* $\alpha(\cdot)$ *is differentiable. Then*

$$\tilde{P}\{\lim_{n\to\infty} \bar{X}'_n(\theta_0) = \alpha'(\theta_0)\} = 1,$$

where $\bar{X}'_n(\theta_0) = n^{-1}\sum_{k=0}^{n-1} X'_k(\theta_0)$ *and* $X'_k(\theta_0)$ *is the right-hand derivative of* $X_k(\cdot)$ *evaluated at* θ_0, *namely*

$$X'_k(\theta_0) = \lim_{h\downarrow 0} \frac{X_k(\theta_0 + h) - X_k(\theta_0)}{h}.$$

Since $\alpha(\cdot)$ is the pointwise limit of a sequence of convex functions, it is evident that $\alpha(\cdot)$ is convex (Theorem 10.8 of Rockafellar 1970). Consequently α is differentiable except (possibly) on a countable subset of Λ (Theorem 25.3 of Rockafellar 1970). Theorem 1 asserts that IPA is consistent wherever the steady-state performance measure is smooth. Smoothness of the function $\alpha(\cdot)$ (at all points $\theta \in \Lambda$) can be established by techniques that are independent of IPA (for example, likelihood ratio methods).

3. Examples

In this section, we illustrate Theorem 1 with some examples that arise as solutions to a certain class of stochastic recursions. Let $g : \mathbb{R}^{k+1} \to \mathbb{R}$ satisfy:

(i) $g(x, y)$ is convex in $(x, y) \in \mathbb{R} \times \mathbb{R}^k$,

(ii) $g(x, y)$ is nondecreasing in $(x, y) \in \mathbb{R}^{k+1}$, in the sense that if $x_1 \le x_2$ and $y_1 \le y_2$ (componentwise), then $g(x_1, y_1) \le g(x_2, y_2)$.

For a given stochastic sequence $Y = (Y_n : n \ge 1)$ $(Y_n \in \mathbb{R}^k)$, consider the real-valued sequence $X = (X_n : n \ge 0)$ defined by $X_0 = x_0$ (x_0 deterministic) and

$$X_{n+1} = g(X_n, Y_{n+1}) \tag{4}$$

for $n \ge 0$. Assume that, under the distribution P_θ, the process $(Y_n : n \ge 1)$ is i.i.d. with common distribution F_θ. It is easily seen that X is then a real-valued Markov chain under P_θ.

To apply IPA to the calculation of steady-state derivatives of the sequence X defined by (4), we shall assume that there exist an r.v. Y^*, a distribution P^*, and a function $h : \mathbb{R}^{k+1} \to \mathbb{R}^k$ such that:

(iii) $F_\theta(y) = P^*\{h(\theta, Y^*) \le y\}$ for all $\theta \in \mathbb{R}$, $y \in \mathbb{R}^k$,

(iv) $h_i(\cdot, y)$ is convex for $y \in \mathbb{R}^k$, $1 \le i \le k$, where h_i is the ith component of h (i.e., $h(\theta, y) = [h_1(\theta, y), \ldots, h_k(\theta, y)]$).

Let $\tilde{P}$ be the distribution under which $\tilde{Y} = (\tilde{Y}_n : n \ge 1)$ is a sequence of i.i.d. copies of Y^* (generated under P^*) and set $X_0(\theta) = x_0$, with

$$X_{n+1}(\theta) = g(X_n(\theta), h(\theta, \tilde{Y}_{n+1})) \quad \text{for} \quad n \ge 0. \tag{5}$$

PROPOSITION 2. *Under assumptions* (i)–(iv) *above, the above construction of* $\tilde{P}$ *and the* r.v.'s $\{X_n(\theta) : n \ge 0, \theta \in \mathbb{R}\}$ *satisfies* (2) *and* (3).

PROOF. Assumption (2) can be verified by showing that $P[X_{n+1}(\theta) \le y | X_n(\theta) = x]$ has the appropriate distribution for each choice of x and y. Two Markov chains having the same transition probabilities must necessarily have the same probability distribution on their respective path spaces. As for (3), we can prove this inductively. Note that $X_0(\cdot)$ is trivially convex, and assume $X_n(\theta)$ is convex in θ. We need to show that $X_{n+1}(\theta) = g(X_n(\theta), h(\theta, \tilde{Y}_{n+1}))$ is convex in θ. But, by assumptions (i), (ii), and (iv), this follows directly from Theorem 4.1 of Rockafellar (1970). □

With Proposition 2 in hand, we need only verify (1) in order to apply Theorem 1. This must be done on a case-by-case basis (since the Markov chains defined by (4) need not be positive recurrent).

EXAMPLE 1. Consider the waiting time sequence $W = (W_n : n \ge 0)$ associated with the $GI/G/1$ single-server queue. As is well known, the waiting time sequence in a single-server first-come, first-serve queue takes the recursive form ($[x]^+ \triangleq \max(x, 0)$)

$$W_{n+1} = [W_n + V_n - U_{n+1}]^+ \tag{6}$$

for $n \ge 0$, where V_n represents the service time of the nth customer ($n \ge 0$) and U_{n+1} corresponds to the interarrival time between the nth and $(n + 1)$st customers to the system. This is a special case of (4), in which $Y_{n+1} = (V_n, -U_{n+1})$ and $g(x, y_1, y_2) = [x + y_1 + y_2]^+$. We note that g is the composition of a convex nondecreasing function (namely, $[x]^+$) and a linear function (namely, $x + y_1 + y_2$), and hence is convex. Furthermore, g is nondecreasing in each of its arguments, so that it therefore satisfies (i) and (ii).

If we assume that $V = (V_n : n \ge 0)$ and $U = (U_n : n \ge 1)$ are independent sequences of i.i.d. random variables under P_θ, then W is a Markov chain. If we further require that

$$P_\theta\{V_n \in \cdot\} = \tilde{P}\{\theta\tilde{V}_n \in \cdot\}, \qquad P_\theta\{U_{n+1} \in \cdot\} = \tilde{P}\{\tilde{U}_{n+1} \in \cdot\}, \tag{7}$$

then the sequence $W_n(\theta)$ takes the form

$$W_{n+1}(\theta) = [W_n(\theta) + \theta\tilde{V}_n - \tilde{U}_{n+1}]^+. \tag{8}$$

Hence, Proposition 2 applies, therby proving that $(W_n(\theta) : n \ge 0, \theta \in \Lambda)$ satisfies (2) and (3) on any open interval Λ of $\mathbb{R}$. Furthermore, it is known (see Wolff 1989) that if $\tilde{E}\tilde{V}_n^2 < \infty$, then the law of large numbers (1) is valid at any θ satisfying $\theta < \tilde{E}\tilde{U}_{n+1}/\tilde{E}\tilde{V}_n$. Under the above conditions, it is therefore evident that Theorem 1 may be applied to the sequence W, proving validity of IPA at all but (possibly) countably many $\theta_0 \in (0, \tilde{E}\tilde{U}_{n+1}/\tilde{E}\tilde{V}_n)$; see Hu (1990) for another proof of this fact.

We further note that (7) can be modified in several ways without affecting the basic validity of IPA. For example, Proposition 2 continues to apply to changes in location in the service time distribution (i.e., $P_\theta\{V_n \in \cdot\} = \tilde{P}\{\tilde{V}_n + \theta \in \cdot\}$) as well as scale/location changes in the interarrival time distribution (i.e., $P_\theta\{U_{n+1} \in \cdot\} = \tilde{P}\{\theta\tilde{U}_{n+1} \in \cdot\}$ or $\tilde{P}\{\tilde{U}_{n+1} + \theta \in \cdot\}$).

EXAMPLE 2. In this example, we consider a class of nonlinear storage processes that was introduced by Klemes (1978). Given a reservoir, we let S_n be the storage at time n, and let Y_{n+1} denote the inflow during period $n + 1$. If the outflow during period $n + 1$ is assumed to be a power of the storage at time $n = 1$ (i.e., outflow equals aS_{n+1}^b for some $a, b > 0$), then we conclude that the sequence $(S_n : n \ge 0)$ must satisfy the mass-balance equation

$$S_{n+1} = S_n + Y_{n+1} - aS_{n+1}^b. \tag{9}$$

Hence, $S_{n+1} = v(S_n + Y_{n+1})$, where v is the inverse function to $u(x) = x + ax^b$. We note that if $S_0 > 0$, the sequence $(S_n : n \ge 0)$ takes values in $(0, \infty)$. Furthermore, u is twice continuously differentiable on $(0, \infty)$ with $u'(x) = 1 + abx^{b-1}$, $u''(x) = ab(b-1)x^{b-2}$. But $v(u(x)) = x$ and hence

$$v'(u(x))u'(x) = 1,$$
$$v''(u(x))[u'(x)]^2 + v'(u(x)u''(x) = 0.$$

from which we may conclude that $v''(y) = -u''(v(y))/[u'(v(y))]^3$. It follows that v is convex (concave) on $(0, \infty)$ if $0 < b \leq 1$ ($b \geq 1$). Thus, if $0 < b \leq 1$, it is evident that $g(x, y) = v(x + y)$ satisfies conditions (i)–(ii) of Proposition 2. If we further require that, under P_θ, $Y = (Y_n : n \geq 0)$ is a sequence of i.i.d. random variables for which

$$P_\theta\{Y_n \in \cdot\} = \tilde{P}\{\theta \tilde{Y}_n \in \cdot\},$$

the conditions of Proposition 2 are in force and the sequence $S(\theta) = (S_n(\theta) : n \geq 0)$ given by $S_0(\theta) = S_0 > 0$,

$$S_{n+1}(\theta) = v(S_n(\theta) + \theta \tilde{Y}_{n+1})$$

satisfies (2) and (3). Furthermore, we prove in the Appendix that if $\tilde{E}\tilde{Y}_n^2 < \infty$, the strong law (1) holds at every $\theta > 0$ (with $S_n(\theta)$ playing the role of X_n). Hence, Theorem 1 proves that IPA applies to steady-state derivative estimation, for the class of storage models discussed here, provided the power law exponent satisfies $b \leq 1$.

In fact, it turns out that IPA also applies when $b > 1$. Let $\hat{S}_n = -S_n$, $\hat{Y}_n = -Y_n$, and $\hat{g}(x, y) = -g(-x, -y)$, so that $\hat{S}_{n+1} = \hat{g}(\hat{S}_n, \hat{Y}_n)$. Assuming $\tilde{E}\tilde{Y}_n^2 < \infty$, one may then apply Theorem 1 to the chain $(\hat{S}_n : n \geq 0)$ to prove consistency of IPA.

EXAMPLE 3. In this example, we prove that IPA is typically a consistent estimator of the derivative of the mean steady-state waiting time (with respect to service time perturbations) at a first-come, first-serve infinite capacity queueing station. We assume that we have a feed-forward network, so that customers cannot loop back to the station with positive probability. The argument hinges on the fact that the recursion (6) continues to hold at such a station. The sequence of interarrival times ($U_n : n \geq 1$), although no longer i.i.d., is unaffected by a perturbation in the service times at the station. In particular, (8) continues to hold when the perturbation considered is a scale change in the distribution of the service times. As a consequence, the $W_n(\theta)$'s, as in Example 1, continue to be convex in θ. Thus, if the strong law (1) can be shown, in the network setting, to hold on some open interval Λ, IPA consistency follows (as in the proof of Theorem 1), in the sense that the IPA derivative estimator will converge except (perhaps) at countably many points $\theta_0 \in \Lambda$.

In light of the power of this technique, it seems worth exploring necessary conditions for its applicability. We note that if a suitable probability space can be constructed on which (1)–(3) hold, then for any nondecreasing convex function f,

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k(\theta))$$

is convex in θ (since the composition of a nondecreasing convex function with a convex function is convex). Suppose that there exists a steady-state distribution $\pi(\theta)$ for which the strong law holds, with limit given by

$$\frac{1}{n} \sum_{k=0}^{n-1} f(X_k(\theta)) \rightarrow \int_{\mathbb{R}} f(x)\pi(\theta, dx)$$

a.s. as $n \rightarrow \infty$. Such strong laws typically hold for Markov chains. Because convexity is preserved under pointwise limits, it follows that in order for a probability space satisfying (1)–(3) to exist, evidently

$$\alpha(f; \theta) \triangleq \int_{\mathbb{R}} f(x)\pi(\theta, dx)$$

must be convex in θ for any convex nondecreasing function f for which the strong law holds. This is a necessary condition for the pathwise convexity argument used in this paper to be applicable.

To conclude this section, we note that the storage system studied in Example 2 need not be regenerative. In particular, let $b = a = 1$ and let $(Y_n : n \geq 1)$ be a sequence of i.i.d. Bernoulli $(\frac{1}{2})$ r.v.'s In this case, it is easy to show that the uniform distribution on $[0, 2]$ is a stationary distribution for $(S_n : n \geq 0)$. Suppose $S_0 = x$. If there existed some embedded regenerative structure in $(S_n : n \geq 0)$ such that $(S_n : n \geq 0)$ could then be viewed as a positive recurrent regenerative process, it would follow that, for any bounded (measurable) f,

$$\frac{1}{n}\sum_{k=0}^{n-1} f(S_k) \to \int_0^2 f(x)dx/2 \qquad \text{a.s.}$$

as $n \to \infty$. Let $B = \{x + j2^{-k} : j, k \in Z\}$, and note that $S_n \in B$ a.s. Setting $f(x) = I(x \in B)$, we find that the left-hand side of the above limit relation is identically one, whereas the right-hand side is zero. We conclude that $(S_n : n \geq 0)$ cannot be viewed as a positive recurrent regenerative sequence. The importance of this point is that the pathwise convexity argument employed in this paper can be used to establish IPA consistency for certain types of nonregenerative systems. Recent work of Glasserman, Hu, and Strickland (1990) provides conditions for consistency of IPA in the regenerative stochastic process setting. Thus, our work in this paper can be viewed as complementary to that of Glasserman et al. (1990).[1]

[1] This research was supported by the U.S. Army Research Office under Contract DAAL-03-88-K-0063.

Appendix

We prove here that if $\tilde{E}\tilde{Y}_n^2 < \infty$, then the storage sequence $(S_n(\theta) : n \geq 0)$ defined by (9) obeys a law of large numbers, in the sense that, for every $\theta \geq 0$,

$$\frac{1}{n}\sum_{k=0}^{n-1} S_k(\theta) \to \alpha(\theta) \qquad \text{a.s.}$$

as $n \to \infty$, where the constant $\alpha(\theta)$ is deterministic. By Glynn (1989), the Markov chain $(S_n(\theta) : n \geq 0)$ has a unique stationary distribution $\pi(\theta)$. Furthermore, $E_{\pi(\theta)}S_n(\theta) < \infty$ under the condition $\tilde{E}\tilde{Y}_n^2 < \infty$ (see p. 575 of Glynn 1989). By applying the ergodic theorem for stationary sequences (see Doob 1953), it follows that if $\tilde{P}\{S_0(\theta) \in \cdot\} = \pi(\theta)$,

$$\frac{1}{n}\sum_{k=0}^{n-1} S_k(\theta) \to Z(\theta)$$

with $\tilde{P}$-probability one, where $Z(\theta)$ is the conditional expectation of $S_0(\theta)$ with respect to the invariant σ-field. Furthermore, Glynn (1989) proves that $S_n(\theta) - \check{S}_n(\theta) \to 0$ a.s. as $n \to \infty$, where $\check{S}_n(\theta)$ is a storage sequence that has initial condition $\check{S}_0(\theta) = x > 0$ and is driven by the same sequence of inflows as $S_n(\theta)$. As a consequence, the above strong law continues to hold with $S_0(\theta)$ distributed arbitrarily. We may also conclude that $r(x) = \check{P}_x\{Z(\theta) \in B\}$ is independent of x (for any B), where $\check{P}_x(\cdot) = \tilde{P}\{\cdot \mid S_0(\theta) = x\}$.

To complete the proof, we need to show that $Z(\theta)$ is a constant a.s. Since $Z(\theta)$ is invariant, it can be represented as $k(S_{n+1}(\theta), S_{n+2}(\theta), \cdots)$ for some deterministic function k which is independent of n. Hence,

$$\begin{aligned}
&\check{P}_x\{(S_0(\theta), \ldots, S_n(\theta)) \in \cdot, Z(\theta) \in B\} \\
&= \check{E}_x\{\check{P}_x\{k(S_{n+1}(\theta), \cdots) \in B \mid S_0(\theta), \ldots, S_n(\theta)\} \cdot I(S_0(\theta), \ldots, S_n(\theta)) \in \cdot\} \\
&= \check{E}_x\{r(S_n(\theta)) \cdot I(S_0(\theta), \ldots, S_n(\theta)) \in \cdot\},
\end{aligned}$$

where we have used the Markov property for the second equality.

Since $r(\cdot)$ is constant, we conclude that the above probability equals

$$\check{P}_x\{(S_0(\theta), \ldots, S_n(\theta)) \in \cdot\}\check{P}_x\{Z(\theta) \in B\}.$$

Hence, if $r(x) > 0$, we obtain

$$\check{P}_x\{(S_0(\theta), \ldots, S_n(\theta)) \in \cdot \mid Z(\theta) \in B\} = \check{P}_x\{(S_0(\theta), \ldots, S_n(\theta)) \in \cdot\}.$$

As a consequence, we have that

$$\check{P}_x\{S(\theta) \in \cdot \mid Z(\theta) \in B\} = \check{P}_x\{S(\theta) \in \cdot\}.$$

But $Z(\theta)$ is a function of $S(\theta) = (S_n(\theta) : n \geq 0)$ so, for any (measurable) A, we get

$$\check{P}_x\{Z(\theta) \in A \mid Z(\theta) \in B\} = \check{P}_x\{Z(\theta) \in A\}.$$

Taking $A = B$, we have proved that if $\check{P}_x\{Z(\theta) \in B\} > 0$, then $\check{P}_x\{Z(\theta) \in B\} = 1$. In other words, for any B, $\check{P}_x\{Z(\theta) \in B\}$ is either zero or one. It is easy to see that this implies $Z(\theta)$ is deterministic. □

References

CAO, X. R., "A Sample Performance Function of Closed Jackson Queueing Networks," *Oper. Res.*, 36 (1988), 128–136.

DOOB, J. L., *Stochastic Processes*, John Wiley, New York, 1953.

GLASSERMAN, P., "Performance Continuity and Differentiability in Monte Carlo Optimization," *Proc.* 1988 *Winter Simulation Conf.*, San Diego, CA, 518–524.

———, J. Q. HU AND S. G. STRICKLAND, "Strongly Consistent Steady-State Derivative Estimates," *Probability in the Engineering and Informational Sciences*, S (1991), 391–413.

GLYNN, J. E., "A Discrete-Time Storage Process with a General Release Rule," *J. Appl. Probab.*, 26 (1989), 566–583.

———, "Stochastic Approximation for Monte Carlo Optimization," *Proc.* 1986 *Winter Simulation Conf.*, Washington, DC, 356–365.

———, "Likelihood Ratio Gradient Estimation: An Overview," *Proc.* 1987 *Winter Simulation Conf.*, Atlanta, GA, 366–375.

———, "Likelihood Ratio Gradient Estimation Stochastic Systems," *Comm. ACM*, 33, 10 (1990), 75–84.

HEIDELBERGER, P., X. R. CAO, M. A. ZAZANIS AND R. SURI, "Convergence Properties of Infinitesimal Perturbation Analysis Estimates," *Management Sci.*, 34 (1988), 1281–1301.

HU, J.-Q., "Convexity of Sample Path Performance and Strong Consistency of Infinitesimal Perturbation Analysis Estimates," *IEEE Trans. Automat. Control*, 37 (1992), 258–262.

———, "Strong Consistency of Infinitesimal Perturbation Analysis Estimates," Ph.D. thesis, Harvard University, 1990.

KLEMES, V., "Physically Based Stochastic Hydrologic Analysis," *Adv. in Hydrosci.*, 11 (1978), 285–356.

REIMAN, M. I. AND A. WEISS, "Sensitivity Analysis for Simulations via Likelihood Ratios," *Oper. Res.*, 37 (1989), 830–844.

ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

RUBINSTEIN, R. Y., "The Score Function Approach for Sensitivity Analysis of Computer Simulation Models," *Math. Comput. Simulation*, 28 (1986), 1–29.

SURI, R., "Perturbation Analysis of General Discrete Event Systems," *J. Assoc. Comput. Mach.*, 34 (1987), 686–717.

WOLFF, R. W., *Stochastic Modeling and the Theory of Queues*, Prentice-Hall, Englewood Cliffs, NJ 1989.