A NEW CLASS OF STRONGLY CONSISTENT VARIANCE ESTIMATORS FOR STEADY-STATE SIMULATIONS

by

Peter W. Glynn and Donald L. Iglehart

TECHNICAL REPORT NO. 16

October 1986

Prepared under the Auspices of

U.S. Army Research Contract
DAAC29-84-K-0030

Approved for public release: distribution unlimited.

Reproduction in Whole or in Part is Permitted for any Purpose of the United States Government

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under National Science Foundation Grants DCR-85-09668 and ECS-84-04809.
A NEW CLASS OF STRONGLY CONSISTENT VARIANCE ESTIMATORS
FOR STEADY-STATE SIMULATIONS

by

Peter W. Glynn
University of Wisconsin

and

Donald L. Iglehart
Stanford University

ABSTRACT

The principal problem associated with steady-state simulation is the estimation of the variance term in an associated central limit theorem. This paper develops several strongly consistent estimates for this term using the strong approximations available for Brownian motion. A comparison of rates of convergence is given for a variety of estimators.

Keywords: Brownian motion, confidence intervals, rates of convergence, regenerative simulation, simulation output analysis, steady-state simulation, strong approximation laws, strongly consistent estimation.

This research was supported by Army Research Office Contract DAAG29-84-K-0030 and National Science Foundation Grants DCR-85-09668 and ECS-84-04809.
1. INTRODUCTION

Let \( X = (X(t) : t \geq 0) \) be a real-valued (measurable) stochastic process representing the output of a simulation. (To incorporate output sequences \( X_n \) into our framework, we set \( X(t) = X_{[t]} \), where \([t]\) is the greatest integer less than or equal to \( t \).) The process \( X \) is said to possess a steady-state if there exists a finite constant \( r \) such that

\[
(1.1) \quad r(t) \equiv \frac{1}{t} \int_0^t X(s)ds \to r
\]

as \( t \to \infty \), where \( \to \) denotes weak convergence. The problem of consistently estimating and producing confidence intervals for the parameter \( r \) is known, in the simulation literature, as the steady-state simulation problem.

The limit theorem (1.1) suggests that \( r(t) \) can be used to consistently estimate \( r \). It turns out that one can frequently establish that the output process \( X \) in fact satisfies a somewhat stronger relation, namely, there exist finite constants \( r \) and \( \sigma \) such that

\[
(1.2) \quad t^{1/2}(r(t) - r) \to \sigma N(0,1).
\]

Suppose now that one can construct an estimator \( s(t) \) such that

\[
(1.3) \quad s(t) \to \sigma
\]

as \( t \to \infty \). Then, (1.2) and (1.3) together imply that if \( \sigma > 0 \), then the interval
\[ (1.4) \quad [r(t) - z(\delta) \frac{s(t)}{\sqrt{t}}, r(t) + z(\delta) \frac{s(t)}{\sqrt{t}}] \]

is an asymptotic 100(1-\delta)% confidence interval for \( r \), provided that \( z(\delta) \) is chosen to solve the equation \( P(N(0,1) \leq z(\delta)) = 1 - \delta/2 \).

The above discussion suggests that in the presence of an output process satisfying (1.2) with \( \sigma > 0 \), the steady-state simulation problem is, in principle, solved, once an estimator \( s(t) \) for \( \sigma \) has been constructed. Thus, the determination of an estimator \( s(t) \) for \( \sigma \) can be viewed as the fundamental problem of steady-state simulation output analysis. (It should, however, be noted that certain output analysis methods employ a different approach, which does not require explicit estimation of \( \sigma \); the methods of batch means (see pp. 85-89 of BRATLEY, FOX, and SCHRAGE (1983) for a description) and standardized time series (SCHRUBEN (1983)) fall into this category).

As a consequence of the above observation, a great deal of attention has been focused on the construction of such estimators for \( \sigma \). In some sense, all currently available estimation methods make use of the fact that if \( X \) is well-behaved and approximately stationary, then (1.2) suggests that

\[ \lim_{t \to \infty} \text{E}(t(r(t) - r))^2 = \sigma^2 \text{E}(N(0,1))^2 \]

(1.5) \[ \text{i.e., } \int_0^\infty \text{E} X_c(0) X_c(t) dt = \sigma^2 \]
where \( X(t) \equiv X(t) - r \). Spectral procedures (see pp. 95-98 of [3]) use kernel estimators to estimate the left-hand side of (1.5), whereas autoregressive methods (see pp. 98-101 of [3]) fit an autoregressive process to \( X \), compute the analog of (1.5) for the fitted process, and use that quantity as an estimator of (1.5) for \( X \). As for the regenerative method (see pp. 89-94 of [3]), it, loosely speaking, uses the special structure of a regenerative process to "truncates" the integral on the left-hand side of (1.5) at a regeneration time, thereby simplifying the estimation problem.

In this paper, we propose a new method for consistently estimating \( \sigma \), which does not make any explicit use of relationship (1.5). Our estimators are based on known limit theorems about the increments of Brownian motion; the constant \( \sigma \) appears as a scaling constant in these limit theorems. We then use strong approximation methods to translate the resulting estimator for \( \sigma \) from the Brownian motion back to the original output process \( X \); the resulting estimator for \( \sigma \) depends only on the observed values of \( X \), and not the Brownian motion.

In some respects, our method is similar to that of SCHRUBEN (1983). The method of standardized time series depends on the fact that \( \sigma \) appears as a scaling constant when one weakly approximates the original process by a Brownian motion; \( \sigma \) is not estimated but is instead "cancelled" out. By contrast our method involves strong approximation results and gives rise to strongly consistent estimators for \( \sigma \). As proved in GLYNN and IGLEHART (1985), consistent estimation of \( \sigma \) enjoys certain asymptotic advantages over standardized time series.

In Section 2, we introduce our estimators \( \hat{\sigma}(t) \) for \( \sigma \); our basic hypothesis on \( X \) is that a suitable strong approximation by Brownian
motion is possible. Section 3 is devoted to discussion of processes \( \mathbf{X} \) which satisfy the strong approximation hypothesis. In Section 4, the rate of convergence of \( s(t) \) to \( \sigma \) is studied, and compared to that available via the regenerative method. Our main contribution here is to suggest that entirely new methods for estimating \( \sigma \) may be worth exploring. Further comparison of these new methods with the old methods should be carried out via numerical examples.

2. A NEW CLASS OF ESTIMATORS FOR \( \sigma \)

Let \( S = \{S(t): t \geq 0\} \), where \( S(t) = \int_0^t X(s)ds \). Throughout this section, we shall assume that:

\((2.1)\) there exists a probability space supporting a process \( S^* \) and a standard Brownian motion such that:

\((i)\) the distribution of \( S \) equals that of \( S^* \)

\((ii)\) \( S^*(t) - rt = \sigma B(t) + O(t^{1/2 - \lambda}) \) a.s. for some constants \( r, \sigma \), and \( \lambda (0 < \lambda < 1/2, \sigma > 0) \) as \( t \to \infty \).

We shall refer to (2.1) as our strong approximation assumption; it says that a process \( S^* \), possessing precisely the same distribution as \( S \), can be a.s. well approximated by a Brownian motion. Note that under (2.1) (ii),

\[(2.2)\]

\[ t^{-1/2}(S^*(t) - rt) - t^{-1/2}\sigma B(t) + O \text{ a.s.} \]

as \( t \to \infty \). By (2.1) (i), it follows that (2.2) also holds with \( S \) taking the role of \( S^* \); this shows that (2.1) automatically implies (1.2).
For $0 < p < 1$, let

$$s_1(t) = \sup_{0 \leq u \leq t - t^p} \frac{S(u + t^p) - S(u)}{\left(2t^p \cdot (1 - p) \cdot \log t\right)^{1/2}}$$

(2.3) **THEOREM.** If $p$ is chosen so that $p + 2\lambda > 1$, $0 < p < 1$, then $s_1(t) \to \sigma$ a.s. as $t \to \infty$.

**PROOF.** CSÖRGÖ and RÉVÉSZ (1979a) showed that

$$\limsup_{t \to \infty} \frac{B(u + t^p) - B(u)}{\left(2t^p \cdot (1 - p) \cdot \log t\right)^{1/2}} = 1 \text{ a.s.}$$

(2.4)

Let $S_1(t) = S(u) - rt$. By (2.1) (ii), $S_1(t) = \sigma B(t) + O(t^{1/2 - \lambda})$ a.s. so it follows that

$$\sup_{0 \leq u \leq t - t^p} |\sigma B(u + t^p) - \sigma B(u) - S_1(u + t^p) - S_1(u)|$$

$$= O(t^{1/2 - \lambda}) \text{ a.s.}$$

(2.5)

Relations (2.4) and (2.5), together with the condition $2p + \lambda > 1$, imply that

$$\limsup_{t \to \infty} \frac{S_1(u + t^p) - S_1(u)}{\left(2t^p \cdot (1 - p) \cdot \log t\right)^{1/2}} = \sigma \text{ a.s.}$$

(2.6)

i.e.,

$$\limsup_{t \to \infty} \frac{S(u + t^p) - S(u)}{\left(2t^p \cdot (1 - p) \cdot \log t\right)^{1/2}} = \sigma \text{ a.s.}$$

Furthermore, the law of the iterated logarithm for Brownian motion implies that
Applying the strong approximation (2.1) (ii) to (2.7), we find that

\[ r^*(t) - r = O\left(\frac{\log \log t}{t}\right)^{1/2} \text{ a.s.} \]  

where \( r^*(t) = S^*(t)/t \). Since \( (\log \log t)^{1/2} \cdot t^{(p-1)} \to 0 \), it follows from (2.6) and (2.8) that

\[ \lim \sup \frac{S(u+t^p) - t^p r^*(t) - S(u)}{(2t^p (1-p) \cdot \log t)^{1/2}} = \sigma \text{ a.s.} \]

But \( S \) has the same distribution as \( S^* \), so the theorem follows immediately from (2.9).

We can further define the following estimators:

\[ s_2(t) = \sup_{0 < u < t^p} \left| \frac{S(u+t^p) - r(t)t^p - S(u)}{(2t^p (1-p) \cdot \log t)^{1/2}} \right| \]

\[ s_3(t) = \sup_{0 < u < t^p} \sup_{0 < v < t^p} \frac{S(u+v) - r(t)v - S(u)}{(2t^p (1-p) \cdot \log t)^{1/2}} \]

\[ s_4(t) = \sup_{0 < u < t^p} \sup_{0 < v < t^p} \left| \frac{S(u+v) - r(t)v - S(u)}{(2t^p (1-p) \cdot \log t)^{1/2}} \right| \]

\[ s_5(t) = \inf_{0 < u < t^p} \sup_{0 < v < t^p} \frac{S(u+v) - r(t)v - S(u)}{a(t)} \]
where

\[ \sigma(t) = (8 \cdot (1-p) \cdot \log t \cdot \pi^2 t^p)^{1/2} \]

(2.10) **THEOREM.** If \( p \) is chosen so that \( p + 2\lambda > 1 \), \( 0 < p < 1 \), then

\[ s_i(t) \to \sigma \text{ a.s. as } t \to \infty, \ 2 \leq i \leq 5. \]

The proof of this theorem follows exactly as for Theorem 2.4; the key is to have available an analog to (2.4) for Brownian motion, for each of the estimators \( s_i(t), \ 2 \leq i \leq 5 \). For \( s_4(t), \ 2 \leq i \leq 4 \), we refer to Csörgö and Révész (1979a); for \( s_5(t) \), the result can be found in Csörgö and Révész (1979b).

3. **STRONG APPROXIMATION OF STOCHASTIC PROCESSES**

We now proceed to discuss conditions under which Assumption 2.1 holds.

**SETTING 1:** Let \( Y = \{Y(t) : t \geq 0\} \) be a (possibly) delayed regenerative process with regeneration times \( 0 \leq T(0) < T(1) < \cdots \). If \( f \) is a real-valued function defined on the state space of \( Y \), set \( X(t) = f(Y(t)) \). Let 

\[ T(k) = T(k) - T(k-1) \text{ for } k \geq 1, \]

and assume that \( Y \) is positive recurrent in the sense that \( ET(1) < \infty \). Suppose further that there exists \( 0 < \delta < 2 \) such that
\[(3.1) \quad (i) \quad E(\frac{T(1)}{T(0)} |X(s)| ds)^{2+\delta} < \infty \]
\[(ii) \quad E\tau(1)^{1+\delta/2} < \infty \]
\[(iii) \quad EZ^2(1) > 0, \text{ where } Z(n) = \int_{T(n-1)}^{T(n)} (X(s) - \mu) ds \text{ and } \mu = E(\int_{T(0)}^{T(1)} X(s) ds)/E\tau(1). \]

Then, (2.1) is satisfied with \( r = \mu, \sigma^2 = EZ^2(1)/E\tau(1), \) and \( \lambda \) satisfying \( 0 < \lambda < \delta/(4 + 2\delta); \) see pp. 117-122 of PHILIPP and STOUT (1975) for a proof in the case where \( Y \) is a countable state irreducible Markov chain (their argument easily adapts to the more general regenerative setting described above).

**SETTING 2:** Let \( X = \{X_n : n \geq 0\} \) be a strictly stationary sequence of r.v.'s for which there exists \( 0 < \delta < 1 \) such that \( E|X_0|^{2+\delta} < \infty. \) Suppose, in addition, that \( X \) is \( \Phi \)-mixing (see Section 20 of BILLINGSLEY (1968) for a discussion) with mixing coefficients satisfying
\[ \sum_{n=1}^{\infty} \phi(n)^{1/2} < \infty. \]

If
\[ (3.2) \quad \alpha = 2 \int_{0}^{\infty} E(X_0 X_c(s)) ds > 0 \]
(the integral (3.2) converges absolutely), then (2.1) is satisfied with \( r = EX(0), \sigma^2 = \alpha, \) and \( \lambda \) satisfying \( 0 < \lambda < \delta/(24 + 12\delta) \) (if \( X_0 \) is bounded a.s., then \( \lambda \) can be chosen to be \( 1/12 \)). This result can be found on pp. 26-38 of [11]. (For a version of this result in the case when \( \sum_{n=1}^{\infty} \phi(n)^{1/2} < \infty, \) see BERKES and PHILIPP (1979).)
Further strong approximation theorems are also available for lacunary trigonometric series, martingale processes, Gaussian sequences, and strongly mixing processes; see [11] for a complete description of such results. Thus, the assumption (2.1) is satisfied by a large class of stochastic processes exhibiting weak time dependencies.

We should also comment that the convergence rates (i.e., the size of \( \lambda \)) quoted above for regenerative processes and \( \Phi \)-mixing sequences can probably be improved. For example, much better results are available for sequences of independent and identically distributed (i.i.d) r.v.'s. In particular, it is reasonable to expect that results for regenerative processes can be obtained for \( \lambda \) arbitrarily close to 1/2.

**SETTING 3:** Let \( X = \{ X_n : n \geq 0 \} \) be a sequence of non-degenerate i.i.d r.v.'s with \( E|X_0|^p < \infty \) for \( p > 2 \). Then, (2.1) is satisfied with \( r = EX_0, \sigma^2 = \text{var}(X_0), \) and \( \lambda = 1 - 1/p; \) see KOMLÓS, MAJOR, and TUSNÁDY (1975, 1976) and MAJOR (1976) for proofs.

4. **COMPARISON OF CONVERGENCE RATES**

It has been shown by RÉVÉSZ (1980) that if \( 0 < p < 1 \), then

\[
\lim_{t \to \infty} (\log t)^{1/2 + \delta} \left| \sup_{0 \leq u \leq -t^p} \frac{B(u + t^p) - B(u)}{(2t^p \cdot (1-p) \cdot \log t)^{1/2}} - \sigma \right| = 0 \text{ a.s.}
\]

\[
\lim_{t \to \infty} (\log t)^{1/2 + \delta} \left| \sup_{0 \leq u \leq -t^p} \sup_{0 < v < t^p} \frac{B(u + v) - B(u)}{(2t^p \cdot (1-p) \cdot \log t)^{1/2}} - \sigma \right| = 0 \text{ a.s.}
\]

for \( \delta = 0; \) it is further indicated in CSÖRGÖ and STEINEBACH (1981) that for \( \delta > 0, \) the above \( \lim sup's \) are infinite.
Using the strong approximation (2.1), it follows that

$$\lim_{t \to \infty} (\log t)^{(1/2)+\delta} \left| s_i(t) - \sigma \right| = 0 \text{ a.s.}$$

for $\delta = 0$ $(i = 1, 3)$, whereas divergence occurs if $\delta > 0$. Thus, the rate of convergence of $s_i(t)$ $(i = 1, 3)$ to $\sigma$ is, roughly speaking, of order $(\log t)^{-1/2}$.

It is instructive to compare this rate of convergence to that available when $\sigma$ is estimated via the regenerative method of simulation (we choose this method as a basis for comparison, since we can do the convergence rate analysis easily in this setting).

Let $Y$ be a regenerative process with regeneration times $0 < T(0) < T(1) < \cdots$; set $X(t) = f(Y(t))$, where $f$ is a real-valued function defined on the state space of $Y$. Put $T(-1) = 0$ and let $N(t) = \max \{ k \geq -1 : T(k) \leq t \}$. The basic regenerative estimator for $\sigma$ is given by

$$s(t) = \begin{cases} 
0 & ; N(t) \leq 0 \\
\frac{1}{t} \sum_{i=1}^{N(t)} (V_i - r(t) \tau_i)^2 ]^{1/2} & ; N(t) \geq 1,
\end{cases}$$

where $V_i = \int_{T(i-1)}^{T(i)} X(s) ds$ and $\tau_i = T(i) - T(i-1)$. 


(4.2) **THEOREM.** If \( \frac{T(i)}{X(s)} \leq \frac{T(i-1)}{1+1} ds < \),

\[
\lim_{t \to \infty} \sqrt{\frac{t}{2 \log \log t}} |s(t) - \sigma| = \beta^{1/2} \quad \text{a.s.,}
\]

where

\[
\beta = \frac{1}{4\sigma^2 \cdot E\tau_1} \cdot E[A_1 - \lambda Z_1]^2,
\]

\[
Z_n = V_n - \tau_n,
\]

\[
A_n = Z_n^2 - \sigma^2 \tau_n, \quad \text{and}
\]

\[
\lambda = 2 \frac{E Z_1 \tau_1}{E \tau_1}.
\]

Recall that for regenerative processes, \( r = EV_1/E\tau_1 \) and \( \sigma^2 = E Z_1^2 / E \tau_1 \), so that \( Z_n \) and \( A_n \) are mean-zero r.v.'s; we will need this fact in our proof of Theorem 4.2. Also, we remark that Theorem 4.2 is a statement of the law of the iterated logarithm for the estimator \( s(t) \).

**PROOF (of Theorem 4.2):** On the event \( \{N(t) \geq 1\} \), observe that if \( v(t) = s^2(t) \), then
(4.3) \[ v(t) - \sigma^2 = \frac{1}{t} \sum_{i=1}^{N(t)} z_i^2 \]
\[ + 2(r - r(t)) \cdot \frac{1}{t} \sum_{i=1}^{N(t)} z_i \tau_i \]
\[ + (r - r(t))^2 \cdot \frac{1}{t} \sum_{i=1}^{N(t)} \tau_i^2 \]
\[ = \frac{1}{t} \sum_{i=1}^{N(t)} (A_i - \lambda Z_i) \]
\[ - 2 \int_{T(N(t))}^t (X(s) - r)ds \cdot \frac{1}{t} \sum_{i=1}^{N(t)} z_i \tau_i \]
\[ + 2 (r - r(t)) \cdot \left( \frac{1}{t} \sum_{i=1}^{N(t)} z_i \tau_i - \frac{E Z_i \tau_i}{E \tau_i} \right) \]
\[ + (r - r(t))^2 \cdot \frac{1}{t} \sum_{i=1}^{N(t)} \tau_i^2 . \]

But \( s(t) = g(v(t)) \) where \( g(x) = x^{1/2} \), so by Taylor's theorem, we have

\[ s(t) = \sigma + g'(v(t)) (v(t) - \sigma^2) \]

where \( \xi(t) \) lies between \( v(t) \) and \( \sigma^2 \) and \( g'(x) = 1/(2x^{1/2}) \). Since \( v(t) + \sigma^2 \) a.s., it follows that \( g'(\xi(t)) + 1/(2\sigma) \) a.s. Thus, to prove the theorem, it suffices to show that

(4.4) \[ \lim_{t \to \infty} \sqrt{\frac{t}{2 \log \log t}} \left| v(t) - \sigma^2 \right| = 2\sigma^{1/2} \text{ a.s.} \]
By the Hartman-Wintner law of the iterated logarithm and the fact that
\[ N(t) \to a.s., \] we have that
\[
\lim_{t \to \infty} \sqrt{\frac{N(t)}{2 \log \log N(t)}} \left| \frac{1}{N(t)} \sum_{i=1}^{N(t)} (A_i - \lambda Z_i) \right| = 2\sigma \beta^{1/2} (E \tau_1)^{1/2} \ a.s.
\]

But \( N(t)/t \to 1/E \tau_1 \ a.s. \) as \( t \to \infty \), so

\[
\lim_{t \to \infty} \sqrt{\frac{t}{2 \log t}} \left| \frac{1}{t} \sum_{i=1}^{N(t)} (A_i - \lambda Z_i) \right| = 2\sigma \beta^{1/2} \ a.s.,
\]

\[
\lim_{t \to \infty} \frac{t}{\log t} \left| (r - r(t)) \cdot \left( \frac{1}{t} \sum_{i=1}^{N(t)} Z_i \tau_i - \frac{E \tau_1}{E \tau_1} \right) \right| < \infty \ a.s., \text{ and}
\]

\[
\lim_{t \to \infty} \frac{t}{\log t} \left| (r - r(t))^2 \cdot \left( \frac{1}{t} \sum_{i=1}^{N(t)} \tau_i^2 \right) \right| < \infty \ a.s.
\]

Furthermore,
\[
\int_0^t |X(s) - r| \, ds \leq \max_{1 \leq k \leq N(t)+1} \left( \int_{T(k)}^{T(k+1)} |X(s)| \, ds + |r| \tau_k \right).
\]

Our moment hypothesis allows one to apply the Borel-Cantelli lemma to obtain
\[
\left( \int_{T(k)}^{T(k+1)} |X(s)| \, ds + |r| \tau_k \right) / k^{1/4} \to 0 \ a.s.
\]
as \( k \to \infty \); this shows that

\[
\frac{1}{t^{1/4}} \max_{1 \leq k \leq N(t)+1} \left( \int_{T(k-1)}^{T(k)} |X(s)| ds + |\mathbf{r}|_{s_k} \right) \to 0 \text{ a.s.}
\]

and \( t \to \infty \). Hence

\[
\lim_{t \to \infty} t^{3/4} \left| \int_{T(N(t))}^{t} (X(s) - \mathbf{r}) ds \cdot \frac{1}{t^2} \sum_{i=1}^{N(t)} Z_i \tau_i \right| = 0 \text{ a.s.}
\]

as \( t \to \infty \). Combining (4.5), (4.6), (4.7), and (4.8), we see that the decomposition (4.3) yields (4.4).

Roughly speaking, Theorem 4.2 says that \( s(t) \) converges to \( \sigma \) at rate \( (\log \log t/t)^{1/2} \). By comparison with the previously obtained convergence rate for \( s_i(t) \) (i=1,3), this is much faster. This does not necessarily, however, imply that the estimators \( s_i(t) \) for \( \sigma \) will behave worse than the regenerative estimator for purposes of confidence interval generation.
REFERENCES


A NEW CLASS OF STRONGLY CONSISTENT VARIANCE ESTIMATORS FOR STEADY-STATE SIMULATIONS

Peter W. Glynn and Donald L. Iglehart

Department of Operations Research
Stanford University
Stanford, CA 94305

U.S. Army Research Office
Post Office Box 12211
Research Triangle Park, NC 27709

DAAG29-84-K-0030

October 1986

UNCLASSIFIED

This document has been approved for public release and sale; its distribution is unlimited.

The principal problem associated with steady-state simulation is the estimation of the variance term in an associated central limit theorem. This paper develops several strongly consistent estimates for this term using the strong approximations available for Brownian motion. A comparison of rates of convergence is given for a variety of estimators.
END
5 - 87
PTIC