AD-A139 315   ESTIMATING TIME AVERAGES VIA RANDOMLY-SPACED   1/1
              OBSERVATIONS(U) WISCONSIN UNIV-MADISON MATHEMATICS
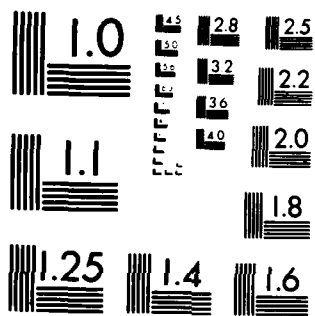   RESEARCH CENTER B L FOX ET AL. JAN 84 MRC-TSR-2638
UNCLASSIFIED DAAG29-80-C-0041   F/G 12/1   NL

END
DATE
FILMED
4 -84
DTIC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #2638

ESTIMATING TIME AVERAGES VIA
RANDOMLY-SPACED OBSERVATIONS

Bennett L. Fox and Peter W. Glynn

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

January 1984

(Received October 14, 1983)

DTIC
SELECTED
MAR 2 2 1984
B

84

DTIC FILE COPY

AD A139315

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

ESTIMATING TIME AVERAGES VIA RANDOMLY-SPACED OBSERVATIONS

Bennett L. Fox[*,1] and Peter W. Glynn[2]

Technical Summary Report #2638
January 1984

ABSTRACT

To estimate continuous-time averages via randomly-spaced observations of discrete-event systems, we develop a point-process framework and use it to generalize both regenerative and stationary-process oriented simulation methodologies. We give consistent estimators, central limit theorems, and an effective bias-reducing jackknife. The impact on indirect estimation of transaction (customer) averages is discussed.

AMS (MOS) Subject Classifications: 62M15, 65C05, 68J05, 60G55

Key Words: Simulation, point processes

Work Unit Number 5 (Optimization and Large Scale Systems)

*Département d'informatique et de recherche opérationnelle, Université de Montréal, Montréal, Québec H3C 3J7

SIGNIFICANCE AND EXPLANATION

In many stochastic systems, one is interested in estimating steady-state expected values. When Monte Carlo simulation is used to estimate such parameters, an assessment of accuracy, in the form of confidence intervals, is often required. Most procedures for producing such confidence intervals require that the simulation be sampled so that the time increments between observations are all equal. This is difficult to accomplish in a discrete-event simulation, since the clock which drives the simulation is incremented in a random fashion. Our purpose, in this paper, is to show how methods for dealing with equally spaced observations can be adapted to run on the random time scale of the driving clock for the simulation.

Accession For

NTIS GRA&I	✓	
DTIC TAB	☐	
Unannounced	☐	
Justification		

By_____
Distribution/
Availability Codes

Dist	Avail and/or Special
A-1	

ESTIMATING TIME AVERAGES VIA RANDOMLY-SPACED OBSERVATIONS

Bennett L. Fox[*,1] and Peter W. Glynn[2]

1. INTRODUCTION.

Let $0 = T_0 < T_1 < \ldots$ be <u>event times</u>. Though the associated sequence $\tilde{X}_1, \tilde{X}_2, \ldots$ with $\tilde{X}_k \in \mathbb{R}^d$ is not necessarily an imbedded Markov chain, we call $\tilde{X}_k$ the <u>state</u> at time T_k - somewhat abusing the term. To define the state $X(t)$ at an arbitrary time t, interpolate:

$$(1.1) \qquad X(t) = \sum_{k=0}^{\infty} \tilde{X}_k I_{[T_k, T_{k+1})}(t)$$

where the indicator I_A is 1 or 0 depending on whether or not $t \in A$. For this definition to make sense, every state change must correspond to an event time. The state does not change continuously. It jumps at discrete (possibly random) times. In other words, we have a <u>discrete-event system</u>. Let f be a real-valued function. Put

$$(1.2) \qquad r(t) = \frac{1}{t} \int_0^t f(X(s))\,ds .$$

We solve the <u>steady-state problem</u>: estimate the limit (when it exists)

$$(1.3) \qquad r = \lim_{t \to \infty} r(t)$$

and construct confidence intervals for r.

To do this, we develop a point-process framework and use it to generalize both regenerative and stationary-process oriented simulation methodologies. Simply averaging the $\tilde{X}_k$'s generally inconsistently estimates r. The T_k's are not necessarily regeneration times. We work with generally dependent observations, in contrast to regenerative approaches.

---
[*]Département d'informatique et de recherche opérationnelle, Université de Montréal, Montréal, Québec H3C 3J7

__1.1 Shifts.__ We use the set of functions $\omega : [0,\infty) \rightarrow \mathbb{R}^d$ that are right continuous and
have left limits to describe the sample space Ω. Define X via

(1.4) $$X(\cdot) = X(\cdot,\omega) = \omega(\cdot)$$

for $\omega \in \Omega$. For any random variable $R : \Omega \rightarrow [0,\infty)$, define a __right shift__ via

(1.5) $$\theta_R = \theta_{R(\omega)}(\omega) = \omega(R(\omega) + \cdot) .$$

Let $A \subset \mathbb{R}^d$, put $S_0 = 0$, and set

(1.6) $$S_{k+1} = \inf\{t > S_k : X(t-) \neq X(t), X(t) \in A\}$$

where $X(t-)$ is the left limit of X at t. If $A = \mathbb{R}^d$, then $S_1 = T_1$. Define

(1.7) $$P_n\{X \in \cdot\} = P\{X \cdot \theta_{S_n} \in \cdot\}$$

and

(1.8) $$P_t\{X \in \cdot\} = P\{X \cdot \theta_t \in \cdot\}$$

For most simulations where the steady-state limit r exists, limit probabilities $\hat{P}$ and
$\tilde{P}$ exist such that weak convergence holds:

(1.9) $$P_n \Longrightarrow \hat{P}$$

and

(1.10) $$P_t \Longrightarrow \tilde{P}$$

where

(1.11) $$\hat{P}\{X \cdot \theta_{S_1} \in \cdot\} = \hat{P}\{X \in \cdot\}$$

and (__shift invariance__)

(1.12) $$\tilde{P}\{X \cdot \theta_t \in \cdot\} = \tilde{P}\{X \in \cdot\}$$

for $t > 0$.

   __Example 1.1.__ Let X be a delayed (resp., non-delayed) regenerative process under
P (resp., $\hat{P}$). Assume that X regenerates when it hits A. Then $P_n = \hat{P}$ for $n > 1$
though generally $P_0 \neq \hat{P}$.

   __Example 1.2.__ Let X be as in example 1.1. If the regenerative process is positive
recurrent and the regeneration-spacing distribution has a non-trivial Lebesgue component,
then (1.10) and (1.12) hold (see Miller (1972)). Also, $\tilde{P}$ and $\hat{P}$ are related by

(1.13)
$$\widetilde{P}\{X \in \cdot\} = \frac{1}{\hat{E}S_1} \int_0^\infty \hat{P}\{X \circ \theta_t \in \cdot; S_1 > t\}dt$$

where $\hat{E}$ denotes expectation under $\hat{P}$. In Section 3 we show that (1.13) holds more generally.

A process X satisfying

(1.14)
$$P\{X \circ \theta_{S_1} \in \cdot\} = P\{X \in \cdot\}$$

is _synchronous_ with respect to the _imbedded point process_ sequence $\{S_n\}$.

Example 1.3. Suppose that $\{\widetilde{X}_n\}$ is a stationary sequence with $X_n = T_n = n$. Then $X(t)$ is synchronous with respect to $\{S_n\}$.

From examples 1.1 and 1.3 we see that synchronous processes generalize both non-delayed regenerative processes and stationary sequences. Assume that the simulator (somehow) chooses the time origin so that everything representing the initial "transient" phase is to its left. This is trivial for regenerative processes but not in general. This deletion assumption translates mathematically as: (1.14) holds.

Let

(1.15)
$$\alpha_n = S_{n+1} - S_n$$

and

(1.16)
$$X_n(t) = \begin{cases} X(S_n + t); & 0 < t < \alpha_n \\ \infty & ; \quad t > \alpha_n . \end{cases}$$

We assume that

(1.17) $\{X_n\}$ is a ϕ-mixing sequence (Billingsley (1968),
pp. 166-168) with ϕ-mixing coefficients satisfying

$$\sum_{k=1}^\infty \phi_k^{1/2} < \infty$$

and

(1.18)
$$E(Y_n(|f|)^2 + \alpha_n^2) < \infty$$

where

-3-

$$(1.19) \qquad Y_n(f) = \int_0^{\alpha_n} f(X_n(t))dt$$

$$= \int_{S_n}^{S_{n+1}} f(X(t))dt \ .$$

The _cycle_ sequence $\{X_n\}$ is not necessarily iid. If it were, we could simply apply regenerative methodology. Three more definitions set the stage:

$$(1.20) \qquad \bar{Y}_n(f) = \sum_{k=0}^{n-1} Y_k(f)/n$$

$$(1.21) \qquad \bar{\alpha}_n = S_n/n$$

$$(1.22) \qquad r_n = \bar{Y}_n(f)/\bar{\alpha}_n \ .$$

1.2. Preview of results. In Section 2 we assume that (1.14), (1.17), and (1.18) hold, usually without further explicit mention. We show that

$$(1.23) \qquad r_n \to r$$

$$(1.24) \qquad r(t) \to r$$

$$(1.25) \qquad r = EY_0(f)/E\alpha_0$$

and generalize the "inspection paradox"

$$(1.26) \qquad \frac{1}{t} \int_0^t I_{[x,\infty)}(\alpha_{N(s)})ds \to \int_x^\infty P\{\alpha_0 > s\}ds/E\alpha_0 \ .$$

The left side of (1.26) is the proportion of time over $[0,t]$ that the cycle in progress had length at least x. Its limit, the right side of (1.26), is precisely that for the regenerative case; e.g., see Bratley, Fox, and Schrage [(1983), problem 3.7.4] or Heyman and Sobel [(1982), §5.5]. Thus, for synchronous processes the cycle in progress at time t tends to be longer than average, thereby biasing $r(t)$.

Next, Section 2 proves a central limit theorem (CLT):

$$(1.27) \qquad n^{1/2}(r_n - r) \Longrightarrow \sigma N(0,1)/E\alpha_0$$

$$(1.28) \qquad t^{1/2}(r(t) - r) \Longrightarrow \sigma N(0,1)/(E\alpha_0)^{1/2}$$

-4-

where $N(0,1)$ is a zero-mean, unit-variance, normal random variable,

$$(1.29) \qquad \sigma^2 = EY_0(\hat{f})^2 + 2 \sum_{k=1}^{\infty} EY_0(\hat{f})Y_k(\hat{f})$$

and

$$(1.30) \qquad \hat{f}(x) = f(x) - r .$$

In other words, $Y_k(\hat{f}) = Y_k(f) - \alpha_k r$. If the synchronous process X is either regenerative or stationary with $S_n = n$, the CLT simplifies. For the former, the second term in (1.29) – corresponding to covariances – vanishes. For the latter, $\alpha_n \equiv 1$. Also

$(1.31) \qquad$ when $A = R^d$,

$$\sigma^2 = E\hat{f}^2(\tilde{X}_0)\tau_0^2 + 2 \sum_{k=1}^{\infty} \hat{f}(\tilde{X}_0)\hat{f}(\tilde{X}_k)\tau_0\tau_k$$

where

$$\tau_k = \tau_{k+1} - \tau_k .$$

To estimate σ^2 in the regenerative case, one uses the knowledge that the covariance terms in (1.29) vanish, to construct the estimator σ_n^2:

$$\sigma_n^2 = \frac{1}{n-1} \sum_{k=1}^{n} (Y_k(f) - r_n\tau_k)^2 ,$$

the estimator σ_n^2 is easily shown to be strongly consistent for σ^2.

In the general case, estimation of σ^2 is more complicated. The parameter σ^2 can be expressed in the form

$$(1.32) \qquad \sigma^2 = \left[\operatorname{var} Y_0(f) + 2 \sum_{k=1}^{\infty} \operatorname{cov}(Y_0(f),Y_k(f)) \right]$$

$$- r\left[\operatorname{cov}(Y_0(f),\alpha_0) + 2 \sum_{k=1}^{\infty} \operatorname{cov}(Y_0(f),\alpha_k) \right]$$

$$- r\left[\operatorname{cov}(\alpha_0,Y_0(f)) + 2 \sum_{k=1}^{\infty} \operatorname{cov}(\alpha_0,Y_k(f)) \right]$$

$$+ r^2\left[\operatorname{var} \alpha_0 + 2 \sum_{k=1}^{\infty} \operatorname{cov}(\alpha_0,\alpha_k) \right]$$

the four bracketed terms appearing in (1.32) can be consistently estimated by $c_{1n}, c_{2n}, c_{3n}, c_{4n}$ (say); this can be accomplished by standard techniques (e.g. batch means, spectral methods, autoregressive procedures). Such methods include parameters (e g. batch size, spectral window, autoregressive order) which must be keyed to sample size. In any case, the estimator σ_n^2 is then given by

$$\sigma_n^2 = c_{1n} - r_n c_{2n} - r_n c_{3n} + r_n^2 c_{4n} .$$

When using $r(t)$, replace n by

(1.33) $\qquad\qquad N(t) = \# \ S_i\text{'s observed in } (0,t]$

Glynn and Iglehart (1981), Jow [(1982), pp. 54-56], and Streller [(1980), Theorem 3.2] prove essentially (1.27) but under significantly different hypotheses.

From (1.27) and (1.28) we get the respective confidence intervals

$$[r_n - z_\delta \sigma_n / \bar{\alpha}_n \sqrt{n}, \ r_n + z_\delta \sigma_n / \bar{\alpha}_n \sqrt{n}]$$

and

$$\left[r(t) - z_\delta \sigma_{N(t)} / (t\bar{\alpha}_{N(t)})^{1/2}, \ r(t) + z_\delta \sigma_{N(t)} / (t\bar{\alpha}_{N(t)})^{1/2} \right]$$

where the percentile z_δ is the unique solution of $P[N(0,1) \leq z_\delta] = 1 - \delta/2$. If σ_n is a strongly consistent estimator for σ, then these are asymptotically exact $100(1 - \delta)\%$ confidence intervals for r.

Under an additional assumption

(1.34 $\qquad\qquad$ there exists $K > 0$ such that $P\{\alpha_0 \leq K\} = 1$

$\qquad\qquad\qquad$ and $\sup\{|f(x)| : x \in R^d\} \leq K$

Section 2 proves that

(1.35) $\qquad\qquad Er_n = r - \beta/(E\alpha_0)^2 n + 0(1/n)$

and finds an expression for β. The form of (1.35) motivates a jackknife:

(1.36) $\qquad\qquad \tilde{r}_{2n} = 2r_{2n} - (r(0,n - 1) + r(n,2n - 1))/2$

where

(1.37) $\qquad\qquad r(a,b) = \left(\sum_{j=a}^{b} Y_j(f) \right) / \left(\sum_{j=a}^{b} \alpha_j \right)$

and

(1.38)
$$\tilde{E r}_{2n} = r + 0(1/n)$$

reducing bias by an order of magnitude.

Sometimes there is no bias. If $S_n = T_n = n$, then $Er_n = r$.

For another case, let $A = R^d$ and suppose that the T_k's see time averages. For the latter, suppose that

(i) $N(t)$ is a stationary Poisson process

(ii) for each $t > 0$, $\{X(s-) : 0 < s < t\}$ is independent of

$\{N(t + u) - N(t) : u > 0\}$.

Wolff (1982) shows that, if $r(t) \to r$ a.s., then

(1.39)
$$\hat{r}_n = \frac{1}{n} \sum_{k=0}^{n-1} f(\tilde{X}_k) \to r \qquad a.s.$$

Since the summands are identically distributed, $\hat{Er}_n = r$. Section 2 concludes by proving that

(1.40)
$$\sqrt{2n} \ (\tilde{r}_{2n} - r) \implies \sigma N(0,1)/E\alpha_0 .$$

Combining (1.27) and (1.40), we see that $\tilde{r}_{2n}$ and r_{2n} have the same asymptotic variance. So our jackknife reduces bias without increasing variance.

Section 3 begins by proving that

(1.41)
$$\frac{1}{t} \int_0^t \hat{P}\{X \cdot \theta_s \in \cdot\} ds \to \tilde{P}\{X \in \cdot\}$$

where $\tilde{P}$ satisfies (1.13), still assuming that (1.14), (1.17), and (1.18) hold. The point is that (1.13) remains valid when X is (merely) synchronous. Franken et al. [(1982), Chapter 1] show that it holds under even weaker conditions.

When X is regenerative, $\hat{P}\{X \in \cdot\} = \tilde{P}\{X \cdot \theta_{S_1} \in \cdot\}$, "inverting" (1.13). This simple inversion generally fails, assuming only (1.9)-(1.12). Recall that under $\tilde{P}$, the cycle trapping a fixed time tends to be longer than a typical cycle. In the regenerative case, only the actual trapping cycle is affected, due to independence of cycles. In general, however, neighboring cycles are also affected due to cycle correlation. Franken

-7-

et al. [(1982), p. 23] show that the general inverse to (1.13) is

$$
(1.42) \qquad \hat{P}\{X \in \cdot\} = \frac{1}{\lambda} \sum_{k=1}^{\infty} \tilde{P}\{X \cdot \theta_{S_k} \in \cdot; \, S_k < 1\}
$$

where $\lambda = \tilde{E}N(1)$ and $\tilde{E}$ denotes expectation under $\tilde{P}$. Comparing (1.11) and (1.14), we see that X is synchronous under $\hat{P}$, the <u>Palm distribution</u> of $\tilde{P}$. Here λ is the <u>intensity</u> of $\tilde{P}$, and X is <u>stationary</u> under $\tilde{P}$.

Chapter 1 of Franken et al. (1982) thoroughly discusses the relationships between $\hat{P}$ and $\tilde{P}$ and proves an intuitive alternative to (1.42):

$$
(1.43) \qquad \hat{P}\{X \in \cdot\} = \lim_{h \downarrow 0} \tilde{P}\{X \cdot \theta_{S_1} \in \cdot \,|\, S_1 < h\} \, .
$$

This shows that the Palm distribution $\hat{P}$ is the stationary dist-   ∈₁n $\tilde{P}$ conditioned on hitting A at time 0.

We want (1.12) to hold with P replacing $\tilde{P}$. This means ' the simulator (somehow) deletes the entire transient phase, choosing the time origin so ₁⊕ ₁nis phase is to its left. Typically this is impossible to do exactly in practice, but to proceed *mathematically we assume it has been done exactly.* This translates as

$$
(1.44) \qquad P\{X \cdot \theta_t \in \cdot\} = P\{X \in \cdot\}
$$

holds. This is stronger than (1.14) because $Er(t) = r$ under (1.44) but not generally under (1.14). In fact (1.44) usually holds literally only if the initial state is generated by the (generally unknown) stationary distribution $\tilde{P}$. Even for regenerative processes, where synchronization via (1.14) is trivial, making (1.44) hold even to a first approximation is generally hard. Nevertheless, in the rest of this section and in Section 3, usually without further explicit mention, we assume that (1.44) holds and that, under $\hat{P}$ but not necessarily under P, (1.14), (1.17), and (1.18) hold. By contrast, Section 2 assumes that (1.14), (1.17), and (1.18) hold under P.

Several counterparts to results in Section 2 are proved in Section 3. There we show that, under $\tilde{P}$, $r(t)$ is (still) strongly consistent and

$$
(1.45) \qquad \sqrt{t} \, (r(t) - r) \Longrightarrow \sigma N(0,1)/(\hat{E}\alpha_0)^{1/2}
$$

where

$$(1.46) \qquad \sigma^2 = \hat{E} Y_0(\hat{f})^2 + 2 \sum_{k=1}^{\infty} \hat{E} Y_0(\hat{f}) Y_k(\hat{f}) \ .$$

This CLT corresponds to (1.28) with $\hat{E}$ (expectation under $\hat{P}$) replacing E. Estimate σ^2 and construct confidence intervals just as before.

The stationarity assumptions (1.14) and (1.44) can be significantly relaxed. Our CLT's can be extended to include certain nonstationary processes by appealing to Billingsley [(1968), Theorem 20.2]. Furthermore, our discussion carries over to the steady-state estimation problem for

$$r = \lim_{t \to \infty} \frac{1}{t} \sum_{k=0}^{\infty} f(X(U_k)) I_{[0, U_k]}(t)$$

(when the limit exists), where the U_k's are an increasing sequence of random times. Such limits are of interest, for example, in queues where lump-sum rewards are payed out to the server at customer departure epochs. Our arguments go through provided that one modifies the definition of $Y_n(f)$ to

$$Y_n(f) = \sum_{j=0}^{\infty} f(X(U_j)) I_{[S_n, S_{n+1})}(U_j) \ .$$

1.4 Transactions. To essentially every time average r there corresponds a transaction (customer) average s and conversely. Heyman and Stidham (1980) and Heyman and Sobel [(1982), §11.3] establish this correspondence explicitly. Thus, every estimator of r yields an _indirect estimator_ of s and conversely. In our framework, Section 4 indicates that on balance estimating s indirectly is often better than estimating it directly. This conclusion appears contrary to some folklore.

Readers wishing to skip to Section 4 can do so without loss of continuity.

2. RESULTS AND PROOFS; I.

Throughout this section we assume that (1.14), (1.17), and (1.18) hold under P.

THEOREM 1. Formulas (1.23), (1.24), and (1.25) hold a.s.

Proof. First, observe that $(Y_n(f), \alpha_n)$ is trivially a functional of $X_n(\cdot)$ and thus $\{(Y_n(f), \alpha_n) : n \geq 0\}$ is ϕ-mixing with the same mixing coefficients as the X_n's. Since any ϕ-mixing sequence is ergodic (Lamperti (1977), pp. 95-96), apply Birkhoff's ergodic theorem (Heyman and Sobel (1982), p. 366) to conclude that

$$(2.1) \qquad \bar{Y}_n(f) \to EY_0(f) \qquad a.s.$$

$$(2.2) \qquad \bar{\alpha}_n \to E\alpha_0 \qquad a.s.$$

proving (1.23) with r given by (1.25). With $N(t)$ defined by (1.33),

$$(2.3) \qquad S_{N(t)}/N(t) \leq t/N(t) \leq S_{N(t)+1}/N(t)$$

from the definitions. From $N(t) \uparrow \infty$ a.s. and (2.2), the extreme terms of (2.3) converge to $E\alpha_0$ a.s.; as $t/N(t)$ gets squeezed, it converges to $E\alpha_0$ a.s. For nonnegative f,

$$(2.4) \qquad \frac{1}{t}\left(\frac{N(t)}{N(t)+1}\right)\bar{Y}_{N(t)}(f) \leq r(t) \leq \frac{1}{t}\left(\frac{N(t)+1}{N(t)}\right)\bar{Y}_{N(t)+1}(f).$$

Apply $N(t) \uparrow \infty$ a.s., (2.1), and $N(t)/t \to 1/E\alpha_0$ a.s. to quickly see that the extreme terms of (2.4) converge to r given by (1.25), proving (1.24) for $f \geq 0$. Split a general f into its positive and negative parts, apply (1.24) to each, and recombine. To justify the last step, use $(E|Y_1(f)|)^2 \leq EY_1(f)^2 < \infty$ by (1.18).

THEOREM 2. Formula (1.26) holds.

Proof. Observe that

$$(2.5) \qquad \frac{1}{n}\int_0^{S_{n+1}} I_{[x,\infty)}(\alpha_{N(s)})\,ds = \frac{1}{n}\sum_{k=0}^{n}\alpha_k I_{[x,\infty)}(\alpha_k) + E\{\alpha_0 ; \alpha_0 > x\} \qquad a.s.$$

-10-

$(E\{X;A\} \triangleq EXI_A$ where I_A is 1 or 0 depending on whether or not A occurs) and then mimic the proof of Theorem 1.

THEOREM 3. Formulas (1.27) and `.28) hold.

Proof. The sequence $\{Y_k(\hat{f})\}$ is ϕ-mixing with zero mean. It satisfies the conditions of Billingsley (1968), Theorem 20.1, and hence

$$(2.6) \qquad \sqrt{n}\, \bar{Y}_n(f) \implies \sigma N(0,1)$$

as $n \to \infty$; for $\sigma = 0$, see Billingsley (1968), p. 177. Rewrite (2.6) as

$$(2.7) \qquad \sqrt{n}\, \bar{a}_n(r_n - r) \implies \alpha N(0,1) .$$

Now (1.27) follows from (2.7 , (2.2), and a routine application of the converging-together lemma (Billingsley (1968), p. 25).

For (1.28), let $\{t_n\}$ be an arbitrary sequence converging to ∞. Apply Billingsley's [(1968), p. 146] random time change theorem to the weak invariance principle version of (2.6) to get

$$(2.8) \qquad \sqrt{N(t_k)}\, \bar{Y}_{N(t_k)}(\hat{f}) \implies \sigma N(0,1)$$

as $k \to \infty$. Another application of the converging-together lemma then yields

$$(2.9) \qquad \sqrt{t_k}\, (r_{N(t_k)} - r) \implies \sigma N(0,1)/\sqrt{Ea_0}$$

using $N(t)/t \to Ea_0$ proved just after (2.3) and $\sqrt{N(t_k)} = \sqrt{t_k}\,\sqrt{N(t_k)/t_k}$. Clearly

$$(2.10) \qquad \sqrt{t_k}\, |r_{N(t_k)} - r(t_k)| \leqslant \sqrt{t_k}\, Y_{N(t_k)}(|f|)/S_{N(t_k)} .$$

Combining a standard Borel-Cantelli argument and $EY_k(|f|)^2 < \infty$ proves that

$$(2.11) \qquad Y_k(|f|)/\sqrt{k} \to 0 \qquad \text{a.s.}$$

But the right-hand side of (2.10) can be rewritten as

$$(t_k/N(t_k))^{1/2} \cdot (Y_{N(t_k)}(|f|)/\sqrt{N(t_k)})(N(t_k)/S_{N(t_k)})$$

which, by (2.11), converges to $(Ea_0)^{1/2} \cdot 0 \cdot (Ea_0)^{-1} = 0$. Apply the converging-together lemma

-11-

to (2.9) and (2.10) to get (1.24) *for* t *going to* ∞ *through the sequence* t_k. Because

that sequence is arbitrary, (1.24) holds without qualification (see Billingsley (1968), p.

16).

Now also assuming (1.34) we have

THEOREM 4. Formula (1.35) holds with

$$(2.12) \qquad \beta = E\alpha_0 Y_0(\hat{f}) + \sum_{k=1}^{\infty} (E\alpha_0 Y_k(\hat{f}) + EY_0(\hat{f})\alpha_k) \ .$$

<u>Proof</u>. The infinite series for β converges absolutely by a remark of Billingsley (1968),

p. 177. Turning to the bias expansion itself, observe that

$$(2.13) \qquad r_n - r = \bar{Y}_n(\hat{f}) g(\bar{\alpha}_n)$$

where

$$(2.14) \qquad g(x) = 1/x \ .$$

Choosing ε small enough,

$$(2.15) \qquad \sup\{g''(x) : |x - E\alpha_0| < \epsilon\} = M < \infty \ .$$

Letting event $B_n = \{|\bar{\alpha}_n - E\alpha_0| < \epsilon\}$, use (1.34) and then Chebyshev's inequality to get

$$(2.16) \qquad |E\{r_n - r; B_n^c\}| < KP\{|\bar{\alpha}_n - E\alpha_0| > \epsilon\}$$

$$< KE(\bar{\alpha}_n - E\alpha_0)^4/\epsilon^4 \ .$$

By Billingsley [(1968), lemma 4, p. 172],

$$(2.17) \qquad E(\bar{\alpha}_n - E\alpha_0)^4 = 0(n^{-2}) \ .$$

Next

$$E\{r_n - r\} = E\{r_n - r; B_n\} + E\{r_n - r; B_n^c\} \ .$$

Combining (2.16)-(2.17) gives

$$(2.18) \qquad E\{r_n - r\} = E\{r_n - r; B_n\} + 0(n^{-2}) \ .$$

On the event B_n, expand $g(\bar{\alpha}_n)$ in a Taylor series around $E\alpha_0$ to get

$$(2.19) \qquad E\{r_n - r; B_n\} = S + T$$

where

-12-

(2.20)
$$s = E\left\{\bar{Y}_n(\hat{f})\left[\frac{1}{E\alpha_0} - \frac{1}{(E\alpha_0)^2}(\bar{\alpha}_n - E\alpha_0)\right] \mid B_n\right\}$$

(2.21)
$$T = E\left\{\bar{Y}_n(\hat{f})g''(\xi_n)(\bar{\alpha}_n - E\alpha_0)^2/2 \mid B_n\right\}$$

and

(2.22)
$$|\xi_n - E\alpha_0| < \epsilon .$$

Apply Cauchy-Schwartz and then Billingsley [(1968), Lemmas 3 and 4, p. 172] to get

(2.23)
$$T \leq ME(Y_n(\hat{f})^2)^{1/2}(E(\bar{\alpha}_n - E\alpha_0)^4)^{1/2} = O(n^{-3/2}) .$$

An argument similar to that justifying (2.18) gives

(2.24)
$$s = -E(\bar{Y}_n(\hat{f})\bar{\alpha}_n)/(E\alpha_0)^2 + O(n^{-2}) .$$

A simple calculation shows that

(2.25)
$$E(\bar{Y}_n(\hat{f})\bar{\alpha}_n) = \beta/n + o(1/n)$$

with β given by (2.12). Combining (2.18)-(2.25) gives (1.35), finishing the proof.

COROLLARY. Jackknifing works: (1.38) holds.

THEOREM 5. Formula (1.40) holds.

Proof. By the converging-together lemma, it suffices to prove that

(2.26)
$$\sqrt{n}\ (r_{2n} - \tilde{r}_{2n}) \to 0$$

in probability. Straightforward algebra shows that the left side of (2.26) equals

(2.27)
$$\sqrt{n}\ B_n(C_n - D_n)/(C_n + D_n)$$

where

$$B_n = r(0, n-1) - r(n, 2n-1)$$

$$C_n = \alpha(0, n-1)/n$$

$$D_n = \alpha(n, 2n-1)/n$$

$$\alpha(a, b) = \sum_{j=a}^{b} \alpha_j .$$

An argument   he proof of (1.27) shows that

(2.28)
$$\sqrt{n}\ B_n \implies \sqrt{2}\ \sigma N(0, 1)/E\alpha_0 .$$

Clearly

(2.29)
$$C_n - D_n \to 0 \quad \text{a.s.}$$

Combining (2.27) through (2.29) verifies (2.26).

3. RESULTS AND PROOFS: II.

Theorem 6. Assume that (1.14), (1.17), and (1.18) hold under $\hat{P}$. Then, there exists a probability $\tilde{P}$ such that:

i) $\frac{1}{t} \int_0^t \hat{P}\{X \cdot \theta_s \in \cdot\}ds \to \tilde{P}\{X \in \cdot\}$ as $t \to \infty$

ii) $\tilde{P}\{X \cdot \theta_u \in \cdot\} = \tilde{P}\{X \in \cdot\}$ for $u > 0$

iii) $\tilde{P}\{X \in \cdot\} = \frac{1}{\hat{E}S_1} \int_0^\infty \hat{P}\{X \cdot \theta_s \in \cdot; S_1 > s\}ds.$

Proof. Let f be a bounded nonnegative function on Ω. For $X \in \Omega$, set $f_u(X) = f(X \cdot \theta_u)$ and put

$$g_u(X) = \int_0^{S_1} f_u(X \cdot \theta_s)ds = g_0(X \cdot \theta_u)$$

Proceeding as in the proof of Theorem 1, Birkhoff's ergodic theorem proves that

$$(3.1) \qquad \frac{1}{n} \sum_{k=0}^{n-1} g_u(X \cdot \theta_{S_k}) = \frac{1}{n} \int_u^{S_n + u} f(X \cdot \theta_s)ds \to \hat{E}g_u(X)$$

$\hat{P}$ a.s., as $n \to \infty$. The "squeeze" argument of (2.4), applied to (3.1), can be readily adapted to show that

$$\frac{1}{t} \int_u^{u+t} f(X \cdot \theta_s)ds \to \frac{1}{\hat{E}S_1} \hat{E}g_0(X \cdot \theta_u) \qquad \hat{P} \text{ a.s.}.$$

It follows, from bounded convergence, that

$$\frac{1}{t} \int_0^t \hat{E}f(X \cdot \theta_s)ds \to \frac{1}{\hat{E}S_1} \hat{E}g_0(X \cdot \theta_u)$$

as $t \to \infty$; in particular, specializing f to indicator functions yields

-15-

$$(3.2) \qquad \frac{1}{t} \int_0^t \hat{P}\{X \circ \theta_s \, \epsilon \, \cdot\} ds \to \frac{1}{ES_1} \int_0^\infty \hat{P}\{X \circ \theta_u \circ \theta_s \, \epsilon \, \cdot; \, S_1 > s\} ds$$

as $t \to \infty$. Setting $u = 0$ in (3.2) proves i) and iii). For ii), observe that the left-hand side of (3.2) is independent of u.

Throughout the remainder of this section, we assume that (1.14), (1.17), and (1.18) are in force for $\hat{P}$, and that P denotes probability under the measure $\tilde{P}$ constructed in Theorem 6.

THEOREM 7.

$$(3.3) \qquad P\{r(t) \to r\} = 1 .$$

Proof. Use definitions and then the fact that (1.24) holds under $\hat{P}$ to get

$$(3.4) \qquad \hat{P}\{r(t) \circ \theta_s \to r; \, S_1 > s\}$$

$$= \hat{P}\left\{\int_s^{s+t} f(X(u)) du/t \to r; \, S_1 > s\right\} = \hat{P}\{S_1 > s\} .$$

Now integrate formula (3.4) with respect to s from 0 to ∞. On the right we get $\hat{E}S_1$, positive by our assumptions. On the left, use dominated convergence to take the limit with respect to t inside the integral and then (1.13) to get $\tilde{P}\{r(t) \to r\}\hat{E}S_1$. Cancelling $\hat{E}S_1$ from both sides gives (3.3).

THEOREM 8. Under P formulas (1.45) and (1.46) hold.

Proof. For any x,

(3.5)
$$\hat{P}\{\sqrt{t}\ (r(t) - r) \cdot \theta_s < x;\ S_1 > s\}$$

$$= \hat{P}\left\{(1/\sqrt{t}\int_s^{s+t} \hat{f}(X(u)) < x|S_1 > s\right\}\{S_1 > s\}du\ .$$

By Billingsley [(1968), Theorem 20.2] the conditional probability above converges to the appropriate normal probability as $t \to \infty$, for every x and s. Integrate both sides of (3.5) with respect to s from 0 to ∞, use (1.13) to see that the left side equals $P\{\sqrt{t}\ (r(t) - r) < x\}\hat{E}S_1$, use dominated convergence on the right to take the limit with respect to t inside the outer integral, and recall that any nonnegative random variable Z has expectation $\int_0^\infty P\{Z > t\}dt$. Cancel $\hat{E}S_1$ from both sides of (3.4) integrated as above to get that $P\{\sqrt{t}\ (r(t) - r) < x\}$ converges to the appropriate normal probability. A routine calculation verifies (1.46).

4. INDIRECT ESTIMATION OF TRANSACTION AVERAGES.

Suppose a performance measure s aggregates costs for transactions moving through the system. Let transaction i arrive at time A_i and leave at time B_i. Put $D_i = B_i - A_i$. Assume that the cost associated with transaction i is $g_i(D_i)$ and that the average cost

$$(4.1) \qquad s = \lim_{n \to \infty} [g_1(D_1) + \ldots + g_n(D_n)]/n$$

exists a.s.; clearly

$$(4.2) \qquad s_n = [g_1(D_1) + \ldots + g_n(D_n)]/n$$

consistently estimates s. Our discussion carries over to more general transaction averages. We use s for concreteness.

In the notation of Section 1, state $\tilde{X}_k$ includes the value of A_i for every transaction i in the system at time T_k and

$$(4.3) \qquad f(X(t)) = \sum g_i'(t - A_i)$$

where the sum is over those transactions i for which $A_i \le t \le B_i$. Here g_i' is the right derivative of g_i and by definition $g_i'(d) = 0$ for $d < 0$ and for $d > D_i$. This setup allows rather general g_i's, though not indicators.

Now call the arrival rate λ. Assuming all limits exist, $r = \lambda s$ by Heyman and Stidham (1980). Usually λ is simply the reciprocal of the expected arrival spacing. In more general cases such as batch arrivals with batch size and spacing dependent, the theory developed earlier in this paper shows how to estimate λ.

We can choose whether to observe discretized observations $Y_n(f)$ as detailed earlier or transaction observations $g_i(D_i)$. Our choice is based on four criteria:

(i) Ease of data collection. Formula (4.3) indicates that gathering discretized observations at state-change epochs is more work, but probably not much more because the evolution of the system has to be simulated in any case. Gathering discretized observations $f(X(t))$ at (reasonably short) equally-spaced intervals that do not necessarily correspond to state-change epochs may be much more work.

(ii) True variance. Some believe that λs_n always has smaller variance than r_m where n is the number of transactions completely processed up to event epoch m. To some

-18-

extent the results of Carson and Law (1980) <u>for special cases</u> support this folklore, but
Cooper's [(1981), pp. 293-296] example shows that it is untrustworthy.

<u>(iii) Variance estimation</u>. Folklore has it that the sequence of transaction observations
$\{g_i(D_i)\}$ is covariance stationary if and only if the sequence of discretized observations
$\{Y_n(\hat{t})\}$ is covariance stationary. Thus, loosely speaking, any variance-estimation
technique (e.g., via batch means, spectral analysis, autoregressive representations,
functional limit theorems) applies to both or to neither. In practice, however, one has to
ask at what sample sizes asymptotic results <u>reasonably</u> apply: the faster covariance falls
off with increasing lag, the better. Generally, for discretized observations, covariance
does drop fairly quickly. By contrast, especially in systems that are not first-in, first-
out (FIFO), a series of transaction observations scrambles past, present, and future and
cuts connections between time spacing and index spacing. Thus, covariance between widely-
spaced observations may well be significant. This messy covariance structure is hard to
handle, at least with practical samaple sizes.

<u>(iv) Bias</u>. The scrambling effect mentioned in point (iii) above may make it harder to
detect (and hence attenuate) initialization bias for transaction observations.
Incompletely processed transactions cause termination bias for transaction observations,
exacerbated by the "inspection paradox" discussed earlier. For discretized observations,
r_n is generally biased but our jackknife using $\tilde{r}_n$ makes this bias nearly negligible.

We conclude that, at least for non-FIFO systems, indirect estimation via discretized
observations has more advantages than disadvantages relative to direct estimation via
transaction observations. The results developed in the preceding sections make statistical
analysis of the former possible and practical.

Asymptotically, $s = r/\lambda$. Estimate s by plugging in consistent estimators for r
and λ. If, as is usual, λ is known, then divide confidence limits for r by λ to get
confidence limits for s. If λ has to be estimated, a straightforward ratio-estimator
analog of the methods discussed in this paper must be used to obtain an asymptotically
exact confidence interval.

-19-

REFERENCES

1. P. BILLINGSLEY (1968), <u>Convergence of Probability Measures</u>, Wiley, New York.

2. P. BRATLEY, B. L. FOX, and L. E. SCHRAGE (1983), <u>A Guide to Simulation</u>, Springer-Verlag, New York.

3. J. S. CARSON and A. M. LAW (1980), <u>Conservation equations and variance reduction in queuing simulations</u>, Oper. Res., 28, pp. 535-546.

4. R. B. COOPER (1981), <u>Introduction to Queueing Theory</u>, 2nd edn., North Holland, New York.

5. P. FRANKEN, D. KÖNIG, U. ARNDT, and V. SCHMIDT (1982), <u>Queues and Point Processes</u>, Wiley, New York.

6. P. W. GLYNN and D. L. IGLEHART (1981), <u>Simulation output analysis for general state space Markov chains</u>, Proc. of the ORSA-TIMS Special Interest Meeting on Applied Probability - Computer Science, The Interface.

7. D. P. HEYMAN and M. J. SOBEL (1982), <u>Stochastic Models in Operations Research, Vol. 1</u>, McGraw-Hill, New York.

8. D. P. HEYMAN and S. STIDHAM, Jr. (1980), <u>The relation between customer and time averages in queues</u>, Oper. Res., 28, pp. 983-994.

9. L. JOW (1982), <u>An autoregressive method for simulation output analysis</u>, Technical Report 21, Dept. of Operations Research, Stanford University, Stanford, California.

10. J. LAMPERTI (1977), <u>Stochastic Processes</u>, Springer-Verlag, New York.

11. D. R. MILLER (1972), <u>Existence of limits in regenerative processes</u>, Ann. Math. Stat., 43, pp. 1275-1282.

12. A. STRELLER (1980), <u>A generalization of cumulative processes</u>, Elektron. Informationsverarb. Kybernetik, 16, pp. 449-460.

13. R. WOLFF (1982), <u>Poisson arrivals see time averages</u>, Oper. Res., 30, pp. 223-231.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 2638	2. GOVT ACCESSION NO. AD-A139 315	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ESTIMATING TIME AVERAGES VIA RANDOMLY-SPACED OBSERVATIONS		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Bennett L. Fox and Peter W. Glynn		8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of 610 Walnut Street Wisconsin Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 5 - Optimization and Large Scale Systems
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709		12. REPORT DATE January 1984
		13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Simulation, point processes

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

To estimate continuous-time averages via randomly-spaced observations of discrete-event systems, we develop a point-process framework and use it to generalize both regenerative and stationary-process oriented simulation methodologies. We give consistent estimators, central limit theorems, and an effective bias-reducing jackknife. The impact on indirect estimation of transaction (customer) averages is discussed.