

LARGE-SAMPLE THEORY FOR STANDARDIZED TIME SERIES: AN OVERVIEW

Peter W. Glynn  
 Department of Industrial Engineering  
 University of Wisconsin, Madison, Wisconsin 53706

Donald L. Iglehart  
 Department of Operations Research  
 Stanford University, Stanford, California 94305

There are two basic approaches to constructing confidence intervals for steady-state parameters from a single simulation run. The first is to consistently estimate the variance constant in the relevant central limit theorem. This is the approach used in the regenerative, spectral, and autoregressive methods. The second approach (standardized time series, STS) due to SCHRUBEN [10] is to "cancel out" the variance constant. This second approach contains the batch means method as a special case. Our goal in this paper is to discuss the large-sample properties of the confidence intervals generated by the STS method. In particular, the asymptotic (as run size becomes large) expected value and variance of the length of these confidence intervals is studied and shown to be inferior to the behavior manifested by intervals constructed using the first approach.

1. INTRODUCTION

Standardized time series has attracted considerable attention in the simulation community, since its introduction by SCHRUBEN [10]. Although the method has been exploited for the purpose of detecting initialization bias (see [9]), our intent here is to focus solely on the application of standardized time series to generation of confidence intervals for steady-state simulations. We shall describe, for these confidence intervals, certain large-sample results first obtained in GLYNN and IGLEHART [5]. We start, in Section 2, by outlining the steady-state confidence interval problem. In Section 3, the method of standardized time series is described, while Section 4 is devoted to the large-sample theory for the confidence intervals generated. Section 5 concludes with a description of future research problems associated with the method of standardized time series.

2. THE STEADY-STATE CONFIDENCE INTERVAL PROBLEM

Let  $Y = \{Y(t) : t \geq 0\}$  be a real-valued stochastic process representing the output of a simulation. (To incorporate stochastic sequences  $\{Y_n : n \geq 0\}$  into our framework, set  $Y(t) = Y[\lfloor t \rfloor]$ , where  $\lfloor t \rfloor$  is the greatest integer less than or equal to  $t$ .) Assume that  $Y$  possesses a steady-state, in the sense that

$$r(t) \equiv \frac{1}{t} \int_0^t Y(s) ds \Rightarrow r \quad (2.1)$$

as  $t \rightarrow \infty$  ( $\Rightarrow$  denotes weak convergence), where  $r$  is a finite deterministic constant.

The steady-state estimation problem of simulation has basically two parts:

- i) construct a consistent estimator for the steady-state parameter  $r$
- ii) assess the variability of the estimator constructed in i).

Given (2.1), the solution to part i) is trivial: namely, given an observation of the process  $Y$  over the interval  $[0, t]$ , use  $r(t)$  (or some asymptotically equivalent variant thereof) as our point

estimator for  $r$ . This method for estimating  $r$  is known as a **single-replicate** method. Although multiple replicate procedures for estimating  $r$  are also available (see GLYNN [4] for the relevant large-sample theory), we shall restrict our attention here to single-replicate methods.

Part ii) is the crux of the steady-state estimation problem; it has attracted a great deal of attention in the simulation literature over the last ten years. The traditional variability assessment technique, which has been widely adopted in the simulation community, is to construct confidence intervals for the point estimator. It should be emphasized, however, that there are other possible methods available for assessing point estimator variability. For example, one could use an estimate of the mean square error (MSE) of the point estimator for the same purpose.

In any case, confidence interval generation has generally been regarded as the most reasonable means for assessing point estimator variability. Adopting this viewpoint, most published analyses have required assuming a strengthened version of (2.1), namely existence of a finite positive constant  $\sigma$  such that

$$t^{1/2}(r(t) - r) \Rightarrow \sigma N(0,1) \quad (2.2)$$

as  $t \rightarrow \infty$ , where  $N(0,1)$  is a normal random variable (RV) with mean zero and unit variance.

To use the central limit theorem (CLT) (2.2) directly to generate confidence intervals, one needs a process  $\{s(t) : t \geq 0\}$ , with the following properties:

- i)  $s(t)$  may be constructed from the process  $Y$  observed over the interval  $[0, t]$
- ii)  $s(t) \Rightarrow \sigma$  as  $t \rightarrow \infty$ . (2.3)

Property (2.3) ii) says that  $s(t)$  consistently estimates the parameter  $\sigma$ . Given such a process  $\{s(t) : t \geq 0\}$ , the interval

$$[r(t) - z_\delta \frac{s(t)}{\sqrt{t}}, r(t) + z_\delta \frac{s(t)}{\sqrt{t}}]$$

can easily be shown to be an asymptotic  $100(1-\delta)\%$  confidence interval for  $r$ , provided that  $z_\delta$  is chosen as a solution to  $P\{N(0,1) \leq z_\delta\} = 1-\delta/2$ . Thus, under the hypothesis (2.2), the confidence interval generation problem is reduced to construction of consistent estimators for  $\sigma$ . We will call confidence interval procedures which follow this line of attack **consistent estimation procedures**.

Note that  $\sigma$  has an important statistical interpretation. If the process  $\{(r(t)-r)^2 : t > 0\}$  is uniformly integrable (this should be viewed as a mathematical regularity condition which, from a practical viewpoint, is just slightly stronger than the assertion (2.2)), then we can pass expectations through the limit theorem obtained by squaring both sides of (2.2), thereby yielding

$$t E(r(t)-r)^2 \rightarrow \sigma^2, \tag{2.4}$$

as  $t \rightarrow \infty$ . In other words, the MSE of the point estimator  $r(t)$  is asymptotic to  $\sigma^2/t$ . Thus, a consistent estimation procedure not only produces a confidence interval for  $r$ , but also an estimate for the MSE of the point estimator  $r(t)$ , namely  $s^2(t)/t$ .

We view the following well-known techniques as consistent estimation procedures:

- i) the regenerative method
- ii) spectral procedures (with bandwidth going to zero as sample size  $t \rightarrow \infty$ )
- iii) autoregressive methods (with model order going to infinity as sample size  $t \rightarrow \infty$ )
- v) batch means (with both the batch size and number of batches going to infinity as sample size  $t \rightarrow \infty$ ).

Mathematical theory verifying consistency of the regenerative estimator is available (see, for example, IGLEHART [8]); we are not, however, aware of such consistency results for the other three techniques mentioned above, in the context of the non-stationary processes which arise in simulation.

A totally different approach to the confidence interval generation problem is, however, available. A glance at the limit theorem (2.2) indicates that  $\sigma$  acts, in some sense, as a scaling parameter. This suggests that one can eliminate  $\sigma$  entirely from the limit theorem (2.2) by passing to a "dimensionless" version of the result: this is the basic idea behind standardized time series.

More specifically, suppose that one can find a non-vanishing process  $\{Z(t) : t \geq 0\}$  such that

$$(t^{1/2}(r(t)-r), Z(t)) \Rightarrow (\sigma N(0,1), \sigma Z) \tag{2.4}$$

as  $t \rightarrow \infty$ ; it follows, from the continuous mapping theorem (see p. 31 of BILLINGSLEY [1]), that

$$t^{1/2}(r(t)-r)/Z(t) \Rightarrow N(0,1)/Z \tag{2.5}$$

as  $t \rightarrow \infty$ .

At first glance, it appears that (2.5) should be enough to construct a confidence interval for  $r$ . However, it is clear that one possible choice for  $\{Z(t) : t > 0\}$  is  $Z(t) \equiv \sigma$ . While this is, of course, a mathematically reasonable possibility, it cannot be implemented from a practical viewpoint since  $Z(t)$  depends on the unknown parameter  $\sigma$ .

One might also attempt to use processes  $\{Z(t) : t > 0\}$ , in which  $Z(t) = Z(t,r)$  is a function of the parameter  $r$ . Clearly, then,  $Z(t,r)$  must have a very special dependence on  $r$ , in order that the limit theorem (2.4) may be "unfolded" to reveal a confidence interval for  $r$  of the form:

$$[r(t) - \frac{L(t)}{t^{1/2}}, r(t) + \frac{U(t)}{t^{1/2}}], \tag{2.6}$$

when  $L(t), U(t)$  do not depend on  $r$  (or any other unknown parameters). To the best of our knowledge, the only confidence interval of the form (2.6), in which the scaling process  $Z(t)$  depends explicitly on  $r$ , is the Fieller interval (see IGLEHART [7]); the construction of this interval is possible only because of the quadratic dependence of  $Z(t)$  on  $r$ .

In light of the above discussion, it seems reasonable to limit ourselves to processes  $\{Z(t) : t \geq 0\}$  such that:

$$Z(t) \text{ does not depend on any unknown parameters.} \tag{2.7}$$

Because (2.7) is not mathematically precise, we shall need to describe this assumption in more detail in the next section.

In any case, (2.4) and (2.7) form the core of the method of standardized time series. In view of (2.5), we shall call procedures based on (2.4) and (2.7) **cancellation methods**. If we adopt the convention that  $Z$  cannot be deterministic, then cancellation methods and consistent estimation procedures are disjoint families of techniques, and we arrive at a decomposition of steady-state confidence interval methods which is described by Figure 1.

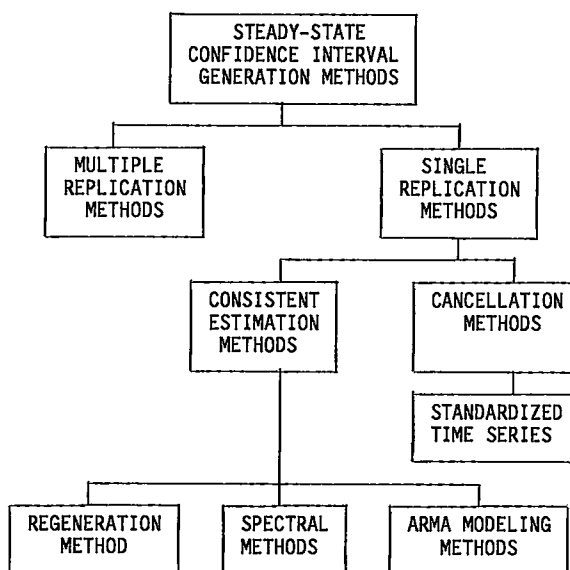


Figure 1

### 3. STANDARDIZED TIME SERIES

As indicated in the previous section, standardized time series is a cancellation method; this point was made by SCHRUBEN in [10]. Recall that  $\sigma$  is a measure of the variability of  $r(t)-r$ . Thus, if  $Z(t)$  is to be proportional to  $\sigma$ , it seems reasonable for  $Z(t)$  to look at the fluctuations of  $r(t)-r$  as a function of time. This suggests that one should base one's analysis on the process  $r(\cdot)-r$ , as opposed to merely the point estimate  $r(t)-r$ .

In keeping with this viewpoint, it is natural to extend the CLT (2.2) to a functional central limit theorem (FCLT). A FCLT provides a distributional approximation for the entire process  $r(\cdot)-r$ . To be more precise, let

$$\bar{Y}_n(t) = \frac{1}{n} \int_0^{nt} Y(s) ds, \quad 0 \leq t \leq 1.$$

Given that  $Y$  possesses a steady-state (see (2.1)), it follows, upon observing that  $\bar{Y}_n(t) = t r(nt)$ , that

$$\bar{Y}_n(t) \Rightarrow rt$$

as  $n \rightarrow \infty$ . To obtain some sense of the random fluctuations of  $\bar{Y}_n(t)$  around  $rt$ , it is necessary to "scale up" the difference  $\bar{Y}_n(t) - rt$ . To be consistent with the CLT (2.2), the scaling factor must clearly be  $n^{1/2}$ . So, consider the random function

$$X_n(t) \equiv n^{1/2}(\bar{Y}_n(t) - rt),$$

for  $0 < t < 1$ . An FCLT provides an approximation, in distribution, for the process  $X_n = \{X_n(t) : 0 \leq t \leq 1\}$ .

To determine the limit  $X$  of the sequence  $X_n$ , observe that the continuity of  $X_n$  suggests that the limit  $X$  ought to be continuous in  $t$ . Furthermore, consider the increments of  $X_n$ :

$$X_n(t+s) - X_n(t) = n^{1/2} \left( \frac{1}{n} \int_{nt}^{n(t+s)} Y(u) du - rs \right)$$

$$X_n(t) - X_n(0) = n^{1/2} \left( \frac{1}{n} \int_0^{nt} Y(u) du - rt \right). \quad (3.1)$$

Note that  $X_n(t) - X_n(0) \approx X_n(t-\epsilon) - X_n(0)$  for  $\epsilon$  small and positive. From (3.1), it follows that  $X_n(t+s) - X_n(t)$  depends only on the evolution of  $Y$  after time  $nt$ , whereas  $X_n(t-\epsilon) - X_n(0)$  depends only on  $Y$  up to time  $nt-\epsilon$ . As  $n \rightarrow \infty$ , one expects that events occurring  $\epsilon$  time units apart are asymptotically independent; consequently, the limit process  $X$  should have independent increments, on the sense that for  $t_1 < t_2 < \dots < t_n$ , the random variables  $X(t_2) - X(t_1), \dots, X(t_n) - X(t_{n-1})$  are independent. (Extending our heuristic argument above from two increments to  $n$  increments is easy.) Finally, if the process  $Y$  is time-homogeneous, it makes sense (see (3.1)), that

$$X_n(t+s) - X_n(t) \stackrel{D}{\approx} X_n(s) - X_n(0)$$

( $\stackrel{D}{\approx}$  means "has approximately the same distribution as"). Passing to the limit we therefore expect that  $X$  will have stationary increments, in the sense that  $X(t+s) - X(t) \stackrel{D}{\approx} X(s) - X(0)$ . ( $\stackrel{D}{\approx}$  denotes equality in distribution.)

To summarize, we expect that the limit process  $X$ :

- i) is continuous
- ii) has independent increments
- iii) has stationary increments. (3.2)

It turns out that (3.2) uniquely characterizes the possible limit  $X$ ; namely,  $X$  must be a Brownian motion (see BREIMAN [2], Chapter 12, for definitions and results).

A Brownian motion process must necessarily be Gaussian, which is consistent with the CLT (2.2).

We can now state our FCLT assumption on  $Y$ :

$$X_n(\cdot) \Rightarrow \sigma B(\cdot) \quad (3.3)$$

as  $n \rightarrow \infty$ , where  $B(\cdot)$  is a standard Brownian motion (i.e.,  $B(0) = 0$ ,  $\text{var } B(t) = t$ ). Note that (3.3) refers to weak convergence of the random function  $X_n(\cdot)$  to the random function  $\sigma B(\cdot)$ . Thus, our FCLT assumption requires an understanding of the notion of weak convergence of random functions, as opposed to the more standard notion of weak convergence of random variables. We shall not go into further detail on this point, except to mention that a brief discussion of the relevant convergence issues is given in Section 2 of GLYNN and IGLEHART [5].

Of more practical importance is the need to have an understanding of when a FCLT holds. As our heuristic argument indicates, one basically needs a time-homogeneous process with an appropriate asymptotic independence structure. To be more concrete, assume that  $Y$  is derived from a stochastic sequence obeying a recursion of the form

$$Y_{n+1} = h(n, Y_n, \eta_{n+1}) \quad (3.4)$$

where  $\eta_{n+1}$  is independent of  $Y_0, Y_1, \dots, Y_n$ . In order that  $Y$  be time-homogeneous, it is necessary that the  $\eta_n$ 's be identically distributed and that  $h(n, y, \eta)$  not depend explicitly on  $n$ .

From a simulation standpoint, the time-homogeneity requirement can be stated as:

The state transition rules and clock-resetting rules should not contain any explicit time-dependence.

Thus, a simulation which explicitly incorporates time trends is not time-homogeneous.

The asymptotic independence assumption is easily understood, from an intuitive standpoint: events in the distant past should not affect events in the distant future too much. Mathematically, this requirement is often stated in terms of mixing conditions:

$\phi$ -mixing: This is a very restrictive type of mixing condition, which basically says that

$$|P(A|B) - P(A)|$$

should be uniformly small in both A and B, provided that A and B are events separated by a long time period. (See p. 166-167 of BILLINGSLEY [1] for a rigorous definition.)

**Strong mixing:** This is a less restrictive type of mixing condition, stating that

$$|P(A \cap B) - P(A)P(B)|$$

ought to be uniformly small in both A and B, provided that A and B are events separated by a long time period. (See HALL and HEYDE [6] for a definition.)

Typically, in order to obtain a FCLT, certain rate conditions are imposed on the mixing coefficients; these conditions tend to be more stringent for strong mixing processes than  $\phi$ -mixing processes.

It turns out that regenerative processes that are positive recurrent (i.e., finite expected time between regenerations) are automatically strong mixing. (See GLYNN [3] for a proof.) For a regenerative process to be  $\phi$ -mixing, it is necessary that regenerations occur "uniformly fast" throughout the state space. More precisely, if the regenerative process is time-homogeneous Markov on a state space S, then we require that there exist  $n > 1$  such that state space S, then we require that there exist  $n \geq 1$  such that

$$\sup\{P\{T > n | Y_0 = x\} : x \in S\} < 1, \quad (3.5)$$

where T is the first regeneration time for Y. Note that (3.5) disqualifies any birth-death type process from being  $\phi$ -mixing, since the hitting time of any fixed state cannot be uniformly bounded; in fact, the hitting time of any fixed state tends to infinity as the initial state goes to infinity. By similar reasoning, virtually any simulation of an open queueing network (with an unlimited number of potential customers) will give rise to a process which is not  $\phi$ -mixing. (However, such processes will frequently be strong mixing.)

In terms of mathematical conditions guaranteeing that Y satisfy the FCLT (3.3), the reader is referred to Section 3 of [5]. It turns out that, mathematically speaking, one "almost always" gets a FCLT whenever a CLT holds, so that the FCLT assumption is only slightly stronger than the CLT assumption. From a practical standpoint, we repeat that an FCLT holds whenever one has time-homogeneity and asymptotic independence, asymptotic independence usually being described by mixing hypotheses.

We need now to discuss how the FCLT (3.3) is used to obtain the scaling process  $\{Z(t) : t > 0\}$ . Let g be a real-valued function defined on a function space, so that  $g(X_n)$  makes sense as a random variable. Assume that:

- i)  $g(\alpha x) = \alpha g(x)$  for all  $\alpha > 0$  and functions x (this is the scaling property for Z(t)).
- ii)  $g(B) > 0$  w.p.1 (this guarantees that Z(t) is non-vanishing for t sufficiently large) (3.6)
- iii)  $g(x + \beta k) = g(x)$  for all  $\beta$ , where  $k(t) = t$  ( $x + \beta k$  is the function whose value at t is  $x(t) + \beta k(t)$ ).

This guarantees that  $g(X_n)$  does not depend implicitly on the unknown parameter r (see (2.7)).

In addition, we want g to be suitably continuous, but we shall not be precise here about this. (See Section 3 of [5] for details.)

Let  $h(x) = x(1)/g(x)$ . By the continuous mapping lemma (this is where we need g to be suitably continuous), and (3.3)

$$h(X_n) \Rightarrow h(\sigma B),$$

as  $n \rightarrow \infty$ . But  $h(\sigma B) = \sigma B(1)/g(\sigma B) = \sigma B(1)/\sigma g(B) = B(1)/g(B)$ , so that the  $\sigma$  has "cancelled out." On the other hand,

$$\begin{aligned} h(X_n) &= \frac{n^{1/2}(\bar{Y}_n(1) - \mu)}{g(n^{1/2}(\bar{Y}_n - k\mu))} \\ &= \frac{(\bar{Y}_n(1) - \mu)}{g(\bar{Y}_n - k\mu)} \\ &= \frac{\bar{Y}_n(1) - \mu}{g(\bar{Y}_n)}. \end{aligned}$$

So, assuming that Y satisfies an FCLT and g satisfies (3.6), the interval

$$[\bar{Y}_n(1) - z(1-\alpha/2) g(\bar{Y}_n),$$

$$\bar{Y}_n(1) - z(\alpha/2) g(\bar{Y}_n)] \quad (3.7)$$

is an asymptotic 100 (1- $\delta$ )% confidence interval for r, where  $H(z(\alpha)) = \alpha$ , and  $H(\cdot) = P\{B(1)/g(B) < \cdot\}$  (we implicitly assumed here that H was a strictly increasing continuous distribution function, but this turns out to be justified; see Section 3 of [5]).

The process  $(\bar{Y}_n - rk)/g(\bar{Y}_n)$  is called a **standardized time series**. Letting  $\mathcal{M}$  be the set of g's satisfying (3.6), we find that each  $g \in \mathcal{M}$  gives rise to a standardized time series. Among possible choices for g are:

- i) batch means with the number of batches fixed as sample size tends to infinity (in other words, there exists a choice of g giving rise to batch means confidence intervals).
- ii) Schruben's standardized sum process (see [10]).
- iii) Schruben's standardized maximum process (see [10]).

For more detail on possible choices for g, see Section 4 of [5].

It should be noted that in order to incorporate the method of batch means, with the number of batches increasing with the sample size, it is necessary to consider Z(t)'s of the form  $g_n(X_n)$ , where  $g_n \in \mathcal{M}$ .

One then needs to examine weak convergence of  $X_n(1)/g_n(X_n)$ ; this type of weak convergence problem tends to be much more complicated to analyze mathematically than the case in which  $g_n(\cdot)$  is independent of  $n$ . Fortunately, with our definition of standardized time series, there is no need to consider such convergence problems.

To summarize the above discussion, we have shown that each  $g \in \mathcal{M}$  gives rise to a standardized time series confidence interval which is asymptotically valid.

#### 4. FURTHER PROPERTIES OF STANDARDIZED TIME SERIES

As indicated in Section 3, each standardized time series procedure gives rise to asymptotically valid confidence intervals. We now wish to discuss the desirability of these intervals. The discussion of desirability will focus on two issues:

- i) asymptotic mean of the confidence interval half-length.
- ii) asymptotic variability of the confidence interval half-length.

Let  $L_n(g)$  be the length of the confidence interval based on observing  $Y$  over the interval  $[0, n]$  and using a standardized time series based on  $g \in \mathcal{M}$ . Clearly,

$$L_n(g) = (z(1-\alpha/2) - z(\alpha/2)) g(\bar{Y}_n) ,$$

by (3.7). Noting that

$$\begin{aligned} g(\bar{Y}_n) &= g(\bar{Y}_n - kr) = n^{-1/2} g(n^{1/2}(\bar{Y}_n - kr)) \\ &= n^{-1/2} g(X_n) , \end{aligned}$$

we see that

$$n^{1/2} L_n(g) \Rightarrow (z(1-\alpha/2) - z(\alpha/2)) \sigma g(B) ,$$

as  $n \rightarrow \infty$ . Thus,  $L_n(g)$  is of order  $n^{-1/2}$ . As is clear from Section 2, this is the same order convergence rate as that obtained for consistent estimation procedures.

However, a more precise comparison can be made. Under appropriate uniform integrability hypotheses (this should be viewed as a mild additional mathematical assumption on  $Y$ ), one can pass expectations through (4.1):

$$n^{1/2} E L_n(g) \rightarrow (z(1-\alpha/2) - z(\alpha/2)) \cdot \sigma E g(B) \quad (4.2)$$

as  $n \rightarrow \infty$ . Clearly, one would like to select  $g \in \mathcal{M}$  so that the right-hand side of (4.2) is as small as possible. It turns out that

$$\inf\{z(1-\alpha/2) - z(\alpha/2) E g(B) : g \in \mathcal{M}\} = 2z_\delta \quad (4.3)$$

where  $z_\delta$  is the  $(1-\delta/2)\%$  quantile of a  $N(0,1)$  r.v. (see Section 5 of [5]). The right-hand side of (4.3) is the asymptotic confidence interval length obtained via a consistent estimation technique. We therefore have that:

Asymptotically, a standardized time series confidence interval is at least as long, in expected value, as a confidence interval obtained via consistent estimation.

In fact, it turns out that standardized time series confidence intervals are strictly longer than consistent estimation intervals, because the infimum in (4.3) is not attained within  $\mathcal{M}$ .

What about asymptotic variability of  $L_n(g)$ ? Note that

$$\text{var } g(\bar{Y}_n) = n^{-1} \text{var } g(X_n)$$

so that

$$n \text{var } g(\bar{Y}_n) \rightarrow \sigma^2 \text{var } g(B)$$

as  $n \rightarrow \infty$  (again, provided that appropriate uniform integrability for  $Y$  is available). It turns out that if  $g \in \mathcal{M}$ , then  $g(B)$  can not be deterministic (this justifies calling standardized time series a cancellation procedure; see Section 2), so  $\text{var } g(B) > 0$ . Thus, the variance of  $L_n(g)$  is of order  $n^{-1}$ .

On the other hand, if consistent estimation is used, we claim that the variance of  $L_n(g)$  is of smaller order than  $n^{-1}$ . The length of such a confidence interval is given by

$$L_n = 2 z_\delta s(n) / \sqrt{n} ;$$

so

$$\text{var } L_n = \frac{1}{n} 4 z_\delta^2 \text{var } s(n) .$$

But  $s(n) \Rightarrow \sigma$  as  $n \rightarrow \infty$ , so that (under appropriate uniform integrability),  $\text{var } s(n) \rightarrow \text{var } \sigma = 0$ , showing that  $\text{var } L_n$  is of smaller order than  $n^{-1}$ . In fact, in [5] it is shown that for the regenerative method,  $\text{var } L_n$  is of order  $n^{-2}$ . We conclude that:

Asymptotically, a standardized time series confidence interval is more variable than a confidence interval obtained via consistent estimation.

It is of interest to return to a philosophical point discussed in Section 2. It was mentioned there that confidence intervals are but one way to measure the variability of a point estimate. Another alternative is to just estimate the MSE of the point estimate, namely  $\sigma^2$ . Clearly, consistent estimation techniques give such an estimate, in addition to a confidence interval for  $r$ .

However, it does not appear that standardized time series gives consistent estimators for  $\sigma^2$ . For observe that

$$n^{1/2} g(\bar{Y}_n) \Rightarrow \sigma g(B)$$

Since  $g(B)$  is non-deterministic if  $g \in \mathcal{M}$ , it follows that  $n^{1/2} g(\bar{Y}_n)$  does not consistency estimate  $\sigma$ .

## 5. SUMMARY AND FUTURE WORK

Schruben's method of standardized time series for generating steady-state confidence intervals has provided a totally new approach for attacking one of the hardest problems in output analysis. It appears to behave well empirically, and seems to be a robust procedure. However, it suffers two defects from a mathematical viewpoint: longer confidence intervals, on average, and more variable confidence intervals, than consistent estimation methods. On the other hand, these results pertain only to large-sample asymptotics, and therefore it is possible that standardized time series procedures may strictly dominate consistent estimation methods in a small-sample context.

As for future work to be done in the area, the following problems seem interesting:

1. Mathematical analysis of small-sample theory for standardized time series (this would involve obtaining Berry-Esseen type results for FCLT's and would be hard).
2. Analyzing confidence interval methods in which one studies  $(\bar{Y}_n(1) - r)/g_n(\bar{Y}_n)$  (this would permit increasing the number of batches with the sample size).
3. Extending standardized time series to multivariate output.
4. Sequential stopping rules for standardized time series (classical work on sequential stopping rules requires consistent estimation of  $\sigma^2$ ).

## ACKNOWLEDGEMENTS

Peter W. Glynn was supported by Army Research Office Contracts DAAG29-80-C0041 and DAAG29-84-K-0030 and by National Science Foundation Grant ECS-8404809. Donald L. Iglehart was supported by Army Research Office Contract DAAG29-84-K-0030 and by National Science Foundation Grant MCS-8203483.

## REFERENCES

- [1] Billingsley, P., **Convergence of Probability Measures**, John Wiley, New York, 1968.
- [2] Breiman, L., **Probability**, Addison-Wesley, Reading, Mass., 1968.
- [3] Glynn, P.W., "Some New Results in Regenerative Process Theory," Technical Report No. 60, Dept. of Operations Research, Stanford University, Stanford, CA, 1982.
- [4] Glynn, P.W., "Limit Theorems for the Method of Replication," Technical Summary Report No. 23, Mathematics Research Center, Madison, WI, 1985.
- [5] Glynn, P.W. and D.L. Iglehart, "The Theory of Standardized Time Series," Technical Report No. 32, Dept. of Operations Research, Stanford University, Stanford, CA, 1985.
- [6] Hall, P. and C.C. Heyde, **Martingale Limit Theory and its Applications**, Academic Press, New York, 1980.
- [7] Iglehart, D.L., "Simulating Stable Stochastic Systems, V: Comparison of Ratio Estimators," **Nav. Res. Logistics Quart.**, **22**, 1975, 553-565.
- [8] Iglehart, D.L., "The Regenerative Method for Simulation Analysis," in **Current Trends in Programming Methodology -- Software Modeling**, (K.M. Chandy and R.T. Hey, editors), Prentice-Hall, Englewood Cliffs, N.J., 1978.
- [9] Schruben, L., "Detecting Initialization Bias in Simulation Output," **Opns. Res.**, **30**, 1982, 569-590.
- [10] Schruben, L., "Confidence Interval Estimation Using Standardized Time Series," **Opns. Res.**, **31**, 1983, 1090-1108.

## PETER W. GLYNN

Peter W. Glynn received his Ph.D. degree in Operations Research from Stanford University in 1982 and is currently an Assistant Professor of Industrial Engineering at the University of Wisconsin at Madison. His research interests include simulation, queueing theory, and dynamic programming. He is a member of the TIMS, ORSA, Sigma Xi, and the Statistical Society of Canada.

Peter W. Glynn  
Department of Industrial Engineering  
University of Wisconsin-Madison  
1513 University Avenue  
Madison, WI 53706  
608/263-6790

## DONALD L. IGLEHART

Dr. Iglehart has been a Professor of Operations Research at Stanford University since 1967. Prior to that he was an Assistant and Associate Professor in the Department of Operations Research at Cornell University. He was educated at Cornell University (Bachelor of Engineering Physics), Stanford University (M.S. and Ph.D. in Statistics), and Oxford University (Postdoctoral Research Fellowship). He has published papers in the areas of inventory theory, queueing theory, weak convergence of probability measures, and simulation output methodology. His current research interests center on developing new probabilistic and statistical methods for analyzing the output of computer simulations.

Donald L. Iglehart  
Department of Operations Research  
Stanford University  
Stanford, CA 94305  
415/497-4032