UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

# RANDOMIZED ESTIMATORS FOR TIME INTEGRALS

Peter W. Glynn

Technical Summary Report #2603

November 1983

ABSTRACT

Let $\{X(t) : t \geq 0\}$ be a real-valued stochastic process and set $\alpha = \int_0^\infty X(t)\, G(dt)$, where $G$ is a (non-random) distribution function. If the support of $G$ is large, standard Monte Carlo techniques for estimating $\alpha$ are inefficient, since $X$ must be simulated over the entire support of $G$. To avoid this difficulty, randomization schemes are derived that require simulation of $X$ over random subsets of the support of $G$. Large-sample behavior of randomized estimators is studied in detail. Some variance reduction techniques are also presented.

AMS (MOS) Subject Classifications: 60G57, 68J05

Key Words: Simulation, variance reduction, random measures

Work Unit Number 5 - Optimization and Large Scale Systems

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

## SIGNIFICANCE AND EXPLANATION

Consider a stochastic system for which one needs an estimate of the expected discounted cost over the infinite horizon. Standard Monte Carlo procedures do not apply since the parameter to be estimated involves values of the process over an infinitely long time interval. In this paper, we present Monte Carlo estimation techniques, based on randomization, that can be used in the above setting. The techniques developed turn out to be more efficient than the standard approach, even when the parameter to be estimated cumulates costs over a finite time interval.

---

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

RANDOMIZED ESTIMATORS FOR TIME INTEGRALS

Peter W. Glynn

## 1. Introduction

Let $\{X(t) : t > 0\}$ be a real-valued stochastic process representing the output of a simulation, and consider a time integral of the form

(1.1) $$I = \int_0^\infty X(t)\, G(dt) \quad ,$$

where $G$ is a (deterministic) non-decreasing function. Our objective, in this paper, is to study Monte Carlo techniques for estimating the parameter $\alpha = EI$.

The time integral (1.1) includes several commonly studied performance criteria.

(1.2). Example. If $G(t) = 1 - e^{-\alpha t} (\alpha > 0)$, then $I$ corresponds to discounting $X(t)$ at rate $\alpha$ over an infinite horizon. Such time integrals occur frequently in inventory models.

(1.3). Example. If $G(t) = \min(t/T, 1)$, then $I$ is the average of $X(t)$ over the interval $[0,T]$. Such averages are often of interest, in a queueing context.

(1.4). Example. Let $\{Y(t) : t > 0\}$ be a stochastic process and $f(t,y)$ be a real-valued performance measure which may depend explicitly on $t$. By setting $X(t) = f(t,Y(t))$, criteria of the form

$$I = \int_0^\infty f(t,Y(t))\, G(dt)$$

can be incorporated as a special case of (1.1).

(1.5). Example. Grassman (1982) has recently developed a Monte Carlo technique for estimating $Ef(Z(t))$, where $Z(t)$ is a uniformizable Markov jump process. The idea is to represent $Z(t)$ as $Y(N(t))$, where $N(t)$ is a

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

Poisson process with rate $\lambda$ (say), and $\{Y(k) : k > 0\}$ is an independent discrete-time Markov chain. Thus,

$$\text{(1.6)} \qquad Ef(Z(t)) = \sum_{k=0}^{\infty} Ef(Y(k))e^{-\lambda t}(\lambda t)^k/k! \quad .$$

Setting $X(t) = f(Y([t]))$ ($[t]$ = greatest integer less than or equal to $t$), and $G(dt)$ = Poisson measure, we see that the representation (1.6) is a special case of (1.1).

In Section 2, we shall briefly discuss the direct method for estimating $\alpha$; this involves generation of independent variates $I_1, I_2, \ldots,$ each having the distribution of $I$. The parameter $\alpha$ is then estimated by $\bar{I}(n)$, where

$$\bar{I}(n) = \frac{1}{n} \sum_{k=1}^{n} I_k \quad .$$

The difficulty with the direct method is that if $T = \sup\{t : G(t) < G(\infty)\}$ ($G(\infty) = \lim_{t\to\infty} G(t)$) is large, then generation of variates is expensive. Thus, in Section 3, a general framework for randomized estimation of $\alpha = EI$ is presented.

(1.7). <u>Definition</u>. $\bar{R}(n)$ is a randomized $\bar{I}(n)$ - estimator if there exists a $\sigma$-field $\mathcal{G}$ such that

$$E\{\bar{R}(n) \mid \mathcal{G}\} = \bar{I}(n) \quad .$$

The definition of conditional expectation implies that $E\bar{R}(n) = E\bar{I}(n) = \alpha$, justifying the description of $\bar{R}(n)$ as an $\bar{I}(n)$ - estimator. Some authors refer to methods based on randomized estimation as conditional Monte Carlo procedures (see, for example, Rubinstein (1981), p. 141). However, we prefer to reserve the term "conditional Monte Carlo" for the "converse" to a randomized estimator.

(1.8). <u>Definition</u>. $\bar{I}(n)$ is a conditional Monte Carlo $\bar{R}(n)$ - estimator if and only if $\bar{R}(n)$ is a randomized $\bar{I}(n)$ - estimator.

The definition (1.8) of the term "conditional Monte Carlo" is consistent with its usage in several recent books; see, for example, Bratley, Fox, and Schrage (1983) or Law and Kelton (1982). The following proposition is a well-known property of conditional expectation (Burrill (1972), p. 392).

(1.9). <u>Proposition</u>. Suppose that $E|\overline{R}(n)| < \infty$ and that $\overline{R}(n)$ is a randomized $\overline{I}(n)$ - estimator. Then,

$$\text{var}(\overline{R}(n)) \geq \text{var}(\overline{I}(n)) \quad .$$

Proposition 1.9 states that a randomized $\overline{I}(n)$ - estimator has larger mean square error (MSE) than $\overline{I}(n)$. For this reason, the conditional Monte Carlo estimator $\overline{I}(n)$ is to be preferred in the case that generating $\overline{R}(n)$ requires the same effort as simulating $\overline{I}(n)$. However, it turns out that in our time integral setting, the time required to generate a randomized $\overline{I}(n)$ - estimator will often be smaller than that required to simulate $\overline{I}(n)$; this property can offset the MSE advantage of $\overline{I}(n)$. This is the theme of Sections 3 through 6. After developing a general framework for randomized estimation in Section 3, three specific randomized estimation algorithms are studied in Section 4 through 6.

## 2. The Direct Method

Development of estimation theory for the parameter $\alpha$ requires some assumptions on the simulation. Let $(\Omega, F, P)$ be the probability space which supports our simulation. We assume that:

A1. there exists a family of processes $\{(X_k, S_k) : k \geq 1\}$ such that $X_k : \mathbf{R}^+ \times \Omega \to \mathbf{R}$ and $S_k : \mathbf{R}^+ \times \Omega \to \mathbf{R}^+$, where $\mathbf{R}^+ = [0,\infty)$

A2. the processes $X_k(\cdot,\omega)$ possess left limits and are right continuous, for each $\omega \in \Omega$

A3. the processes $S_k(\cdot,\omega)$ are non-decreasing and right continuous for each $\omega \in \Omega$; also $P\{S_k(t) > 0\} > 0$ for $t > 0$

A4. $\{(X_k, S_k) : k \geq 1\}$ is a sequence of independent and identically distributed (i.i.d.) random elements

A5. $G : \mathbf{R} \to [0,1]$ is a non-decreasing right continuous function such that $G(0-) = 0$, $G(0) < 1$, and $G(\infty) = 1$ $(G(x-) = \sup\{G(t) : t < x\})$.

A6. $\int_0^\infty E|X_k(t)|\ G(dt) < \infty$ .

Assumption A2 guarantees that $X_k$ is product measurable (see Dellacherie and Meyer (1978), p. 89). Fubini's theorem applied to A6 therefore asserts that if

$$I_k(\omega) \triangleq \begin{cases} \int_0^\infty X_k(t,\omega)G(dt); & \int_0^\infty |X_k(t,\omega)|\ G(dt) < \infty \\ 0 \ ; & \text{else} \end{cases},$$

then $I_k$ is $\mathcal{F}$-measurable (i.e. a random variable) and satisfies

$$I_k = \int_0^\infty X_k(t)\ G(dt) \quad \text{a.s.}$$

The goal is to estimate $\alpha \triangleq EI_k$. The process $S_k(t)$ will be interpreted as the amount of "effort" required to simulate $X_k(\cdot)$ up to time $t$.

(2.1). Example. Suppose that $X_k(t) = t + Y_k(N_k(t))$, where $\{Y_k(j) : j \geq 0\}$ is a Markov chain and $N_k(t)$ is a Poisson process. Assuming that the simulation effort is measured by the number of random variables (r.v's) generated, $S_k(t) = 2N_k(t) + 2$ ($N_k(t) + 1$ for $Y_k(0),\ldots,Y_k(N_k(t))$ and $N_k(t) + 1$ for the exponential variates).

Let $\bar{I}(n) = \sum_{k=1}^{n} I_k/n$ with $\bar{I}(0) = 0$; $\bar{I}(n)$ will be referred to as the direct estimator of $\alpha$. Set $T = \sup\{t : G(t) < 1\}$ and put $N(t) = \max\{k : S_1(T) +\ldots+ S_k(T) \leq t\}$; $N(t)$ is the number of $I_k$'s generated by $t$ units of effort. Then, $I(t) \triangleq \bar{I}(N(t))$ is an estimator for $\alpha$, which can be constructed from $t$ units of effort.

(2.2). <u>Theorem</u>. Assume A1 - A6. Then, $I(t) \to \alpha$ a.s. as $t \to \infty$ if $P\{S_k(T) < \infty\} = 1$.

<u>Proof</u>. The strong law of large numbers guarantees that $\bar{I}(n) \to \alpha$ a.s. By A3 and A5, $P\{S_k(T) > 0\} > 0$ so $N(t) < \infty$ a.s. for all $t$. Furthermore, the assumption $P\{S_k(T) = \infty\} = 0$ assures that $N(t) \to \infty$ a.s. (see Çinlar (1975), p. 290), from which the result follows. ||

It is worth observing that if $P\{S_k(T) = \infty\} > 0$, then $I(t)$ does not, in general, converge to $\alpha$ (see Example 1.2). We will also be interested in rates of convergence for our estimators. For such results, we require a further moment assumption.

A7. $\sigma^2 = E(I_k - \alpha)^2 < \infty$ .

(2.3). <u>Theorem</u>. Assume A1 - A7. If $P\{S_k(T) < \infty\} = 1$, then

(2.4) $$\overline{\lim_{t\to\infty}}\, a(t) \cdot |I(t) - \alpha| = \sigma(ES_k(T))^{1/2} \text{ a.s.}$$

where $a(t) \triangleq (t/2 \log\log t)^{1/2}$ and $0.\infty \triangleq 0$.

<u>Proof</u>. The Hartman-Wintner form of the law of the iterated logarithm implies that

(2.5) $$\overline{\lim_{n\to\infty}}\, a(n)|\bar{I}(n) - \alpha| = \sigma \text{ a.s.}$$

Since $N(t) \to \infty$ a.s., (2.5) yields

(2.6) $$\overline{\lim_{t\to\infty}}\, a(N(t))|I(t) - \alpha| = \sigma \text{ a.s.}$$

But $N(t)/t \to 1/ES_k(T)$ a.s. (see [6], p. 290), which in turn implies that

(2.7) $$a(N(t))/a(t) \to 1/(ES_k(T))^{1/2} \text{ a.s.}$$

Relation (2.4) follows immediately from (2.6) and (2.7). ||

Confidence intervals for $\alpha$ can also be constructed from the direct estimator $I(t)$. The key tool is a central limit theorem (CLT), which is valid under slightly stronger assumptions than Theorem 2.3.

(2.8). <u>Theorem</u>. Assume A1 - A7. If $ES_k(T) < \infty$, then

(2.9) $$t^{1/2}(I(t) - \alpha) \Longrightarrow \sigma(ES_k(T))^{1/2} N(0,1)$$

as $t \to \infty$, where $N(0,1)$ is a mean zero unit variance r.v. and $\Longrightarrow$ denotes weak convergence.

Proof. From the CLT for i.i.d.r.v.'s,

$$n^{1/2}(\bar{I}(n) - \alpha) \Longrightarrow \sigma N(0,1)$$

as $n \to \infty$. Let $t_k$ be an arbitrary sequence converging to infinity. Since $N(t_k)/t_k \to 1/ES_k(T)$ a.s., one can apply Theorem 7.3.2 of Chung (1974) to conclude that

$$N(t_k)^{1/2}(I(t_k) - \alpha) \Longrightarrow \sigma N(0,1)$$

as $k \to \infty$. The converging-together lemma (Billingsley (1968), p. 25) then yields

$$t_k^{1/2}(I(t_k) - \alpha) \Longrightarrow \sigma(ES_k(T))^{1/2} N(0,1)$$

as $k \to \infty$. Since $\{t_k\}$ was arbitrary, we obtain (2.9) ([2], p. 16). ||

If $z_\delta$ solves $P\{N(0,1) < z_\delta\} = 1 - \delta/2$, Theorem 2.8 proves that the random interval

$$[I(t) - z_\delta v^{1/2}(t)/t^{1/2},\ I(t) + z_\delta v^{1/2}(t)/t^{1/2}]$$

is an approximate $100(1-\delta)\%$ confidence for $\alpha$, where $v(t)$ is a consistent estimator for $\sigma^2 ES_k(T)$.

Before concluding this section, we consider a special case of Example 1.4. Suppose that the process $Y(t)$ considered there is a stationary process on state space $S$. Then,

(2.10) $$\alpha = EI = \int_0^\infty Ef(t,Y(t))\, G(dt)$$

$$= \int_0^\infty \int_S f(t,y) P\{Y(t) \in dy\}\, G(dt)$$

$$= \int_S \int_0^\infty f(t,y)\, G(dt)\pi(dy)$$

where $\pi(\cdot) = P\{Y(0) \in \cdot\}$. In general, if the process $Y(t)$ can be simulated, then the distribution $\pi$ is known explicitly so that $\alpha$ can be calculated analytically from (2.10). Thus, the interest in Monte Carlo estimation of $\alpha$ occurs when $Y(\cdot)$ is non-stationary. In the simulation literature, such non-stationary estimation problems are referred to as transient simulations.

## 3. A General Framework for Randomized Estimation of Time Integrals.

In this section, we assume that:

A8. there exists a sequence of processes $\{H_k : k \geq 1\}$ such that $H_k : \mathbf{R} \times \Omega \to \mathbf{R}^+$ and which satisfy:

i.) $H_k(\cdot,\omega)$ is non-decreasing and right continuous, for each $\omega \in \Omega$

ii.) $\{H_k : k \geq 1\}$ is a sequence of i.i.d. random elements, which is independent of the collection $\{(X_k, S_k) : k \geq 1\}$

iii.) $EH_k(t) = G(t)$ for all $t \in \mathbf{R}$.

(3.1). Lemma. Under A5 and A8, $H_k(0-) = 0$ a.s. and $EH_k(\infty) = 1$.

Proof. For $t < 0$, $H_k(t)$ is non-negative and satisfies $EH_k(t) = G_k(t) = 0$, so that $H_k(t) = 0$ a.s., proving that $H_k(0-) = 0$ a.s. For the finiteness of $EH_k(\infty)$, observe that $H_k(n)$ increases to $H_k(\infty)$. Hence, by monotone convergence,

$$EH_k(\infty) = \lim_{n\to\infty} EH_k(n) = \lim_{n\to\infty} G(n) = 1 \quad . \qquad ||$$

Our randomized $\bar{I}(n)$ - estimator will be based on

$$(3.2) \qquad R_k(\omega) = \begin{cases} \int_0^\infty X_k(t,\omega)H_k(dt,\omega); & \omega \notin A_k \\ 0 \quad ; & \omega \in A_k \end{cases}$$

where $A_k = \{\int_0^\infty |X_k(t)|H_k(dt) = \infty\}$. Observe that for each $\omega$, $R_k(\omega)$ is well-defined by formula (3.2), since $X_k$ is product measurable, and $H_k$ satisfies A8 i.). Several later arguments will require the following approximation result. Let

$$X_{km}(t) = \min\{X_k(t), m\}$$

$$X^n_{km}(t) = \min\{X_k(2^{-n}(j+1)), m\}$$

for $j < 2^n t < j+1$.

(3.3). Proposition. Assume A1 - A6, A8, and suppose that $X_k$ is non-negative. Then

$$\lim_{m\to\infty} \lim_{n\to\infty} \int_0^\infty X^n_{km}(t) H_k(dt) = \int_0^\infty X_k(t) H_k(dt) \text{ a.s.} \tag{3.4}$$

Proof. First, it is easily seen that $X_{km}(\cdot,\omega)$ is right continuous. The right continuity of $X_{km}$ implies that

$$\lim_{n\to\infty} X^n_{km}(t) = X_{km}(t)$$

for all $t > 0$. Now, $H_k(\infty) < \infty$ a.s., so that we can apply a.s. the bounded convergence theorem to conclude that

$$\lim_{n\to\infty} \int_0^\infty X^n_{km}(t) H_k(dt) = \int_0^\infty X_{km}(t) H_k(dt) \text{ a.s.}$$

Now, apply the monotone convergence theorem to complete the proof. ||

Since the integrals on the left-hand side of (3.4) are discrete sums, it is clear the the right-hand side of (3.4) is a r.v. and that $A_k \in F$. Observe now that if $X_k(t)$ is right continuous with left limits, then the same property holds for

$$X^+_k(t) = \begin{cases} X_k(t); & X_k(t) > 0 \\ 0 \; ; & \text{else} \end{cases} .$$

Thus, by splitting $X_k$ into its positive and negative parts, and using (3.4), we see that $R_k$ is a r.v.

We will use Proposition 3.3 to prove that $\overline{R}(n) = \sum_{k=1}^{n} R_k/n$ is a randomized $\overline{I}(n)$ - estimator. Let $G = \sigma(X_j : j > 1)$ (the $\sigma$-field generated by the $X_j$'s).

(3.5). Theorem. Assume A1 - A6 and A8. Then $E\{R_k \mid \mathcal{G}\} = I_k$.

Proof. Let $f : \mathbf{R}^p \to \mathbf{R}$ be a non-negative bounded continuous function with compact support. Fix an arbitrary selection of time indices $t_1,\ldots,t_p$, and write $f(X_k)$ for $f(X_k(t_1),\ldots,X_k(t_p))$. Assume, for the moment that $X_k$ is non-negative. Then,

$$(3.6)\quad E\{f(X^n_{km}) \int_0^\infty X^n_{km}(t)H_k(dt)\}$$

$$= \sum_{j=0}^{\infty} E\{f(X^n_{km})X^n_{km}(j2^{-n})(H_k(j2^{-n}) - H_k((j-1)2^{-n}))\}$$

$$= \sum_{j=0}^{\infty} E\{f(X^n_{km})X^n_{km}(j2^{-n})(G(j2^{-n}) - G((j-1)2^{-n}))\}$$

$$= E\{f(X^n_{km}) \int_0^\infty X^n_{km}(t)G(dt)\}$$

where the second equality follows from independence of $X_k$ and $H_k$. Since $X_{km}$ is right continuous, $f(X^n_{km}) \to f(X_{km})$ as $n \to \infty$ by continuity of $f$, so that Proposition 3.3 implies that

$$(3.7)\quad \lim_{n\to\infty} f(X^n_{km}) \int_0^\infty X^n_{km}(t)H_k(dt) = f(X_{km}) \int_0^\infty X_{km}(t)H_k(dt) \text{ a.s.}$$

Now, $X^n_{km}$ is bounded by $m$ and $H_k(\infty)$ is integrable (Lemma 3.1) so evidently the left-hand side of (3.7) is dominated by an integrable r.v. Hence, one can apply dominated convergence to (3.6), yielding

$$(3.8)\quad \begin{aligned} &E\{f(X_{km}) \int_0^\infty X_{km}(t)H_k(dt)\} \\ &= E\{f(X_{km}) \int_0^\infty X_{km}(t)\, G(dt)\} . \end{aligned}$$

Since $f$ has compact support, $f(X_{km}) = f(X_{kM})$ for all $m$ greater than some $M$. Then, the non-negativity of $f$ implies that the r.v.'s on both sides of (3.8) are increasing in m, for $m > M$. Hence, monotone convergence can be applied to (3.8), yielding

$$E\{f(X_k) \int_0^\infty X_k(t)H_k(dt)\}$$

$$= E\{f(X_k) \int_0^\infty X_k(t)\ G(dt)\} \quad .$$

By appealing to the montone class theorem ([7], p.14) and using the fact that the functions $f(X_k)$ generate $\sigma(X_k)$, one obtains

$$E\{Z \int_0^\infty X_k(t)H_k(dt)\} = E\{Z \int_0^\infty X_k(t)G(dt)\}$$

for all bounded $Z \in \sigma(X_k)$. Hence, by definition of conditional expectation,

$$E\{\int_0^\infty X_k(t)H_k(dt) \mid X_k\} = \int_0^\infty X_k(t)G(dt) \quad .$$

In particular, $E\{\int_0^\infty |X_k(t)|H_k(dt)\} = \int_0^\infty E|X_k(t)|G(dt) < \infty$ (see A6), so $P(A_k) = 0$ for $k > 1$. Thus,

$$\int_0^\infty X_k(t)H_k(dt) = R_k \text{ a.s.}$$

so that $E\{R_k \mid X_k\} = I_k$ a.s. Since $R_k \in \sigma(X_k, H_k)$, $R_k$ is independent of $\sigma(X_j : j \neq k)$ and hence ([6], p. 308)

$$E\{R_k \mid G\} = E\{R_k \mid X_k\} = I_k \text{ a.s.} \quad ,$$

completing the proof of the theorem in the case that $X_k > 0$. For the general situation, split $X_k$ into its positive and negative parts, repeat the above argument, and recombine using the fact that $P(A_k) = 0$. ||

(3.9). <u>Corollary</u>. Under A1 - A6 and A8, $ER_k = \alpha$.

We wish to prove limit theorems in terms of the parameter t, where t corresponds to an index of effort. In the applications we will be considering in Sections 4 through 6, the effort required to generate $H_k$ will be negligible compared to the effort necessary to simulate $I_k$. Then, $S_k(\tau_k)$ is the effort required to generate $R_k$, where $\tau_k = \sup\{t : H_k(t) < H_k(\infty)\}$. Note that $\tau_k$ and $S_k(\tau_k)$ are $F$-measurable, due to right continuity of $H_k$ and $S_k$.

(3.10). Proposition. Assume A1 - A6 and A8. Then

$$E\{S_k(\tau_k);\ \tau_k < \infty\} = \int_0^\infty ES_k(t)F(dt) \tag{3.11}$$

where $F(t) = P\{\tau_k \le t\}(E\{S_k(\tau_k);\ A\} \triangleq ES_k(\tau_k)I_A$, where $I_A(\omega)$ is 1 or 0 depending on whether or not $\omega \in A$).

Proof. Let $S_{km}(t) = \min\{S_k(t),m\}$ and set $\tau_k^n = 2^{-n}(j+1)$ on $\{j < 2^n\tau_k \le j+1\}$. Then,

$$\begin{aligned} E\{S_{km}(\tau_k^n);\ \tau_k < \infty\} &\\ &= \sum_{j=0}^{\infty} E\{S_{km}(\tau_k^n);\ 2^{-n}j < \tau_k \le 2^{-n}(j+1)\} \\ &= \sum_{j=0}^{\infty} ES_{km}(2^{-n}(j+1))P\{2^{-n}j < \tau_k \le 2^{-n}(j+1)\} \end{aligned} \tag{3.12}$$

where the second inequality is due to independence of $S_k$ and $\tau_k$. Applying bounded convergence and then monotone convergence to (3.12) yields (3.11). ||

Let $M(t) = \max\{k : S_1(\tau_1) + \ldots + S_k(\tau_k) \le t\}$; then $M(t)$ is the number of $R_k$'s generated with effort $t$. The next lemma shows that $M(t)$ is no smaller than the number of $I_k$'s generated.

(3.13). Lemma. Assume A1 - A6 and A8. Then $M(t) \ge N(t)$ a.s. for $t \ge 0$.

Proof. We need to show $S_k(\tau_k) \le S_k(T)$ a.s.; this inequality is trivial if $T = \infty$. By monotone convergence,

$$E\{H_k(\infty) - H_k(T)\} = \lim_{n\to\infty} E\{H_k(T+n) - H_k(T)\} = \lim_{n\to\infty} G(T+n) - G(T) = 0 \quad .$$

The non-negativity of $H_k(\infty) - H_k(T)$ implies that $H_k(\infty) = H_k(T)$ a.s., proving that $\tau_k \le T$. ||

Let $R(t) \triangleq \bar{R}(M(t))$; $R(t)$ is a randomized estimator constructable from $t$ units of effort.

(3.14). <u>Theorem</u>. Assume A1 - A6 and A8. If $P\{\tau_k < \infty\} = 1$, then $R(t) \to \alpha$ a.s. as $t \to \infty$.

<u>Proof</u>. Since $EH_k(\infty) = 1$ a.s. (Lemma 3.1), $P\{H_k(\infty) > 0\} > 0$, so that $P\{\tau_k > 0\} > 0$. Hence, $P\{S_k(\tau_k) > 0\} > 0$ by Proposition 3.10 and A3. This, in turn, implies that $M(t) < \infty$ a.s. The fact that $\tau_k < \infty$ a.s. forces $S_k(\tau_k)$ to be finite a.s. (see A1), which assures that $M(t) \to \infty$ a.s. as $t \to \infty$. The theorem then follows from the strong law for $\bar{R}(n)$ and Corollary 3. ||

To obtain analogs of Theorems 2.3 and 2.8, we need to analyze $E(R_k - \alpha)^2$. Our expression will involve the function $K(s,t)$, where $K(s,t) = EH_k(s)H_k(t)$.

A9. $EH_k^2(t) < \infty$ for $t \in R^+$ .

(3.15). <u>Lemma</u>. Under A8 - A9, $K(s,t)$ is the distribution function of a $\sigma$-finite measure on $R^2$.

<u>Proof</u>. Note that

$$H_k(s)H_k(t) \leq H_k^2(t)$$

for $s \leq t$. Hence, $K(s,t) \leq K(t,t) < \infty$ so $K(s,t)$ is real-valued. Let $s_n$, $t_n$ decrease to $s$, $t$ respectively. Since $H(s_k)H(t_k) \leq H(s_1)H(t_1)$, the finiteness of $K(s_1,t_1)$ allows application of dominated convergence to prove that $K(s_n,t_n) \to K(s,t)$; $K$ is therefore continuous from above. Also, for $s_1 < s_2$, $t_1 < t_2$,

$$H(s_2,t_2) - H(s_1,t_2) - H(s_2,t_1) + H(s_1,t_1) \geq 0$$

so that

$$K(s_2,t_2) - K(s_1,t_2) - K(s_2,t_1) + K(s_1,t_1) \geq 0 \quad .$$

The lemma then follows from Theorem 12.5 of Billingsley (1979). ||

A10. $\int_0^\infty \int_0^\infty E|X_k(s)X_k(t)|\ K(ds,dt) < \infty$ .

(3.16). Theorem. Under A1 - A6 and A8 - A10,

$$ER_k^2 = \int_0^\infty \int_0^\infty EX_k(s)X_k(t)\ K(ds,dt) < \infty \quad . \tag{3.17}$$

Proof. For $X_k$ non-negative, observe that for any integer T,

$$E\{\int_0^T \int_0^T X_{km}^n(s)X_{km}^n(t)H_k(ds)H_k(dt)\} \tag{3.18}$$

$$= \sum_{j=0}^{2^nT} \sum_{\ell=0}^{2^nT} E\{X_{km}^n(j2^{-n})X_{km}^n(\ell 2^{-n})\Delta H_k(j,\ell,n)\}$$

$$= \sum_{j=0}^{2^nT} \sum_{\ell=0}^{2^nT} EX_{km}^n(j2^{-n})X_{km}^n(\ell 2^{-n})\Delta K(j,\ell,n)$$

$$= \int_0^T \int_0^T EX_{km}^n(s)X_{km}^n(t)K(ds,dt)$$

where

$$\Delta H_k(j,\ell,n) = (H_k(j2^{-n}) - H_k((j-1)2^{-n}))(H_k(\ell 2^{-n}) - H_k((\ell-1)2^{-n}))$$

$$\Delta K(j,\ell,n) = K(j2^{-n},\ell 2^{-n}) - K((j-1)2^{-n},\ell 2^{-n}) - K(j2^{-n},(\ell-1)2^{-n}) + K((j-1)2^{-n},(\ell-1)2^{-n}) \quad .$$

Letting n, m, and T tend to infinity in (3.18) (in that order), bounded and monotone convergence proves (3.17) for $X_k$ non-negative. For $X_k$ of mixed sign, split $X_k$ into its positive and negative parts and recombine using A10. ||

Let $\sigma_R^2 = \int_0^\infty \int_0^\infty EX_k(s)X_k(t)K(ds,dt) - \alpha^2 = E(R_k-\alpha)^2$ and set $s(t) = ES_k(t)$. The proof of the following theorem is identical to those of Theorems 2. and 2.8.

(3.19). Theorem. Assume A1 - A6 and A8 - A10. If $P\{S_k(\tau_k) < \infty\} = 1$, then

$$\overline{\lim_{t\to\infty}}\ a(t) \cdot |R(t) - \alpha| = (\sigma_R^2 \int_0^\infty s(t)F(dt))^{1/2} \text{ a.s.} \tag{3.20}$$

where $\sigma_R^2 < \infty$. If, in addition, $ES_k(\tau_k) < \infty$, then

$$t^{1/2}(R(t) - \alpha) \Longrightarrow \sigma_R \left(\int_0^\infty s(t)F(dt)\right)^{1/2} N(0,1) \tag{3.21}$$

as $t \to \infty$.

On the basis of the CLT (3.21), it is natural to interpret $\sigma_R^2 ES_k(\tau_k)/t$ as the asymptotic variance of the estimator $R(t)$.

In addition to the estimator $R(t)$, one can construct a second estimator based on the framework described thus far. Let $\bar{h}(n) = \sum_{k=1}^{n} H_k(\infty)/n$ and set

$$\tilde{R}(n) = \begin{cases} \bar{R}(n)/\bar{h}(n); & \bar{h}(n) > 0 \\ 0 \quad ; & h(n) = 0 \end{cases} .$$

The estimator $\tilde{R}(n)$ is merely $\bar{R}(n)$ normalized by the random total mass of the first $n$ $H_k$'s; we therefore refer to $\tilde{R}(n)$ as a normalized randomized $\bar{I}(n)$ - estimator, and set $\hat{R}(t) = \tilde{R}(M(t))$.

The following result has an identical proof to that of Theorem 3.14 (recall that $EH_k(\infty) = 1$ by Lemma 3.1 so that $\bar{h}(n) \to 1$ a.s.).

(3.22). Theorem. Assume A1 - A6 and A8. If $P\{\tau_k < \infty\} = 1$, then $\hat{R}(t) \to \alpha$ a.s. as $t \to \infty$.

An analog to Theorem 3.19 is also available, under a certain moment condition.

A11. $\int_0^\infty \int_0^\infty E|\hat{X}_k(s)\hat{X}_k(t)|K(ds,dt) < \infty$ ,

where $\hat{X}_k(s) = X_k(s) - \alpha$ .

(3.23). Theorem. Assume A1 - A6, A8, A9, and A11. If $P\{S_k(\tau_k) < \infty\} = 1$, then

$$\overline{\lim_{t\to\infty}}\, a(t) \cdot |\hat{R}(t) - \alpha| = \left(\hat{\sigma}_R^2 \int_0^\infty s(t)F(dt)\right)^{1/2} \text{ a.s.} \tag{3.24}$$

where $\hat{\sigma}_R^2 = \int_0^\infty \int_0^\infty E\hat{X}_k(s)\hat{X}_k(t)K(ds,dt) = E\hat{R}_k^2 < \infty$ ($\hat{R}_k \triangleq R_k - \alpha H_k$). If, in addition, $ES_k(\tau_k) < \infty$, then

$$t^{1/2}(\hat{R}(t) - \alpha) \Longrightarrow \hat{\sigma}_R \left(\int_0^\infty s(t)F(dt)\right)^{1/2} N(0,1) \tag{3.25}$$

as $t \to \infty$.

Proof. Note that if $X_k(t)$ satisfies A1 - A7, then $\hat{X}_k(t)$ also does. Under A11, one can therefore apply Theorem 3.19 to the process $\hat{X}_k$ to obtain the relation

$$(3.26)\quad \overline{\lim_{t\to\infty}}\ a(t) \cdot \frac{1}{M(t)} \left|\sum_{k=1}^{M(t)} \int_0^\infty \hat{X}_k(s)H_k(ds)\right| = (\hat{\sigma}_R^2 \int_0^\infty s(t)F(dt))^{1/2} \text{ a.s.}$$

But observe that

$$\frac{1}{M(t)} \sum_{k=1}^{M(t)} \int_0^\infty \hat{X}_k(s)H_k(ds) = \frac{1}{M(t)} \sum_{k=1}^{M(t)} \int_0^\infty (X_k(s) - \alpha)H_k(ds)$$
$$= \bar{h}(M(t))(\hat{R}(t) - \alpha) \quad .$$

Since $\bar{h}(M(t)) \to 1$ a.s. as $t \to \infty$, (3.24) follows from (3.26). A similar proof is valid for (3.25). ||

We wish to emphasize that A10 and A11 are not equivalent moment hypotheses. Consider a case in which $X_k(t) = 1$ for $0 < t < 1$ and vanishes elsewhere, with $0 < G(1) < 1$. Then, A10 is always satisfied but A11 is valid only if $K(\infty,\infty) < \infty$. On the other hand, suppose that $X_k(t) = 1$ for $t > 0$. Then, A11 always holds but A10 is valid only if $K(\infty,\infty) < \infty$.

As in the case of $R(t)$, relation (3.25) suggests that $\hat{\sigma}_R^2\, ES_k(\tau_k)/t$ may be interpreted as the asymptotic variance of the estimator $\hat{R}(t)$. The goal, in Sections 4 through 6, will be to determine randomization schemes that make $\sigma_R^2\, ES_k(\tau_k)$ and/or $\hat{\sigma}_R^2\, ES_k(\tau_k)$ as small as possible.

Certain bounds for $\sigma_R^2$ and $\hat{\sigma}_R^2$ are available. The lower bound $\sigma_R^2 > \sigma^2$ can be obtained from Proposition 1.9, under A1 - A8. Theorem 3.5 asserts that

$$E\{\int_0^\infty \hat{X}_k(t)H_k(dt) \mid G\} = I_k - \alpha$$

from which it follows, by Proposition 1.9, that $\hat{\sigma}_R^2 > \sigma^2$. Note that the lower bounds are attained for $H_k = G$ a.s. Upper bounds are also available, under certain conditions.

(3.27). Proposition. Assume A1 - A6 and A8. Then, if $EH_k^3(\infty) < \infty$,

(3.28) $$\sigma_R^2 \le \int_0^\infty EX_k^2(t)G(dt) \cdot EH_k^3(\infty) - \alpha^2$$

(3.29) $$\hat{\sigma}_R^2 \le (\int_0^\infty EX_k^2(t)G(dt) - \alpha^2) \cdot EH_k^3(\infty) \quad .$$

The upper bounds are attained.

Proof. By Cauchy-Schwartz,

$$R_k^2 = (\int_0^\infty X_k(t)H_k(dt))^2 \le (\int_0^\infty X_k^2(t)H_k(dt))H_k(\infty) \quad .$$

Replicating the approximation argument of Theorem 3.16 proves that

(3.30) $$ER_k^2 \le \int_0^\infty EX_k^2(t)K(dt,\infty) \quad .$$

Let $f$ be a bounded continuous function with compact support. Then, the definition of $K$ implies that

$$\sum_{j=0}^{\infty} f(j2^{-n})\{K(j2^{-n},\infty) - K((j-1)2^{-n},\infty)\}$$

$$= E\Big\{\sum_{j=0}^{\infty} f(j2^{-n})\{H_k(j2^{-n}) - H_k((j-1)2^{-n})\}H_k(\infty)\Big\} \quad .$$

Letting $n \to \infty$, we obtain

(3.31) $$\int_0^\infty f(t)K(dt,\infty) = E\{\int_0^\infty f(t)H_k(dt)H_k(\infty)\} \quad .$$

The monotone class theorem extends (3.31) to include all bounded Borel measurable $f$. But, by Cauchy-Schwartz,

$$E\{\int_A H_k(dt) \cdot H_k(\infty)\} = E\{\int_A H_k(dt))^{1/2} \cdot (\int_A H_k(dt) \cdot H_k(\infty)^2)^{1/2}\}$$

$$\le E\{\int_A H_k(dt)\} \cdot E\{\int_A H_k(dt) \cdot H_k(\infty)^2\}$$

$$\le \int_A G(dt) \cdot EH_k^3(\infty)$$

so that we may conclude that $K(dt,\infty) \le G(dt)EH_k^3(\infty)$. Applying this inequality

(3.30) yields (3.28). Inequality (3.29) is obtained by letting $\bar{X}_k(t)$ play the role of $X_k(t)$ in (3.28).

The upper bounds are attained by taking $H_k(t) = 0$ for $t < M_k$ and 1 for $t \geq M_k$, where $\{M_k\}$ is an i.i.d. sequence of r.v.'s having common distribution G. ||

The inequalities (3.28) and (3.29) may be easily adapted to provide bounds on the moments entering hypotheses A10 and A11, respectively.

(3.32). Proposition. Assume A1 - A5 and A8. Then, A6 - A7 and A9 - A11 are satisfied if $EH_k^3(\infty) < \infty$ and $\int_0^\infty EX_k^2(t)G(dt) < \infty$.

Proof. First, observe that $E|X_k(t)| \leq EX_k^2(t)$ so A6 is satisfied. Also, $EH_k^3(\infty) < \infty$ implies that $K(\infty,\infty) = EH_k^2(\infty) < \infty$, yielding A9. As for A10, the argument of Proposition 3.27 is easily adapted to show that

$$(3.33)\qquad \int_0^\infty \int_0^\infty E|X_k(s)X_k(t)|K(ds,dt) \leq \int_0^\infty EX_k^2(t)G(dt) \cdot EH_k^3(\infty) < \infty \quad .$$

For A11, substitute $\hat{X}$ for X in (3.33), and observe that

$$\int_0^\infty E\hat{X}_k^2(t)G(dt) = \int_0^\infty EX_k^2(t)G(dt) - \alpha^2 \quad .$$

As for A7, $\sigma^2$ is a lower bound on $\sigma_R^2$. ||

We turn now to a variance reduction technique that is sometimes applicable to the randomized estimators developed above. Let $0 = t_0 < t_1 <\ldots< t_{\ell-1} < t_\ell = \infty$ and set

$$Q_i(\cdot) = P\{H_k \in \cdot \mid t_{i-1} \leq \tau_k < t_i\}$$

for $i = 1,\ldots,\ell$. Suppose that one can generate independent deviates from the measures $Q_1,\ldots,Q_\ell$ (see Sections 5 and 6 for examples) and that $p_i \triangleq P\{t_{i-1} \leq \tau_k < t_i\}$ is known, for $1 \leq i \leq \ell$. These assumptions will allow us to stratify our sampling scheme.

Let $\{m_k : k \geq 1\}$ be a sequence of integers taking values in $\{1,\ldots,\ell\}$; the $m_k$'s define a sampling order for the $Q_i$'s. If $m_k = 1$, one generates $H_k^s$ from $Q_i$ (independently of the $(X_k,S_k)$'s), and sets

$$R_k^s = \int_0^\infty X_k(t)H_k^s(dt) \quad .$$

If $\omega_{ni} = \{j \leq n : m_j = 1\}$ $(1 \leq i \leq \ell)$, and $k_{ni}$ is the cardinality of $\omega_{ni}$, one sets

$$\overline{R}^s(n) = \sum_{i=1}^{\ell} p_i \sum_{j\varepsilon\omega_{ni}} R_j^s/k_{ni}$$

$$\overline{h}^s(n) = \sum_{i=1}^{\ell} p_i \sum_{j\varepsilon\omega_{ni}} H_j^s(\infty)/k_{ni} \quad .$$

Let $\tau_k^s = \{t : H_k^s(t) < H_k^s(\infty)\}$, and set $N^s(t) = \max\{n : S_1(\tau_1^s) + \ldots + S_n(\tau_n^s) \leq t\}$; $N^s(t)$ is the number of $R_k^s$'s generated in $t$ units of effort. We wish to study the stratified estimators $R^s(t) = \overline{R}^s(N^s(t))$ and $\hat{R}^s(t) = \overline{R}^s(N^s(t))/\overline{h}^s(N^s(t))$.

A12. $k_{ni}/n \to c_i > 0$ as $n \to \infty$, for $1 \leq i \leq \ell$.

Let $d_i = E\{S_k(\tau_k) \mid t_{i-1} \leq \tau_k < t_i\}$, $\sigma_i^2 = E\{R_k^2 \mid t_{i-1} \leq \tau_k < t_i\} - (E\{R_k \mid t_{i-1} \leq \tau_k < t_i\}^2)$, $\hat{\sigma}_i^2 = E\{\hat{R}_k^2 \mid t_{i-1} \leq \tau_k < t_i\} - (E\{\hat{R}_k \mid t_{i-1} \leq \tau_k < t_i\}^2)$. The following theorem is a stratified analog of the CLT's (3.21) and (3.25).

(3.32). <u>Theorem</u>. Assume A1 - A6 and A8 - A12. If $ES_k(\tau_k) < \infty$, then

$$t^{1/2}(R^s(t) - \alpha) \Longrightarrow \sigma_{RS}\Big(\sum_{i=1}^{\ell} c_i d_i\Big)^{1/2} N(0,1) \quad . \tag{3.33}$$

$$t^{1/2}(\hat{R}^s(t) - \alpha) \Longrightarrow \hat{\sigma}_{RS}\Big(\sum_{i=1}^{\ell} c_i d_i\Big)^{1/2} N(0,1) \quad . \tag{3.34}$$

as $t \to \infty$, where $\sigma_{RS}^2, \hat{\sigma}_{RS}^2 < \infty$ and $\sigma_{RS}^2 = \sum_{i=1}^{\ell} p_i^2\sigma_i^2/c_i$, $\hat{\sigma}_{RS}^2 = \sum_{i=1}^{\ell} p_i^2\hat{\sigma}_i^2/c_i$.

*Proof.* Assumptions A10 and A11, together with $ES_k(\tau_k) < \infty$, clearly guarantee the finiteness of $\sigma^2_{RS}$, $\hat{\sigma}^2_{RS}$, and $\sum_{i=1}^{\ell} c_i d_i$. Under A12,

$$(3.35) \qquad \sum_{k=1}^{n} S_k(\tau_k^s)/n = \sum_{i=1}^{\ell} \Big( \sum_{j\epsilon\omega_{ni}} S_j(\tau_j^s)/k_{ni} \Big) \cdot (k_{ni}/n) \to \sum_{i=1}^{\ell} c_i d_i \text{ a.s.}$$

as $n \to \infty$. But

$$(3.36) \qquad \sum_{k=1}^{N^s(t)} S_k(\tau_k^s)/N^s(t) \le t/N^s(t) < \sum_{k=1}^{N^s(t)+1} S_k(\tau_k^s)/N^s(t)$$

on $\{N^s(t) < \infty\}$. However, as in the proof of Theorem 2.2, it is easily argued that $N^s(t) < \infty$ a.s. and $N^s(t) \to \infty$ a.s. as $t \to \infty$. Using this fact in (3.36), (3.35) implies that $t/N^s(t)$ is "squeezed" between two terms converging to $\sum_{i=1}^{\ell} c_i d_i$. Hence,

$$(3.37) \qquad N^s(t)/t \to \Big( \sum_{i=1}^{\ell} c_i d_i \Big)^{-1} \text{ a.s.}$$

The central limit theorem and the converging-together lemma ([2], p. 25) imply that

$$(3.38) \qquad n^{1/2} \Big( \sum_{j\epsilon\omega_{ni}} R_j^s - ER_j^s \Big)/k_{ni} \Longrightarrow (\sigma_i/c_i^{1/2})N(0,1)$$

as $n \to \infty$. Since the r.v.'s $\sum_{j\epsilon\omega_{ni}} R_j^s$ are independent,

$$(3.39) \qquad n^{1/2} (\overline{R}^s(n) - \alpha) \Longrightarrow \Big( \sum_{i=1}^{\ell} p_i^2\sigma_i^2/c_i \Big)^{1/2} N(0,1)$$

as $n \to \infty$. Arguing as in Theorem 2.8, we see that (3.39) and (3.37) together yield (3.33); a similar proof works for (3.34). ||

We now turn to the question of determining the $c_i$'s.

A13. $p_i\sigma_i^2\hat{\sigma}_i^2 > 0$ , $1 \le i \le \ell$ .

(3.40). *Proposition.* Under A13, the minimum of $g(\vec{c}) \triangleq \sigma^2_{RS}\Big( \sum_{i=1}^{\ell} c_i d_i \Big)$ over

$\{\vec{c} : c_i > 0\}$ is $\left(\sum_{i=1}^{\ell} p_i \sigma_i d_i^{1/2}\right)^2$ and is achieved at

$$\underset{\sim}{c}_i = \underset{\sim}{\alpha}\, p_i \sigma_i / d_i^{1/2}$$

(3.41)

$$\underset{\sim}{\alpha} = \sum_{i=1}^{\ell} p_i \sigma_i / d_i^{1/2} \quad .$$

Similarly, the minimum of $h(\vec{c}) \triangleq \hat{\sigma}^2_{RS}\left(\sum_{i=1}^{\ell} c_i d_i\right)$ over $\{\vec{c} : c_i > 0\}$ is $\left(\sum_{i=1}^{\ell} p_i \hat{\sigma}_i^2 d_i^{1/2}\right)^2$ and is achieved at

$$\tilde{c}_i = \tilde{\alpha} p_i \hat{\sigma}_i / d_i^{1/2}$$

(3.42)

$$\tilde{\alpha} = \sum_{i=1}^{\ell} p_i \hat{\sigma}_i / d_i^{1/2} \quad .$$

The proof involves a simple application of Lagrange multipliers; see Theorem 3 of Glynn (1983) for a similar argument. We shall now show that stratification always provides a variance reduction.

(3.43). Proposition. Assume A1 - A6, A8 - A12, and $ES_k(\tau_k) < \infty$. Then,

(3.44)
$$\left(\sum_{i=1}^{\ell} p_i \sigma_i d_i^{1/2}\right)^2 \leq \sigma_R^2\, ES_k(\tau_k)$$

(3.45)
$$\left(\sum_{i=1}^{\ell} p_i \hat{\sigma}_i d_i^{1/2}\right)^2 \leq \hat{\sigma}_R^2\, ES_k(\tau_k) \quad .$$

Proof. By Cauchy-Schwartz

$$\left(\sum_{i=1}^{\ell} p_i \sigma_i d_i^{1/2}\right)^2 \leq \left(\sum_{i=1}^{\ell} p_i \sigma_i^2\right)\left(\sum_{i=1}^{\ell} p_i d_i\right)$$

(3.46)

$$= \left(\sum_{i=1}^{\ell} p_i \sigma_i^2\right) ES_k(\tau_k) \quad .$$

Using Cauchy-Schwartz again yields

$$\sum_{i=1}^{\ell} p_i\sigma_i^2 = ER_k^2 - \sum_{i=1}^{\ell} p_i(E\{R_k;\ t_{i-1} < \tau_k < t_i\}^2/p_i^2) \tag{3.47}$$

$$> ER_k^2 - \left(\sum_{i=1}^{\ell} p_i E\{R_k;\ t_{i-1} < \tau_k < t_i\}/p_i\right)^2$$

$$= ER_k^2 - \alpha^2 = \sigma_R^2 \quad ,$$

proving (3.44); (3.45) is proved similarly. $||$

Utilization of the constants $\underset{\sim}{c}_i$, $\tilde{c}_i$ specified by (3.41) and (3.42) requires a "trial run", in order to obtain estimates for $\sigma_i$, $\tilde{\sigma}_i$, and $d_i$. If one wishes to dispense with a "trial run" and is willing to accept a sub-optimal choice of the $c_i$'s, consider using $c_i = p_i$. Observe that if $c_i = p_i$, then

$$\sigma_{RS}^2\left(\sum_{i=1}^{\ell} c_i d_i\right) = \left(\sum_{i=1}^{\ell} p_i\sigma_i^2\right)\left(\sum_{i=1}^{\ell} p_i d_i\right)$$

$$< \sigma_R^2 \cdot ES_k(\tau_k) \quad ;$$

the inequality derives from (3.46) and (3.47). Hence, the suboptimal choice $c_i = p_i$ yields a variance reduction over a nonstratified sampling plan.

One final remark is in order. Suppose that there exists $T$ such that $G(t) = 0$ for $t < T$ and $1$ for $t > T$. Then, an argument similar to that used in Lemma 3.13 proves that $H_k = G$ a.s., in which case $R_k = I_k$ a.s. Hence, our randomized framework yields no new estimators when $I = EX(T)$ for some non-random $T$.

## 4. Poisson Process Randomization

In this section, we consider a class of randomized time integral estimators for which the $H_k$'s previously defined are non-homogeneous Poisson processes. Specifically, we shall assume that

$$H_k(t) = \frac{1}{\lambda} N_k(\lambda G(t)) \tag{4.1}$$

for some $\lambda > 0$, where $\{N_k : k \geq 1\}$ is a sequence of i.i.d. Poisson processes for which $EN_k(t) = t$; of course, the family $\{N_k : k \geq 1\}$ is assumed to be independent of $\{(X_k, S_k) : k \geq 1\}$ satisfying A1 - A5. The following proposition is an immediate consequence of standard properties of the Poisson process.

(4.2). Proposition. The sequence $\{H_k : k \geq 1\}$, as defined through (4.1), satisfies A8 and A9. Furthermore, $EH_k^n(\infty) < \infty$ for $n \geq 1$.

Calculation of the distribution $K$ is easy.

(4.3). Proposition. For $0 \leq s \leq t$,

$$K(s,t) = \frac{1}{\lambda} G(s) + G(s)G(t) \quad .$$

Proof. The independent increments property of the Poisson process implies that for $s \leq t$,

$$\begin{aligned} K(s,t) &= \frac{1}{\lambda^2} E(N(\lambda G(s))N(\lambda G(t)) \\ &= \frac{1}{\lambda^2} E(N(\lambda G(s))(N(\lambda G(s)) + N(\lambda G(t)) - N(\lambda G(s))) \\ &= \frac{1}{\lambda^2} EN^2(\lambda G(s)) + \frac{1}{\lambda^2} (\lambda G(s)(\lambda G(t) - \lambda G(s)) \\ &= \frac{1}{\lambda} G(s) + G(s)G(t). \qquad || \end{aligned}$$

Theorems 3.19 and 3.23 require moment hypotheses on the process $X_k$.

B1. $$\int_0^\infty EX_k^2(t)G(dt) < \infty \quad .$$

Propositions 3.32 and 4.2 prove that B1 is sufficient to guarantee A6 - A7 and A10 - A11.

(4.4). Theorem. Under A1 - A5 and B1,

(4.5) $$\sigma_R^2 = \sigma^2 + \frac{1}{\lambda}\int_0^\infty EX_k^2(t)G(dt)$$

(4.6) $$\hat{\sigma}_R^2 = \sigma^2 + \frac{1}{\lambda}\left(\int_0^\infty EX_k^2(t)G(dt) - \alpha^2\right) \quad .$$

Proof. Note that

$$\begin{aligned}\sigma^2 &= E\{\int_0^\infty \int_0^\infty X_k(s)X_k(t)G(ds)G(dt)\} - \alpha^2 \\ &= \int_0^\infty \int_0^\infty EX_k(s)X_k(t)G(ds)G(dt) - \alpha^2 \\ &= \int_0^\infty \int_0^\infty E\hat{X}_k(s)\hat{X}_k(t)G(ds)G(dt) \quad .\end{aligned}$$

Hence, the result follows from Proposition 4.3 and the fact that the distribution $K_1(s,t) = G(\min(s,t))$ induces a measure on $\mathbb{R}^2$ which has support on the diagonal $s = t$. ||

Theorem 4.4 suggests that the normalized estimator $\hat{R}(t)$ has better large sample properties than $R(t)$; consequently, we shall consider only $\hat{R}(t)$ for the remainder of this section. We shall now calculate $F$.

(4.7). Proposition. The distribution $F(x) = P\{\tau_k < x\}$ is given by

$$F(x) = \exp(-\lambda(1 - G(x))) \quad .$$

Proof. Evidently,

(4.8) $$\begin{aligned}\tau_k &= \sup\{t : H_k(t) < H_k(\infty)\} \\ &= \inf\{t : H_k(t) > H_k(\infty)\} \\ &= \inf\{t : N_k(\lambda G(t)) > N_k(\lambda)\} \\ &= \inf\{t : G(t) > T_k(N_k(\lambda))/\lambda\}\end{aligned}$$

where $T_k(j) = \inf\{t : N_k(t) > j\}$. Let $A(t) = P\{t - T_k(N_k(t)) > u\}$. Then $A(t)$ solves the renewal equation

(4.9) $$A(t) = a(t) + \int_0^t A(t-s)e^{-s}ds$$

where

$$a(t) = \begin{cases} 0 \ ; & t < u \\ e^{-t} \ ; & t > u \ . \end{cases}$$

Since the renewal function $M(t) = EN_k(t) = t$ is known, (4.9) can be solved explicitly (see [6], p. 294):

$$(4.10) \qquad A(t) = \begin{cases} 0 & ; \; t < u \\ e^{-u} & ; \; t > u \end{cases} .$$

Now, (4.8) implies that $\{\tau_k < x\} = \{T_k(N_k(\lambda))/\lambda > G(x)\}$ so (4.10) yields

$$F(x) = P\{\lambda - T_k(N_k(\lambda)) > \lambda(1 - G(x))$$
$$= \exp(-\lambda(1 - G(x)). \quad ||$$

We shall now assume that:

B2. $\qquad s(t) = ES_k(t)$ is strictly increasing.

(4.11). Proposition. Under B2,

$$ES_k(\tau_k) = \int_0^\infty \{1 - \exp(-\lambda(1 - G(s^{-1}(x))))\}dx \quad .$$

The proof is immediate from Propositions 3.10 and 4.5.

Theorem 4.4 and Proposition 4.11 can be used to optimize the value of $\lambda$ so as to minimize the asymptotic variance of $\hat{R}(t)$. For example, suppose that $G(t) = t/T$ for $t < T$ and 1 for $t > T$, and let $s(t) = t$. Then the goal is to find $\lambda^*$ to minimize (see Theorem 3.23)

$$(\sigma^2 + \frac{1}{\lambda}\beta)(\int_0^T \{1 - \exp(-\lambda(1 - x/T))\}dx)$$

where $\beta = \int_0^\infty E\hat{X}_k^2(t)G(dt)$; differentiation leads to a non-linear equation which $\lambda^*$ must solve.

In any case, the inequality $\exp(-\lambda u) > 1 - \lambda u$ for $\lambda u > 0$ leads to the bound

$$ES_k(\tau_k) < \lambda \int_0^\infty (1 - G(s^{-1}(x)))dx \quad .$$

Hence, Poisson randomization with intensity $\lambda$ is more efficient than $\overline{I}(n)$ if

$$(\sigma^2 + \beta/\lambda)(\lambda \int_0^\infty (1 - G(s^{-1}(x)))dx) < \sigma^2 s(T) \quad .$$

5. Point Mass Randomization

Suppose that $\{(X_k, S_k) : k \geq 1\}$ is a family of processes satisfying A1 - A5 and

C1. $$\int_0^\infty EX_k^2(t)G(dt) < \infty \quad .$$

Let h be a non-negative function for which:

C2. $$\{t : h(t) > 0\} = [\gamma,\infty), \int_0^\infty h(t)G(dt) = 1$$

C3. $$\int_0^\infty I_{[\gamma,\infty)}(u)/h^2(u)G(du) < \infty \quad .$$

The family $\{H_k : k \geq 1\}$ of point mass randomized estimates will be defined through the rule

(5.1) $$H_k(t) = \begin{cases} \min\{G(t), G(\gamma-)\}; & t < M_k \\ G(\gamma-) + 1/h(M_k); & t \geq M_k \end{cases}$$

where $\{M_k : k \geq 1\}$ is a sequence of i.i.d. r.v.'s, independent of the $(X_k, S_k)$'s, having common distribution $P\{M_k \in dt\} = h(t)G(dt)$. With the $H_k$'s defined through (5.1),

$$R_k = \int_0^\infty I_{[0,\gamma)}(u)X_k(u)G(du) + X(M_k)/h(M_k) \quad .$$

(5.2). Proposition. Under C1 - C3, A6 - A11 hold.

Proof. First, observe that $H_k(t) = G(t)$ for $t < \gamma$. For $t \geq \gamma$, C1 proves that

$$\begin{aligned} EH_k(t) &= G(\gamma-) + E\{1/h(M_k); M_k \leq t\} \\ &= G(\gamma-) + \int_0^\infty I_{[\gamma,t]}(u)/h(u) \cdot h(u)G(du) \\ &= G(t) \end{aligned}$$

so that $H_k$ satisfies A8. For A6 - A7 and A9 - A11, we apply Proposition 3.32. It is evident that $H_k(\infty) = G(\gamma-) + 1/h(M_k)$ so that it suffices to prove that $E(1/h(M_k))^3 < \infty$. But

$$Eh(M_k)^{-3} = \int_0^\infty I_{[\gamma,\infty)}(u)h(u)^{-3} \cdot h(u)G(du)$$

which is finite by C3. ||

It should be clear that the moment assumptions C1 and C3 are considerably stronger than is necessary to obtain A10 and A11.

(5.3). <u>Theorem</u>. Under C1 - C3,

$$(5.4) \quad K(s,t) = \begin{cases} G(s)G(t) & ; \ \min(s,t) < \gamma \\ G(\gamma-)(G(s) + G(t) - G(\gamma-)) + \int_0^\infty I_{[\gamma,\min(s,t)]}(u)/h(u)G(du); & \min(s,t) \geq \gamma \end{cases}$$

$$(5.5) \quad \sigma_R^2 = \beta_1(X_k) + \beta_2(X_k) - \alpha^2$$

$$(5.6) \quad \hat{\sigma}_R^2 = \beta_1(\hat{X}_k) + \beta_2(\hat{X}_k)$$

$$(5.7) \quad ES_k(\tau_k) = \int_0^\infty s(t)h(t)G(dt)$$

where

$$\beta_1(Y) = \int_0^\infty \int_0^\infty I_{[0,\gamma)}(s)I_{[0,\gamma)}(t)EY(s)Y(t)G(ds)G(dt)$$

$$+ 2\int_0^\infty \int_0^\infty I_{[0,\gamma)}(s)I_{[\gamma,\infty)}(t)EY(s)Y(t)G(ds)G(dt)$$

$$\beta_2(Y) = \int_0^\infty I_{[\gamma,\infty)}(t)EY^2(t)/h(t)G(dt) \quad .$$

<u>Proof</u>. Relation (5.7) is immediate from the definition of $H_k$. For (5.4), observe that if $s < \gamma \leq t$,

$$EH_k(s)H_k(t) = G(s) \cdot EH_k(t) = G(s)G(t) \quad .$$

If $\gamma \leq s \leq t$, then

$$EH_k(s)H_k(t)$$

$$= E\{(G(\gamma-) + X(M_k)/h(M_k)I_{[0,s]}(M_k))H_k(t)\}$$

$$= G(\gamma-)EH_k(t) + E\{X_k^2(M_k)/h^2(M_k)I_{[0,s]}(M_k)\}$$

$$+ E\{X(M_k)/h(M_k)I_{[0,s]}(M_k)\}G(\gamma-)$$

$$= G(\gamma-)G(t) + \int_0^\infty I_{[\gamma,s]}(u)EX_k^2(u)/h(u)G(du)$$

$$+ G(\gamma-)(G(s) - G(\gamma-))$$

which is (5.4). The expressions for $\sigma_R^2$ and $\hat{\sigma}_R^2$ are obtained from Theorems 3.16 and 3.23 by appropriately integrating against $K(ds,dt)$ as defined by (5.4), and by using the fact that the integrands are symmetric in $(s,t)$. ||

A certain amount of analytical optimization can be performed in our current setting. Specifically, one can determine the function $h$ which minimizes the asymptotic variance of $\hat{R}(t)$ over all point mass estimators in which $\gamma = 0$. Note that if $\gamma = 0$, the asymptotic variance is $\hat{v}_1(h)/t$, where

(5.8) $$\hat{v}_1(h) = \{\int_0^\infty E\hat{X}_k^2(t)/h(t)G(dt)\}(\int_0^\infty s(t)h(t)G(dt)) \quad .$$

(5.9). <u>Proposition</u>. $\hat{v}_1(h) \geq (\int_0^\infty (s(t)E\hat{X}_k^2(t))^{1/2} G(dt))^2$. The minimum is attained by

(5.10) $$h(t) = \alpha(E\hat{X}_k^2(t)/s(t))^{1/2}, \ \alpha > 0 \quad .$$

<u>Proof</u>. Observe that

$$(\int_0^\infty (s(t)E\hat{X}_k^2(t))^{1/2} G(dt))^2$$

$$= (\int_0^\infty (E\hat{X}_k^2(t)/h(t) \cdot s(t)h(t))^{1/2} G(dt))^2$$

$$\leq \int_0^\infty E\hat{X}_k^2(t)/h(t)G(dt) \cdot \int_0^\infty s(t)h(t)G(dt)$$

by the Cauchy-Schwartz inequality. To check that $h(t)$, as defined by (5.10), attains the minimum, it is only necessary to observe that $h(t)$ is undefined only on a set of zero measure with respect to $s(t)G(dt)$; similarly, $1/h(t)$ is undefined on a set of zero measure with respect to $E\hat{X}_k^2(t)G(dt)$. ||

Hence, if $\int_0^\infty (\hat{EX}_k^2(t)/s(t))^{1/2} G(dt) < \infty$, the density $\hat{h}$ defined by

(5.11) $$\hat{h}(t) = (\hat{EX}_k^2(t)/s(t))^{1/2} / \int_0^\infty (\hat{EX}_k^2(t)/s(t))^{1/2} G(dt) \quad .$$

minimizes the asymptotic variance of $\hat{R}(t)$ over all point mass estimators with $\gamma = 0$. Strictly speaking, (5.11) may not be in the class of randomized estimators discussed thus far, since the equality

$$E\{1/\hat{h}(M_k);\ M_k < t\} = G(t)$$

may not hold due to the possibility of $\hat{EX}_k^2(t)$ vanishing on the support of G. If that occurs, set $\tilde{G}(t) = I_A(t)G(dt)/\int_A G(dt)$, where $A = \{t : \hat{EX}_k^2(t) > 0\}$. Note that $\tilde{G}$ satisfies A6 and A11 if and only if G does; furthermore,

$$E\{1/\hat{h}(M_k);\ M_k < t\} = \tilde{G}(t) \quad .$$

$$\alpha = \int_0^\infty EX_k(t)\tilde{G}(dt) \quad .$$

Thus, $\hat{h}$ always gives rise to a randomized estimator if $\tilde{G}$ is allowed to play the role of G.

One interesting property of the optimal density $\hat{h}$ is that (5.11) is independent of G, modulo the normalization constant. Furthermore, although $\hat{h}(t)$ can be estimated via a trial run, that may be undesirable. In such a case, it may be reasonable to assume that $\hat{EX}_k^2(t)$ is approximately constant and s(t) linear, leading to the approximately optimal density $t^{-1/2}/\int_0^\infty t^{-1/2} G(dt)$.

Analytical information about the optimal density $\hat{h}$ is available also in the case that $\gamma$ is positive. The idea is to define the functionals

$$J_1(h) = \hat{\sigma}_R^2(h)\left(\int_0^\infty th(t)G(dt)\right)$$

$$J_2(h) = \int_0^\infty h(t)G(dt) \quad ,$$

where $\hat{\sigma}_R^2(h)$ is given by formula (5.6). We wish to minimize $J_1(h)$ subject to the constraint $J_2(h) = 1$ over functions h having support $[\gamma,\infty)$.

Necessary conditions for a minimizing $h$ may be obtained using variational arguments (see Smith [12]). Formal analysis shows that a minimizing $\hat{h}$ should be of the form

$$\hat{h}(t) = \alpha(E\hat{X}_k^2(t)/(s(t) + d))^{1/2} ,$$

for some constants $\alpha$ and $d$, when $\gamma > 0$.

Before concluding this section, we offer an extended version of the point mass estimator studied thus far. For a density $h$ having support $[\gamma,\infty)$ and satisfying C3, let $\{M_{ik} : i > 1, k > 1\}$ be a family of i.i.d.r.v.'s for which $P\{M_{ik} \varepsilon dt\} = h(t)G(dt)$. Set

$$H_{ik}(t) = \begin{cases} \min\{G(t), G(\gamma-)\}; \ t < M_{ik} \\ G(\gamma-) + 1/h(M_{ik}); \ t > M_{ik} \end{cases} \tag{5.12}$$

and let $H_k(t) = \sum_{i=1}^{n} H_{ik}(t)/n$. In the presence of C1, A6 - A11 then hold. Let $\sigma_R^2(n)$, $\hat{\sigma}_R^2(n)$ be the corresponding values of $\sigma_R^2$, $\hat{\sigma}_R^2$ for the current randomization scheme.

(5.13). <u>Proposition</u>. Under the assumptions of the above paragraph,

$$\sigma_R^2(n) = \sigma_R^2/n + (n-1)\sigma^2/n \tag{5.14}$$

$$\hat{\sigma}_R^2(n) = \hat{\sigma}_R^2/n + (n-1)\sigma^2/n \tag{5.15}$$

where $\sigma_R^2$, $\hat{\sigma}_R^2$ are defined by (5.5) and (5.6).

<u>Proof</u>. Let $R_{ik} = \int_0^\infty X_k(t)H_{ik}(dt)$ and observe that $R_k = \sum_{i=1}^{n} R_{ik}/n$. Then, the exchangeability of $\{R_{ik} : i > 1\}$ proves that

$$ER_k^2 = (\sum_{i=1}^{n} ER_{ik}^2 + \sum_{j\neq i} ER_{ik}R_{jk})/n^2$$

$$= ER_{1k}^2/n + (n-1)ER_{1k}R_{2k}/n \quad .$$

But $R_{1k}$ has the same distribution as the randomized estimator defined through (5.1); hence $ER_{1k}^2 = \sigma_R^2$. To calculate the expectation of the product term, note that

$$
\begin{aligned}
E\{R_{1k}R_{2k} \mid X_k\} &= E\{R_{1k}E\{R_{2k} \mid X_k, H_{1k}\} \mid X_k\} \quad , \\
&= E\{R_{1k}E\{R_{2k} \mid X_k\} \mid X_k\} \qquad (5.16) \\
&= E\{R_{1k} \mid X_k\}E\{R_{2k} \mid X_k\} = I_k^2 \quad ;
\end{aligned}
$$

the second equality follows from the conditional independence of $\sigma(X_k, H_{1k})$ and $\sigma(X_k, H_{2k})$ given $\sigma(X_k)$ ([6], p. 308); the final equality was actually proven during the argument of Theorem 3.5. Relation (5.16) yields $ER_{1k}R_{2k} = EI_k^2$, which proves (5.14); (5.15) is obtained from (5.14) by substituting $\hat{X}$ for $X$. ||

The final ingredient in calculating the asymptotic variance of estimators based on (5.12) is the value of $ES_k(\tau_k)$. But $\tau_k = \max\{M_{ik} : 1 \leq i \leq n\}$, so that

$$P\{\tau_k \leq x\} = (\int_0^\infty I_{[0,x]}(t)h(t)G(dt))^n \quad ;$$

thus, if $s(t)$ is strictly increasing,

$$ES_k(\tau_k) = \int_0^\infty (\int_0^\infty I_{(s^{-1}(x),\infty)}(t)h(t)G(dt))^n dx \quad .$$

It should be clear that point mass randomization is amenable to variance reduction via stratification. Observe that if $0 = t_0 \leq T < t_1 < t_2 < \ldots < t_n = \infty$, then simulating $M_k$ from the conditional distribution $h(u)I_{[t_{i-1},t_i)}(u)G(du)/\int_{t_{i-1}}^{t_i} h(u)G(du)$ and defining $H_k$ via the rule (5.1) yields a deviate from $Q_i$.

## 6. Integrals with Random Endpoints

Assume that $\{(X_k,S_k) : k \geq 1\}$ is a family of processes satisfying A1 - A5 and

D1. $$\int_0^\infty EX_k^2(t)G(dt) < \infty \quad .$$

Let $L$ be a probability distribution function satisfying:

D2. $$L(0-) = 0, \ \sup\{t : L(t) < 1\} = T$$

D3. $$\int_0^\infty G(du)/\bar{L}(u-) < \infty, \quad \text{where} \quad \bar{L}(t) = 1 - L(t) \quad .$$

We define $\{H_k : k \geq 1\}$ by the rule

(6.1) $$H_k(t) = \begin{cases} \int_0^\infty I_{[0,t]}(u)G(du)/\bar{L}(u-); & t < N_k \\ \int_0^\infty I_{[0,N_k]}(u)G(du)/\bar{L}(u-); & t \geq N_k \end{cases}$$

where the sequence $\{N_k : k \geq 1\}$ is a family of i.i.d.r.v.'s, independent of the $(X_k,S_k)$'s, such that $P\{N_k \in dt\} = L(dt)$. With the $H_k$'s defined through (6.1),

$$R_k = \int_0^\infty I_{[0,N_k]}(u)X_k(u)/\bar{L}(u-)G(du) \quad .$$

(6.2). <u>Proposition</u>. Under D1 - D3, A6 - A11 hold.

<u>Proof</u>. It is evident that

$$\begin{aligned} EH_k(t) &= \int_0^\infty \int_0^\infty I_{[0,t]}(u)I_{(t,\infty)}(s)/\bar{L}(u-)G(du)L(ds) \\ &+ \int_0^\infty \int_0^\infty I_{[0,s]}(u)I_{[0,t]}(s)/\bar{L}(u-)G(du)L(ds) \\ &= \int_0^\infty I_{[0,t]}(u)\bar{L}(t)/\bar{L}(u-)G(du) \\ &+ \int_0^\infty I_{[0,t]}(u)(\bar{L}(u-) - \bar{L}(t))/\bar{L}(u-)G(du) \\ &= G(t) \quad ; \end{aligned}$$

the second equality is Fubini's theorem, whereas the third uses the fact that $\bar{L}(t-) > 0$ on $[0,T]$. Hence, $H_k$ satisfies A8. For A6 - A7 and A9 - A11, we apply Proposition 3.32 - it will suffice to prove that $EH_k^3(\infty) < \infty$. But

$$EH_k^3(\infty) = E\{\int_0^\infty I_{[0,N_k]}(u)/\bar{L}(u-)G(du)\}^3$$

$$= \int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty \prod_{i=1}^{3} I_{[0,s]}(u_i)/\bar{L}(u_i-)G(du_i)L(ds)$$

$$= \int_0^\infty\int_0^\infty\int_0^\infty \bar{L}(\max(u_1,u_2,u_3)-) \prod_{i=1}^{3} 1/\bar{L}(u_i-)G(du_i)$$

$$\leq \int_0^\infty\int_0^\infty \prod_{i=1}^{2} 1/\bar{L}(u_i-)G(du_i)$$

$$= (\int_0^\infty 1/L(u-)G(du))^2 < \infty \quad . \qquad ||$$

(6.3). <u>Theorem</u>. Assume D1 - D3. Then,

$$K(ds,dt) = \bar{L}(\max(s,t)-)/\bar{L}(s-)\bar{L}(t-)G(ds)G(dt) \tag{6.4}$$

$$\sigma_R^2 = \int_0^\infty a(X_k;t)/\bar{L}(t-)G(dt) - \alpha^2 \tag{6.5}$$

$$\hat{\sigma}_R^2 = \int_0^\infty a(\hat{X}_k;t)/\bar{L}(t-)G(dt) \tag{6.6}$$

$$ES_k(\tau_k) = \int_0^\infty s(t)L(dt) \tag{6.7}$$

where

$$a(Y;t) = 2\int_0^\infty I_{(t,\infty)}(s)EY(s)Y(t)G(ds) + EY^2(t)(G(t) - G(t-)) \quad .$$

<u>Proof</u>. For $0 \leq s \leq t$,

$$K(s,t) = E\{\int_0^\infty I_{[0,N_k]}(u)I_{[0,s]}(u)/\bar{L}(u-)G(du) \cdot \int_0^\infty I_{[0,N_k]}(v)I_{[0,t]}(v)/\bar{L}(v-)G(dv)\}$$

$$= \int_0^\infty\int_0^\infty\int_0^\infty I_{[0,x]}(u)I_{[0,x]}(v)I_{[0,s]}(u)I_{[0,t]}(v)/\bar{L}(u-)\bar{L}(v-)G(du)G(dv)L(dx)$$

$$= \int_0^\infty\int_0^\infty I_{[0,s]}(u)I_{[0,t]}(v)\bar{L}(\max(u,v)-)/\bar{L}(u-)\bar{L}(v-)G(du)G(dv)$$

proving (6.4). For (6.5), we use the symmetry of $EX_k(s)X_k(t)$ in $(s,t)$ to write

$$ER_k^2 = \int_0^\infty\int_0^\infty EX_k(s)X_k(t)\bar{L}(\max(s,t)-)/\bar{L}(s-)\bar{L}(t-)G(ds)G(dt)$$

$$= 2\int_0^\infty\int_0^\infty I_{(t,\infty]}(s)EX_k(s)X_k(t)/\bar{L}(t-)G(ds)G(dt)$$

$$+ \int_0^\infty\int_0^\infty I_{\{t\}}(s)EX_k^2(t)/\bar{L}(t-)G(ds)G(dt)$$

$$= \int_0^\infty a(X_k;t)/\bar{L}(t-)G(dt)$$

proving (6.5); for (6.6), apply (6.5) to $\hat{X}_k$. Since $L(\infty) = 1$, $\tau_k < \infty$ a.s. so

$$ES_k(\tau_k) = \int_0^\infty s(t)F(dt) = \int_0^\infty s(t)P\{\tau_k \in dt\}$$

$$= \int_0^\infty s(t)L(dt) \quad . \qquad ||$$

As in Section 5, one can try to analytically optimize the choice of L. We shall require several additional assumptions.

D4. $s(t) < \infty$ for $t > 0$

D5. there exists a non-negative function $u(t)$ such that $a(\hat{X}_k;t)/u(t)G(dt) = s(dt)$ (by D4 and A3, $s(t)$ is a finite non-decreasing function).

The asymptotic variance associated with $\hat{R}(t)$, for a fixed L, is given by $\hat{v}_2(L)/t$ where

$$\hat{v}_2(L) = (\int_0^\infty a(\hat{X}_k;t)/\bar{L}(t-)G(dt))(\int_0^\infty s(t)L(dt)) \quad .$$

(6.8). Proposition. Under D1 - D5,

$$\hat{v}_2(L) > (\int_0^\infty a(\hat{X}_k;t)/u^{1/2}(t)G(dt))^2 \quad .$$

Proof. Observe that

$$(\int_0^\infty a(\hat{X}_k;t)/u^{1/2}(t)G(dt))^2$$

$$= (\int_0^\infty (a(\hat{X}_k;t)/\bar{L}(t-))^{1/2}(\bar{L}(t-)a(\hat{X}_k;t)/u^{1/2}(t)G(dt))^2$$

$$< \int_0^\infty a(\hat{X}_k;t)/\bar{L}(t-)G(dt) \cdot \int_0^\infty \bar{L}(t-)a(\hat{X}_k;t)/u(t)G(dt)$$

by Cauchy-Schwartz. But

$$\int_0^\infty \bar{L}(t-)a(\hat{X}_k;t)/u(t)G(dt) = \int_0^\infty\int_0^\infty I_{[t,\infty)}(x)a(\hat{X}_k;t)/u(t)L(dx)G(dt)$$

$$= \int_0^\infty\int_0^\infty I_{[0,x]}(t)a(\hat{X}_k;t)/u(t)G(dt)L(dx)$$

$$= \int_0^\infty s(x)L(dx)$$

where the second equality is by Fubini, and the final equality uses D5. ||

Assumption D5 requires absolute continuity of $s(dt)$ with respect to $G(dt)$ on $[0,T]$ and non-negativity of $a(\hat{X}_k;t)$. The non-negativity of $a(\hat{X}_k;t)$ clearly arises when $\hat{X}_k$ has non-negative correlations in time. However it also holds under other conditions.

D6. $\{\hat{X}_k(t) : t \geq 0\}$ is a stationary process.

(6.9). Proposition. Suppose D1 and D6 hold for $G(t) = 1 - \exp(-\lambda t)$, $\lambda > 0$. Then $a(\hat{X}_k;t)$ is non-negative and decreasing to zero.

Proof. By stationarity,

$$a(\hat{X}_k;t) = 2\int_t^\infty E\hat{X}_k(0)\hat{X}_k(s-t)\lambda e^{-\lambda s}ds$$

(6.10)

$$= 2\lambda\int_0^\infty E\hat{X}_k(0)\hat{X}_k(u)e^{-\lambda u}du \cdot e^{-\lambda t} \quad .$$

By D1 and Cauchy-Schwartz,

$$E(\lambda\int_0^\infty|\hat{X}_k(s)|e^{-\lambda s}ds)^2 \leq \int_0^\infty E\hat{X}_k(s)^2\lambda e^{-\lambda s}ds < \infty \quad .$$

Hence, Fubini proves that

$$0 \leq E(\int_0^\infty \hat{X}_k(s)\lambda e^{-\lambda s}ds)^2 = \int_0^\infty\int_0^\infty E\hat{X}_k(s)\hat{X}_k(t)\lambda^2 e^{-\lambda(s+t)}dsdt$$

$$= 2\int_0^\infty\int_t^\infty E\hat{X}_k(0)\hat{X}_k(s-t)\lambda e^{-\lambda s}ds\,\lambda e^{-\lambda t}dt$$

$$= \lambda \cdot \int_0^\infty E\hat{X}_k(0)\hat{X}_k(u)e^{-\lambda u}du \quad ,$$

from which it follows, using (6.10), that $a(\hat{X}_k;t)$ is non-negative and decreasing to zero. ||

Assume that $\bar{G}(t) = e^{-\lambda t}$ and that $s(t) = st$. Proposition 6.9 can be used to find a distribution $L$ which is approximately optimal, provided that $\hat{X}_k$ is reasonably close to stationarity. Substituting the expression for $a(\hat{X}_k;t)$ derived in the proof of Proposition 6.9 into the defining relation for $u(t)$, one gets

$$(2\lambda \int_0^\infty E\hat{X}_k(0)\hat{X}_k(t)e^{-\lambda u}du) \cdot e^{-2\lambda t}dt = u(t)s\, dt \quad ;$$

hence $u(t) = \beta e^{-2\lambda t}$ for some $\beta$. The following property holds in our current setting.

D7. $u(t)$ is a continuous function which decreases to zero as $t \to \infty$.

(6.11). <u>Proposition</u>. Under D1 - D5 and D7, $L_1(t) = 1 - (u(t)/u(0))^{1/2}$ is a distribution function which minimizes $\hat{v}_2(L)$.

<u>Proof</u>. The continuity of $u$ implies that $L$ has the required right continuity, so that $L$ is a distribution function. From Proposition 6.8,

$$\hat{v}_2(L) \geq (\int_0^\infty a(\hat{X}_k;t)/u^{1/2}(t)G(dt))^2$$

$$= (\int_0^\infty a(\hat{X}_k;t)/\bar{L}_1(t)G(dt))(\int_0^\infty a(\hat{X}_k;t)\bar{L}_1(t)/u(t)G(dt))$$

$$= (\int_0^\infty a(\hat{X}_k;t)/\bar{L}_1(t)G(dt))(\int_0^\infty\int_0^\infty I_{[0,u]}(t)a(\hat{X}_k;t)/u(t)G(dt)L_1(du))$$

$$= (\int_0^\infty a(\hat{X}_k;t)/\bar{L}_1(t)G(dt))(\int_0^\infty s(u)L_1(du)) = \hat{v}_2(L_1) \quad ,$$

where the second equality is by Fubini, and the third follows from the defining relation for $u$. ||

As a consequence of Proposition 6.11, we see that $L(dt) = \lambda e^{-\lambda t}dt$ is approximately optimal for a simulation problem in which $\hat{X}$ is approximately stationary and $G(dt) = \lambda e^{-\lambda t}dt$.

We now turn to an application of antithetic variance reduction to integral estimators with random endpoints. Let $L^{-1}(x) \triangleq \inf\{u : L(u) \geq x\}$ and assume that there exists a sequence of i.i.d.r.v.'s $\{U_k : k \geq 1\}$, independent of $\{(X_k,S_k) : k \geq 1\}$. Set $N_{k1} = L^{-1}(U_k)$, $N_{k2} = L^{-1}(1 - U_{k+1})$ for $k$ odd; $N_{k1} = L^{-1}(1 - U_{k-1})$, $N_{k2} = L^{-1}(U_k)$ for $k$ even. Let

$$R_{ki} = \int_0^{N_{ki}} X_k(s)/\bar{L}(s-)G(ds)$$

for $i = 1,2$; the pairs $(R_{k1}, R_{k+1,1})$ and $(R_{k2}, R_{k+1,2})$ ($k$ odd) are said to be antithetic. Let $R_1(n) = \sum_{k=1}^{n} R_{k1}/n$. The following theorem involves a proof similar to that of Theorem 3.19.

(6.12). Theorem. Assume D1 - D4. If $\int_0^\infty s(t)L(dt) < \infty$, then

(6.13) $$t^{1/2}(R_1(M_1(t)) - \alpha) \Rightarrow \sigma^2(R_{11} + R_{21})\int_0^\infty s(t)L(dt)/2$$

as $t \to \infty$, where $M_1(t) = \max\{2n : N_{11} + \ldots + N_{2n,1} \leq t\}$.

Based on Theorem 6.3, the antithetic estimator $R_1(M_1(t))$ is more efficient than $R(t)$ if $\sigma^2(R_{11} + R_{21})/2 \leq \sigma_R^2 = \sigma^2(R_{11})$.

D8. $X_k(t) \geq 0$ a.s.

(6.14). Proposition. Assume D1 - D4 and D8. Then, $\sigma^2(R_{11} + R_{21}) \leq 2\sigma_R^2$.

Proof. Clearly, it is sufficient to prove that $\mathrm{cov}(R_{11}, R_{21}) \leq 0$. By D8,

$$R_{11} - R_{22} = \int_0^{L^{-1}(u_1)} X_1(s)/\bar{L}(s-)G(ds) - \int_0^{L^{-1}(u_2)} X_1(s)/\bar{L}(s-)G(ds)$$

has the same sign as

$$R_{12} - R_{21} = \int_0^{L^{-1}(1-u_2)} X_2(s)/\bar{L}(s-)G(ds) - \int_0^{L^{-1}(1-u_1)} X_2(s)/\bar{L}(s-)G(ds) \quad .$$

Hence,

$$0 \leq E(R_{11} - R_{22})(R_{12} - R_{21}) = -2\,\mathrm{cov}(R_{11}, R_{21}) \quad . \quad ||$$

Variance reduction for integrals with random endpoints can also be accomplished via stratification. Note that if $0 = t_0 < t_1 < \ldots < t_n = \infty$, the simulation of $N_k$ from the conditional distribution $L(du)/(L(t_i-) - L(t_{i-1}-))$ $(t_{i-1} \leq u < t_i)$ and definition of $H_k^S$ via (6.1) yields a deviate from $Q_i$.

References:


[1] Billingsley, P. (1968). Convergence of Probability Measures. Wiley, New York.

[2] Billingsley, P. (1979). Probability and Measure. Wiley, New York.

[3] Bratley, P., Fox, B. L., and Schrage, L. E. (1983). A Guide to Simulation. Springer-Verlag, New York.

[4] Burrill, C. W. (1972). Measure, Integration, and Probability. McGraw-Hill, New York.

[5] Chung, K. L. (1974). A Course in Probability Theory. Academic Press, New York.

[6] Çinlar, E. (1975). Introduction to Stochastic Processes. Prentice-Hall, Englewood Cliffs, NJ.

[7] Dellacherie, C. and Meyer, P. A. (1978). Probabilities and Potential. North-Holland, New York.

[8] Glynn, P. W. (1983). On confidence intervals for cyclic regenerative processes. Operations Res. Lett. 2.

[9] Grassman, W. (1982). Simulating transient solutions of Markovian systems. Technical Report, Dept. of Computer Science, University of Saskatchewan, Saskatoon.

[10] Law, A. M. and Kelton, W. D. (1982). Simulation Modelling and Analysis. McGraw-Hill, New York.

[11] Rubinstein, R. Y. (1981). Simulation and the Monte Carlo Method. Wiley, New York.

[12] Smith, D. R. (1974). Variational Methods in Optimization. Prentice-Hall, Englewood Cliffs, NJ.


PWG/jvs