Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

RISK-EFFICIENT SEQUENTIAL SIMULATION ESTIMATORS

Raghu Pasupathy

Department of Statistics
Purdue University
West Lafayette, IN 47906, USA

Yingchieh Yeh

Institute of Industrial Management
National Central University
Taoyuan, TAIWAN

ABSTRACT

Using steady state mean estimation as the prototypical context, we present a decision-theoretic framework for sequentially estimating quantities associated with an observable discrete-time stochastic process. Our framework includes weights for estimator quality and a linear cost of sampling. We first show that the optimal time to stop sampling in the hypothetical case when the autocovariance function of the process is known is the square root of the relative cost and the area under the autocovariance function. This expression inspires a sequential procedure that uses a partially overlapping batch means estimator to "stand-in" for the area under the autocovariance function. The sequential procedure is asymptotically optimal in the sense that the ratio of its risk and that of the optimal risk in the hypothetical scenario approaches unity in a certain asymptotic regime. The nature of our analysis hints at a general optimality principle that may be more generally prevalent.

1 INTRODUCTION

Let $\{Y_j, j \in \mathbb{Z}\}$ be a real-valued discrete-time stochastic process, and $(\Omega, \mathscr{F}, \mathscr{F}_{k\geq 1}, \mathbb{P})$ an associated filtered probability space, representing some phenomenon of interest. Suppose for simplicity that we wish to estimate the "steady state" mean $\mu := \lim_{n\to\infty} \mathbb{E}[n^{-1}\sum_{j=1}^{n} Y_j]$, that we assume exists. Suppose also that our objective is to estimate μ sequentially, that is, as the data $Y_j, j = 1,2,\ldots$ reveal themselves one by one. Importantly, we wish to perform such estimation in a manner that might be considered economical in some reasonable sense.

To make the setting precise, suppose $T \in \mathbb{Z}$ is a stopping time with respect to the σ-algebra formed by the process $\{Y_j, j \in \mathbb{Z}\}$ and $\bar{Y}_T := T^{-1}\sum_{j=1}^{T} Y_j$ is the desired estimator of the steady-state mean μ. Thus, the stopping time T encodes the strategy used to stop the data collection process, so that the estimator $\bar{Y}_T$ of μ can be constructed and returned to a user. We will propose a stopping time T and evaluate the resulting estimator's quality by analyzing the behavior of the risk function

$$R_T(c) := A\,\mathbb{E}[(\bar{Y}_T - \mu)^2] + c\,\mathbb{E}[T], \tag{1}$$

where A and c are known constants representing the trade-off between the increase in quality and the increase in cost associated with each additional observation. In particular, we will be concerned with the behavior of $R_T(c)$ in relation to the best achievable risk under an idealized scenario that we outline in Section 2, especially as the constant c becomes small compared to the constant A.

Strictly, we can dispense with the constant A in (1) and let c connote the *relative weight* of sampling effort to estimator quality; we have chosen to retain A to conform to the original paper by Ghosh and Mukhopadhyay (1979). Also, choices other than (1) and summary measures other than the steady state mean μ might be more reasonable depending on the context. However, as we shall see, much of the analysis in this paper can be greatly generalized — our choice of the structure of the risk function and the summary measure is made largely to ensure that the burden of mathematical exposition does not impede the transfer of insight to the reader.

We emphasize that, unlike in Ghosh and Mukhopadhyay (1979), we do not assume that the observations $Y_j, j = 1,2,\ldots$ are independent and identically distributed, but that $\{Y_j, j \in \mathbb{Z}\}$ satisfies a certain strong approximation property:

$$S_n - n\mu = \sigma B(n) + O(n^{1/2-\lambda}) \text{ a.s.}; \quad S_n := \sum_{j=1}^{n} Y_j, \tag{SA}$$

where μ is the steady state mean introduced earlier, $\sigma \in \mathbb{R}^+$ is variously called the *time-averaged variance constant* (TAVC) or simply the *variance parameter* (Aktaran-Kalaycı, Alexopoulos, Goldsman, and Wilson 2009), $\lambda \in (0,1/2]$ is a known constant, $\{B(t), t \in [0,\infty)\}$ is standard Brownian motion (Billingsley 1995), and S_n is the partial sum formed by observations from the process $\{Y_j, j \in \mathbb{Z}\}$. The assumption in (SA), while appearing stringent, is widely satisfied — see, for example, Philipp and Stout (1975) and Glynn and Iglehart (1988).

To clarify what (SA) says about $\{Y_j, j \in \mathbb{Z}\}$ in terms of satisfying a functional central limit theorem, and second-order stationarity, we state the following result.

Theorem 1 Let the process $\{Y_j, j \in \mathbb{Z}\}$ satisfy (SA). Also, define the interpolated process $\{S_n(t), t \in [0,1]\}$ as follows:

$$S_n(t) := \begin{cases} 0 & t = 0; \\ \sum_{j=1}^{nt} Y_j, & nt = 1,2,\ldots,n; \\ \sum_{j=1}^{\lfloor nt \rfloor} Y_j + (t - \frac{\lfloor nt \rfloor}{n}) Y_{\lfloor nt \rfloor + 1}, & \text{otherwise.} \end{cases}$$

Then, letting $\{B(t), t \in [0,1]\}$ denote the standard Brownian motion (Billingsley 1995), the following assertions hold.

(i) $S_n(t) - \mu nt \xrightarrow{d} \sigma\sqrt{n}B(t), t \in [0,1]$, as $n \to \infty$, where $\xrightarrow{d}$ here refers to weak convergence on the space $D[0,1]$ of right continuous functions having left limits;

(ii) if the sequence of sample means $\{n^{-1}S_n, n \geq 1\}$ is uniformly integrable, that is,

$$\lim_{\alpha \to \infty} \sup_n \mathbb{E}\left[|n^{-1}S_n| \mathbb{I}\left\{|n^{-1}S_n| > \alpha\right\}\right] = 0,$$

then $\mathbb{E}[n^{-1}S_n] \to \mu$ as $n \to \infty$; and

(iii) if the sequence $\{n(n^{-1}S_n - \mu)^2, n \geq 1\}$ is uniformly integrable, then $n\,\text{Var}(n^{-1}S_n) \to \sigma^2$ as $n \to \infty$.

The assertion in (i) follows from Donsker's theorem (Billingsley 1999, pp. 90). Furthermore, with some calculation and assuming the process $\{Y_j, j \in \mathbb{Z}\}$ is second-order stationary (Hoel, Port, and Stone 1972), it can be shown that σ^2 appearing in (SA) satisfies

$$\sigma^2 = \sum_{h=-\infty}^{\infty} \gamma(h), \quad \gamma(h) := \text{cov}(Y_j, Y_{j+h}).$$

We do not prove Theorem 1 but note that an almost identical result appears in Damerdji (1994) without a proof. Also, to ease exposition of all results that follow, we make the following two further assumptions about the process $\{Y_j, j \in \mathbb{Z}\}$.

A.1 The process $\{Y_j, j \in \mathbb{Z}\}$ is second-order stationary, implying that for any $\tau \in \mathbb{Z}$, the process defined by $X_j^\tau = Y_{j+\tau}$ and the process $\{Y_j, j \in \mathbb{Z}\}$ have the same mean and autocovariance functions:

$$\mathbb{E}[X_j^\tau] = \mathbb{E}[Y_j] =: \mu; \quad \text{cov}(X_j^\tau, X_{j+h}^\tau) = \text{cov}(Y_j, Y_{j+h}) =: \gamma(h), \quad h = 0, \pm 1, \pm 2, \ldots$$

(It is seen by selecting $\tau = -h$ that the function $\gamma(\cdot)$ is symmetric about zero, that is, $\gamma(h) = \gamma(-h)$.)

A.2 The autocovariance function $\gamma(h), h=0,1,2,\ldots$ is such that

$$\sum_{h=0}^{\infty} |h|\,\gamma(h) < \infty.$$

2 A HYPOTHETICAL PROCEDURE

We now consider a hypothetical scenario where the autocovariance function

$$\gamma(h) := \mathrm{cov}(Y_j, Y_{j+h}), \quad h = 0, \pm 1, \pm 2, \ldots$$

is known, implying that the quantities

$$\sigma^2 := \sum_{h=-\infty}^{\infty} \gamma(h); \quad \sigma_1^2 := \sum_{h=-\infty}^{\infty} |h|\,\gamma(h)$$

are known. (σ^2 and σ_1^2 exist by the assumptions in A.1 and A.2.) Let's now consider only "naive procedures," that is, procedures $T = n \in \{1,2,\ldots,\}$ that choose to stop after a pre-specified amount of sampling n. Among such procedures, the optimal procedure can be identified. We notice that

$$\begin{aligned} R_n(c) &= A\,\mathrm{var}(\bar{Y}_n) + cn \\ &= An^{-1} \sum_{h=-n}^{n} \left(1 - \frac{|h|}{n}\right)\gamma(h) + cn, \end{aligned}$$

and treating n as a continuous variable,

$$\begin{aligned} \frac{dR_n(c)}{dn} &= -A\left(n^{-2}\sum_{h=-n}^{n}\left(1-\frac{|h|}{n}\right)\gamma(h) - n^{-3}\sum_{h=-n}^{n}|h|\gamma(h)\right) + c \\ &= -An^{-2}\sum_{h=-n}^{n}\left(1-2\frac{|h|}{n}\right)\gamma(h) + c. \end{aligned}$$

Hence the optimal sample size n^*, ignoring non-integrality, solves the equation

$$-An^{-2}\sum_{h=-n}^{n}\left(1-2\frac{|h|}{n}\right)\gamma(h) + c = 0.$$

Notice, in particular, that

$$n^{*2} \sim \frac{A}{c}\sigma^2 + 2\sqrt{\frac{A}{c}}\,\frac{\sigma_1^2}{\sigma} \quad \text{as } c \to 0, \tag{2}$$

where the notation $a_n \sim b_n$ mean $a_n/b_n \to 1$ as $n \to \infty$ for non-negative real-valued sequences $\{a_n, n \in \mathbb{Z}\}$ and $\{b_n, n \in \mathbb{Z}\}$. Notice also that for independent and identically distributed data,

$$n^* = \sqrt{\frac{A}{c}}\,\gamma(0)$$

where $\gamma(0)$ is the asymptotic marginal variance. Furthermore, the incurred risk at the optimal sample size given by (2) satisfies

$$R_{n^*}(c) \sim 2cn^*. \tag{3}$$

The expression for the asymptotic optimal sample size in (2) suggests that the principal factor affecting optimal stopping is the variance parameter $\sigma^2 := \sum_{h=-\infty}^{\infty}\gamma(h)$, and to a lesser extent the "first moment" σ_1 associated with the autocovariance function. The procedure we will outline next reflects this insight.

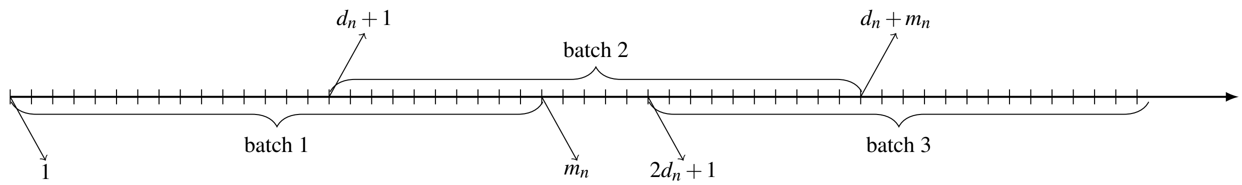

Figure 1: The figure depicts partially overlapping batches. Batch 1 consists of observations $\xi_j, j = 1,2,\ldots,m_n$; batch 2 consists of observations $\xi_j, j = d_n+1, d_n+2,\ldots,d_n+m_n$, and so on, with batch i consisting $\xi_j, j = (i-1)d_n+1,(i-1)d_n+2,\ldots,(i-1)d_n+m_n$. There are $k_n := d_n^{-1}(n-m_n)$ batches in total, where n is the size of the dataset.

3 PROPOSED SEQUENTIAL ESTIMATOR

The hypothetical optimal procedure in the previous section suggests a procedure that might be effective in a real context where the autocovariance function $\gamma(h), h = 0,\pm 1,\pm 2,\ldots,$ is not known. Specifically, when the *relative cost* A/c is large, the hypothetical procedure suggests that the optimal sample size is simply the square root of the relative cost multiplied by the sum of autocovariances. Since the sum of autocovariances σ^2 is unknown, one might expect that a successful sequential procedure might continuously monitor an estimate $\hat{\sigma}_T^2$ of $\sigma^2 := \sum_{h=-\infty}^{\infty}\gamma(h)$, constructed perhaps using a method such as *fully overlapping batch means* (Aktaran-Kalaycı, Goldsman, and Wilson 2007), and then stopping when the collected amount of data exceeds the product of the square root of A/c and $\hat{\sigma}_T^2$, that is,

$$T := \min\{n \geq \sqrt{\frac{A}{c}}\left(\hat{\sigma}_n + n^{-\beta}\right)\}, \tag{4}$$

where $\hat{\sigma}_n^2$ is an estimate of $\sigma^2 = \sum_{h=-\infty}^{\infty}\gamma(h)$, the exact nature of which we will detail next, and $\beta > 0$ is some positive constant. The expression in (4) thus implicitly mimics (2) after slightly inflating the estimate of σ^2 to account for the “early stopping” effect (Chow and Robbins 1965; Glynn and Whitt 1992) that invariably accompanies sequential stopping procedures.

3.1 Estimating TAVC

The proposed procedure estimates the steady-state mean μ as $\bar{Y}_T := T^{-1}\sum_{j=1}^{T} Y_j$, where T is specified through (4). In this section, we briefly outline the *partially overlapping batch means* procedure to estimate the TAVC σ^2. Overlapping batch means as a method to estimate TAVC is a well studied area — see especially Aktaran-Kalaycı et al. (2009) for an overview.

Define the batch means $\bar{Y}_j, j = 1,2,\ldots,k_n$, depicted in Figure 1, as follows.

$$\bar{Y}_i := \frac{1}{m_n}\sum_{j=(i-1)d_n+1}^{(i-1)d_n+m_n} Y_j, \quad i = 1,2,\ldots,k_n := 1+\frac{n-m_n}{d_n}; \quad 1 \leq m_n < n; \quad 1 \leq d_n \leq m_n. \tag{5}$$

The i-th batch mean $\bar{Y}_i$ is thus the sample mean of $m_n < n$ consecutive observations, and the last $m_n - d_n$ observations of each batch overlaps with the previous batch. When $d_n = 1$ we say the batches are *fully overlapping* and when $d_n = m_n$ we say that the batches are *non-overlapping*. For convenience, we ignore the possible non-integrality of the number of batches $k_n := 1+(n-m_n)/d_n$. Also, we use the following notation for convenience:

$$\varepsilon_n := \frac{m_n}{n}; \quad \delta_n := 1-\frac{d_n}{m_n}.$$

In practice, m_n and d_n are conveniently chosen; for instance, when the batch size is chosen as a fixed fraction ν of the total data size and we use fully overlapping batches, we get $\varepsilon_n = \nu n$ and $\delta_n = 1-\varepsilon_n$. Likewise, when the batch size is chosen as a fixed fraction ν of the total data size and we use non-overlapping batches, we get $\varepsilon_n = \nu$ and $\delta_n = 0$.

Given (5), we estimate the variance parameter as the sample variance of the batch means:

$$\hat{\sigma}_n^2(\beta_1,\beta_2) := \frac{1}{k_n}\sum_{j=1}^{k_n}(\bar{Y}_j - \bar{Y})^2, \quad \bar{Y} := k_n^{-1}\sum_{j=1}^{k_n}\bar{Y}_j = n^{-1}\sum_{i=1}^{n}Y_i + E_n. \tag{6}$$

The estimator $\bar{Y}$ of the steady-state mean μ appearing in (6) is the average of the individual batch means $\bar{Y}_j, j = 1,2,\ldots,k_n$ and is sometimes called the *batching* estimator. As implied in (6), the batching estimator of the steady-state mean μ deviates from the *sectioning* estimator $n^{-1}\sum_{i=1}^{n} Y_i$ by some "end effect" E_n. It is easy to see that when batches are non-overlapping, that is, when $\delta_n = 0$, the end effect $E_n = 0$ and the batching, sectioning estimators coincide.

3.2 Asymptotic Properties of the TAVC Estimator

It seems intuitively clear that the characteristics of the stopping time T and the quality of the estimator $\bar{Y}_T$ will be heavily dictated by the properties of the estimator in (6). Since the strong approximation assumption (SA) holds, it seems reasonable that $\hat{\sigma}_n^2$ should converge in some precise sense to the corresponding quantities formed by the Brownian motion; following the notation of Damerdji (1994), we thus let

$$\begin{aligned}\bar{A}_j &= m_n^{-1}\sum_{i=1}^{m_n} B((j-1)d_n + i) - B((j-1)d_n + (i-1))\\ &= m_n^{-1}B((j-1)d_n + m_n) - B((j-1)d_n), \quad j = 1,2,\ldots,k_n.\end{aligned}$$

and

$$\bar{A} = n^{-1}(B(n) - B(0)) = n^{-1}B(n).$$

to "stand-in" for $\bar{Y}_j$ and $\bar{Y}$, respectively. We then have the following result that characterizes the behavior of the estimator $\hat{\sigma}_n^2$ under certain conditions on the batch size sequence and the extent of the overlapping in the batches.

Theorem 2 Suppose the strong approximation assumption (SA) and Assumptions A.1–A.2 hold.

1. Let non-overlapping batches be used, that is, $\delta_n = 0$. Let the batch sizes $\{m_n, n \geq 1\}$ be chosen such that for some $a \in (0,\infty)$,
$$\sum_{n=1}^{\infty}\varepsilon_n^a < \infty; \quad \varepsilon_n := \frac{m_n}{n}. \tag{7}$$
Then
$$\hat{\sigma}_n^2 \xrightarrow{\text{a.s.}} \sigma^2.$$
Furthermore, for any $t \in (0,\infty)$ and $\nu \in (0,\infty)$,
$$P\left(\left|\frac{\hat{\sigma}_n^2}{\sigma^2} - 1\right| > t\right) \leq \frac{C(\nu)}{t^\nu}\,\varepsilon_n^\nu,$$
where $C(\nu)$ is uniformly bounded.
2. Let fully-overlapping batches be used, that is, $\delta_n = 1 - \frac{1}{n}$, and suppose $m_n \to \infty$ as $n \to \infty$. Then
$$\hat{\sigma}_n^2 \xrightarrow{d} \sigma^2\Lambda(\beta_1),$$
where
$$\Lambda(\beta_1) := \frac{1}{\beta_1}\int_0^{1-\beta_1}\left(B(t+\beta_1) - B(t) - \beta_1 B(1)\right)^2 dt.$$

Theorem 2 characterizes the asymptotic behavior of TAVC for the two extreme cases of non-overlapping and fully-overlapping batch means. (We will not explain the cryptic notation $\Lambda(\beta_1, 1)$ here but only state that $\Lambda(\beta_1, 1)$ is a special case of a two-parameter random variable that results from more general batching procedures.) Notably, we see that strong consistency is ensured only if the condition (7) is satisfied, that is, when the batches are non-overlapping, the batch sizes go to infinity but not too fast. As we shall see, risk optimality depends crucially on the TAVC estimator exhibiting strongly consistent behavior, and that too with the tail risk decaying sufficiently fast.

We do not provide a proof of Theorem 2 but instead refer the reader to (Glynn and Whitt 1991; Damerdji 1994; Aktaran-Kalaycı et al. 2009) for clear indications on how to obtain the result.

4 OPTIMALITY OF THE PROPOSED SEQUENTIAL ESTIMATOR

We are now ready to fully characterize the quality of the proposed sequential estimator. Crucial to such characterization is the nature of the stopping time (4), specially on how its moments behave. If the moments of T become large because of excessive early stopping behavior, this is bound to affect the risk of the resulting sequential estimator, potentially making it sub-optimal. The following theorem, analogous to the lemma appearing on page 640 of Ghosh and Mukhopadhyay (1979), fully characterizes the behavior of the stopping time T.

Theorem 3 Suppose the strong approximation assumption (SA) holds. Let non-overlapping batches be used in constructing $\hat{\sigma}_n^2$. Let the batch sizes $\{m_n, n \geq 1\}$ be chosen such that for some $\kappa > 0$, the following condition holds.

$$\lim_{n\to\infty} n^{\kappa}\varepsilon_n = 0. \tag{8}$$

Then, the following assertions hold.

1. $P(T < \infty) = 1$;
2. T is $\downarrow$ in c; $\lim_{c\to 0} T = \infty$ a.s.;
3. $\lim_{c\to 0} T/n^* = 1$ a.s.;
4. $\lim_{c\to 0} \mathbb{E}[(T/n^*)^m] = 1$ for any $m > 0$;

Proof Sketch. The proof of the first part follows in a straightforward manner from the expression in (4) and since $\hat{\sigma}_n^2 \to \sigma^2$ a.s. as $n \to \infty$ (Grams and Serfling 1973). The second part of the theorem also follows in a straightforward manner from the expression in (4). The third part of the theorem follows upon noticing that

$$\sqrt{\frac{A}{c}}\,\hat{\sigma}_T \leq T \leq \sqrt{\frac{A}{c}}\left(\hat{\sigma}_{T-1} + (T-1)^{-\beta}\right).$$

A proof of the last part of the theorem is nuanced and relies heavily on the tail-probability bound that appears in the first part of Theorem 2. □

We finally conclude with the main theorem of the paper which asserts that the risk associated with the proposed estimator is asymptotically optimal in the sense that the ratio of its risk to that of the optimal risk in the hypothetical scenario described in Section 2 tends to unity as c becomes small.

Theorem 4 If $\sigma^2 > 0$ and $0 < \beta < 1/4$, then

$$\lim_{c\to 0} \frac{R_{n^*}(c)}{R_T(c)} = 1,$$

where the expression for $R_{n^*}(c)$ appears in (3).

Proof Sketch. A proof of this theorem relies heavily on Theorem 3 and starts with the arguments of the proof of Theorem 1 in Ghosh and Mukhopadhyay (1979). □

5 SUMMARY AND CONCLUDING REMARKS

We have outlined a simple sequential procedure to estimate the steady-state mean of a discrete time stochastic process. The procedure mimics the optimal stopping procedure for a hypothetical context where the time-averaged variance constant of the stochastic process is known, and estimates the time-average variance constant using an appropriate batching procedure. Interestingly, such sequential estimation of the steady-state mean turns out to be risk optimal when the relative cost of sampling is small and the batches used in the batching procedure satisfy a certain condition that stipulates that they are neither too large nor too small.

It should be clear that while we have treated only the steady-state mean context here, the analogous procedure for other quantities, e.g., steady-state quantile, follows in a very direct fashion — follow the same stopping procedure after substituting the batch means $\bar{Y}_j$ with the batch estimator of the quantity of interest. In contexts involving quantities other than the steady-state mean, different "centering" choices become available — see Dong and Nakayama (2020), Asmussen and Glynn (2007). Also, how to correctly "center" the estimator for the time-averaged variance constant in such contexts turns out to be non-trivial. Nevertheless, our outlined procedure seems to lay out a general blueprint for constructing risk-optimal estimators in sequential simulation contexts.

REFERENCES

Aktaran-Kalaycı, T., C. Alexopoulos, D. Goldsman, and J. R. Wilson. 2009. "Optimal Linear Combinations of Overlapping Variance Estimators for Steady-State Simulation". In *Advancing the Frontiers of Simulation*, edited by C. Alexopoulos, D. Goldsman, and J. R. Wilson. NY: Springer.

Aktaran-Kalaycı, T., D. Goldsman, and J. R. Wilson. 2007. "Linear Combinations of Overlapping Variance Estimators for Simulation". *Operations Research Letters* 35(4):439–447.

Asmussen, S., and P. W. Glynn. 2007. *Stochastic Simulation: Algorithms and Analysis*. New York, NY: Springer.

Billingsley, P. 1995. *Probability and Measure*. New York, NY: Wiley.

Billingsley, P. 1999. *Convergence of Probability Measures*. New York, NY: Wiley.

Chow, Y. S., and H. E. Robbins. 1965. "On the Asymptotic Theory of Fixed-Width Confidence Intervals for the Mean". *Annals of Mathematical Statistics* 36:457–462.

Damerdji, H. 1994. "Strong Consistency of the Variance Estimator in Steady-State Simulation Output Analysis". *Mathematics of Operations Research* 19(2):494–512.

Dong, H., and M. Nakayama. 2020. "A Tutorial on Quantile Estimation via Monte Carlo". In *Monte Carlo and Quasi-Monte Carlo Methods*, edited by B. Tuffin and P. L'Ecuyer, 3–30: Springer International Publishing.

Ghosh, M., and N. Mukhopadhyay. 1979. "Sequential Point Estimation of the Mean when the Distribution Is Unspecified". *Communications in Statistics - Theory and Methods* 8:637–652.

Glynn, P. W., and D. L. Iglehart. 1988. "A New Class of Strongly Consistent Variance Estimators Simulations". *Stochastic Processes and Their Applications* 28:71–80.

Glynn, P. W., and W. Whitt. 1991. "Estimating the Asymptotic Variance with Batch Means". *Operations Research Letters* 10:431–435.

Glynn, P. W., and W. Whitt. 1992. "The Asymptotic Validity of Sequential Stopping Rules for Stochastic Simulations". *The Annals of Applied Probability* 2(1):180–197.

Grams, W. F., and R. Serfling. 1973. "Convergence Rates for U-Statistics and Related Statistics". *The Annals of Statistics* 1(1):153–160.

Hoel, P. G., S. C. Port, and C. J. Stone. 1972. *Introduction to Stochastic Processes*. New York, New York: Houghton Mifflin Company.

Philipp, W., and W. Stout. 1975. *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables*. American Mathematical Society.

AUTHOR BIOGRAPHIES

RAGHU PASUPATHY is an associate professor in the Department of Statistics at Purdue University. His research interests lie broadly in Monte Carlo methods with a specific focus on simulation optimization. He is a member of INFORMS, IIE, and ASA, and serves as an associate editor for Operations Research and INFORMS Journal on Computing. He is the Area Editor for the Simulation Desk at IIE Transactions. His email address is pasupath@purdue.edu and his web page is http://web.ics.purdue.edu/ pasupath.

YINGCHIEH YEH is an associate professor in Institute of Industrial Management at National Central University, Taiwan. He received his Ph.D. from the School of Industrial Engineering at Purdue University. His primary research interests include simulation output analysis, applied probability and statistics, and applied operations research. His email address is yeh@mgt.ncu.edu.tw.