(1)

CONDITIONS FOR THE APPLICABILITY OF THE REGENERATIVE METHOD

by

Peter W. Glynn and Donald L. Iglehart

TECHNICAL REPORT No. 71

SEPTEMBER 1991

Prepared under the Auspices
of
U.S. Army Research Contract
DAAL-03-91-G-0101

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

CONDITIONS FOR THE APPLICABILITY OF THE REGENERATIVE METHOD

by

Peter W. Glynn

and

Donald L. Iglehart
Stanford University

Abstract.

The regenerative method for estimating steady-state parameters is one of the basic methods in the simulation literature. This method depends on central limit theorems for regenerative processes and weakly consistent estimates for the variance constants arising in the central limit theorem. A weak sufficient (and probably necessary) condition for both the central limit theorems and consistent estimates is given. Several references have claimed (without proof) strong consistency of the variance estimates under our condition. This claim seems to us unjustified and we hope this note helps clarify the situation. Also discussed is the relationship between conditions for the validity of the regenerative method and those for the validity of the standardized time series methods.

Keywords.

Consistent variance estimators, regenerative method, regenerative processes, simulation, standardized time series, steady-state estimation.

This work was supported by the Army Research Office contract DAAL-03-91-G-0101, National Science Foundation grant NSF DDM-9101580, and the Stanford Institute for Manufacturing Automation.

1. Introduction.

The regenerative method (RM) for estimating steady-state parameters via simulation has been widely studied; cf. Bratley, Fox, and Schrage (1987), p. 95, Crane and Lemoine (1977), and Law and Kelton (1991), p. 557. Our goal in this paper is to develop the weakest known condition under which the RM is valid. [This condition is probably necessary, as well as sufficient.] This condition is of interest since a number of errors on this subject have appeared in the literature. Furthermore, since the RM is perhaps the cleanest setting for simulation output analysis, it is important to have a good understanding of the required condition. This paper also discusses the relationship between conditions for the validity of the regenerative methods and those for the validity of the standardized times series methods (STSM).

Let $X = (X(t): t \geq 0)$ be a (possibly delayed) regenerative process on state space S with regeneration times $T(-1) = 0 \leq T(0) < T(1) < \cdots$. Denote the regenerative cycle lengths by $\tau_n = T(n) - T(n-1)$, $n \geq 0$. Let $f: S \to \mathbb{R}$ be a given measurable function and define the sequence of random variables

$$Y_n(f) = \int_{T(n-1)}^{T(n)} f[X(s)]ds, \qquad n \geq 0.$$

Often we abbreviate $Y_n(f)$ as Y_n, when no ambiguity arises. Throughout this paper we assume that the regenerative process X is positive recurrent, so that $E\tau_1 < \infty$. The condition that follows is the one we use to guarantee the validity of the RM.

Condition A.

There exists a finite constant α *such that* $E[Y_1 - \alpha\tau_1] = 0$ *and* $0 < E[(Y_1 - \alpha\tau_1)^2] < \infty$.

Let $Z_n \equiv Y_n - \alpha\tau_n$, $n \geq 1$. The regenerative property of X implies that the sequence $(Z_n : n \geq 1)$ is independent and identically distributed (i.i.d.). Condition A implies that $E|Z_1| < \infty$, $E|Y_1| < \infty$, and $\alpha = EY_1/E\tau_1$.

Two point estimators for α arise in the RM. The first, $\alpha(t)$, is based on a simulation from time 0 to t. The second, α_n, is based on a simulation of n regenerative cycles. These estimators are defined as

$$\alpha(t) \equiv t^{-1}\int_0^t f[X(s)]ds, \qquad t > 0,$$

and

$$\alpha_n \equiv \bar{Y}_n/\bar{\tau}_n,$$

where $\bar{Y}_n = n^{-1}\sum_{k=1}^{n} Y_k$ and $\bar{\tau}_n = n^{-1}\sum_{k=1}^{n}\tau_k$. The fact that $\alpha(t)$ is strongly consistent for α when $E|Y_1| < \infty$ and $E\tau_1 < \infty$ (converges to α a.s. as $t \to \infty$) is well known in the regenerative process literature; cf. Chung (1960), p. 87, for a proof in the Markov chain case. Under the same condition α_n is strongly consistent for α by the strong law of large numbers for i.i.d. random variables (r.v.).

The goal of the RM is to produce asymptotically valid confidence intervals for α as the length of the simulation becomes large ($t \to \infty$ or $n \to \infty$). This requires the following central limit theorems (c.l.t.) for both $\alpha(t)$ and α_n:

(1.1) **Theorem.** *If condition A holds, then*

$$t^{1/2}[\alpha(t) - \alpha] \Rightarrow \sigma_1 N(0,1) \text{ as } t \to \infty, \tag{1.2}$$

$$n^{1/2}[\alpha_n - \alpha] \Rightarrow \sigma_2 N(0,1) \text{ as } n \to \infty, \tag{1.3}$$

where $\Rightarrow$ *denotes weak convergence,* $N(0,1)$ *is a normal r.v. mean* 0 *variance* 1, $\sigma_1^2 = EZ_1^2/E\tau_1$, *and* $\sigma_2^2 = EZ_1^2/(E\tau_1)^2$.

Again, this theorem is well established; cf. Chung (1960), p. 94. To complete the establishment of asymptotic confidence intervals for α we need weakly consistent estimators $v(t)$ and v_n for σ_1^2 and σ_2^2 respectively. Given such estimators, it follows from Theorem 1.1 and the continuous mapping theorem (Billingsley (1968), p. 30) that the intervals

$$\left[\alpha(t) - \frac{z(\delta)v(t)^{1/2}}{t^{1/2}},\ \alpha(t) + \frac{z(\delta)v(t)^{1/2}}{t^{1/2}}\right] \tag{1.4}$$

$$\left[\alpha_n - \frac{z(\delta)v_n^{1/2}}{n^{1/2}},\ \alpha_n + \frac{z(\delta)v_n^{1/2}}{n^{1/2}}\right] \tag{1.5}$$

are asymptotic $100(1-\delta)\%$ confidence intervals for α, where $z(\delta)$ is chosen so that $P[N(0,1) \le z(\delta)] = 1 - \delta/2$. In Section 2 we show that condition A is sufficient to guarantee the existence of weakly consistent estimators for σ_1^2 and σ_2^2. We note that Condition A does not imply that $EY_1^2(f) < \infty$ and $E\tau_1^2 < \infty$, so that standard arguments based on the law of large numbers do not apply.

2. Consistency of Regenerative Variance Estimators

The two variance estimators we consider for σ_1^2 and σ_2^2 respectively are

$$v(t) = t^{-1} \sum_{i=1}^{N(t)} [Y_i - \alpha(t)\tau_i]^2, \qquad t > 0$$

and

$$v_n = \frac{n^{-1} \sum_{i=1}^{n} [Y_i - \alpha_n \tau_i]^2}{(\bar{\tau}_n)^2}, \qquad n \geq 1$$

where $N(t) = \sup\{n \geq -1 : T(n) \leq t\}$, the number of completed regenerative cycles in the interval $[0,t]$. Our main result, contained in the next theorem, is to show weak consistency of $v(t)$ and v_n under the same condition as that required for the c.l.t.'s (see Theorem 1.1).

(2.1) **Theorem.** *If condition A holds, then*

(2.2) $$v(t) \Rightarrow \sigma_1^2 \qquad \text{as } t \to \infty$$

and

(2.3) $$v_n \Rightarrow \sigma_2^2 \qquad \text{as } n \to \infty.$$

Proof. We shall only prove (2.2) here, as the proof for (2.3) is similar. First we rewrite $v(t)$ as

(2.4) $$\begin{aligned} v(t) &= t^{-1} \sum_{i=1}^{N(t)} [Y_i - \alpha\tau_i - (\alpha(t) - \alpha)\tau_i]^2 \\ &= t^{-1} \sum_{i=1}^{N(t)} Z_i^2 - 2t^{-1} \sum_{i=1}^{N(t)} Z_i \tau_i (\alpha(t) - \alpha) \\ &\quad + t^{-1}(\alpha(t) - \alpha)^2 \sum_{i=1}^{N(t)} \tau_i^2. \end{aligned}$$

The first term on the right-hand side (r.h.s.) of (2.4) converges to σ_1^2 a.s. by the strong laws for i.i.d. partial sums and for renewal processes. Hence we need to show that terms two and three on the r.h.s. of (2.4) converge weakly to zero, and then apply the "Converging Together" theorem; cf. Billingsley (1968), p. 25. For term two we know that $t^{1/2}(\alpha(t) - \alpha) \Rightarrow \sigma_1 N(0,1)$. So it suffices to show that

(2.5) $$t^{-\frac{3}{2}} \sum_{i=1}^{N(t)} Z_i \tau_i \to 0 \text{ a.s.} \qquad \text{as } t \to \infty$$

in order to conclude that the second term converges weakly to 0. Since $EZ_1^2 < \infty$ by assumption, $Z_n^2/n \to 0$ a.s. by the Borel-Cantelli lemma; cf., Chung (1968) and apply Theorem 3.2.1, 4.2.1, and 4.2.2. This in turn implies that $Z_n/n^{1/2} \to 0$ a.s., $\max_{1 \leq k \leq n} |Z_k|/n^{1/2} \to 0$ a.s., and $\max_{1 \leq k \leq N(t)} |Z_k|/t^{1/2} \to 0$ a.s.

Now observe that

$$|t^{-\frac{3}{2}}\sum_{i=1}^{N(t)} Z_i\tau_i| \le t^{-\frac{3}{2}}\sum_{i=1}^{N(t)} |Z_i\tau_i|$$
$$\le (t^{-\frac{1}{2}}\max_{1\le k\le N(t)} |Z_k|)\cdot(t^{-1}\sum_{i=1}^{N(t)}\tau_i).$$

As shown above the first term on the r.h.s. above converges to 0 a.s., while the second term converges a.s. to 1 by the strong laws used above. For the third term on the r.h.s. of (2.4), we note that $t^{1/2}(\alpha(t)-\alpha)^2 \Rightarrow \sigma_1^2 N(0,1)^2$ from (1.2). Thus it suffices to show that

(2.6) $$t^{-2}\sum_{i=1}^{N(t)}\tau_i^2 \to 0 \text{ a.s.} \qquad \text{as } t\to\infty.$$

Since $E\tau_1 < \infty$ by assumption, the same Borel-Cantelli argument used above shows that $\max_{1\le k\le N(t)} \tau_k/t \to 0$ a.s.

Thus

$$t^{-2}\sum_{i=1}^{N(t)}\tau_i^2 \le (t^{-1}\max_{1\le k\le N(t)}\tau_k)\cdot(t^{-1}\sum_{k=1}^{N(t)}\tau_k) \to 0 \text{ a.s.}$$

establishing (2.6). Using these results in conjunction with (2.4) shows that $v(t) \Rightarrow \sigma_1^2$. ∎

There has been the claim in the simulation literature that Condition A implies strong consistency of $v(t)$ and v_n; cf. Bratley, Fox, and Schrage (1987), p. 118-119; Crane and Lemoine (1977), p. 42; Law and Kelton (1991), p. 559; and Shedler (1987), p. 29. No proof of this claim has been given, and we suspect that it is not true based on the argument required to prove Theorem 2.1. Certainly the claim is true under the added condition $E\tau_1^2 < \infty$, which may be what was intended in the references above.

The argument developed above also establishes the weak consistency of v_n for the general ratio estimation problem of the form $\alpha = EY_1/E\tau_1$. Here we assume that the pairs $\{(Y_i,\tau_i): i\ge 1\}$ are i.i.d., that $E|\tau_1| < \infty$, there exists a finite constant α such that $E[Y_1-\alpha\tau_1]=0$, and $E[(Y_1-\alpha\tau_1)^2]<\infty$. Confidence intervals for α can now be constructed as was done in the equation (1.5).

3. Conditions for Standardized Time Series Method

There are two principal approaches for estimating a steady-state parameter from a single simulation run. The RM is an example of the first approach in which weakly consistent estimation of the variance constant (σ_1^2 or σ_2^2 in (1.2) and (1.3)) is required. Other examples of this approach are the spectral and autoregressive methods. The second approach, proposed by Schruben (1983), are the standardized time series method (STSM). In this approach the variance constant is "canceled out" in a manner reminiscent of the t-statistic. The popular batch means method is a special case of the STSM. The starting point for the STSM is the existence of a functional central limit theorem (f.c.l.t.); see Glynn and Iglehart (1990) for this development.

To discuss the f.c.l.t. we introduce the random functions

$$X_n(t) \equiv n^{-1}\int_0^{nt} f[X(s)]ds$$

for $n \geq 1$ and $t \geq 0$, where $X = (X(t): t \geq 0)$ is the regenerative process introduced in Section 1. Sample functions of $X_n(\cdot)$ lie in the space $C[0,\infty)$ of real-valued continuous functions on $[0,\infty)$. Necessary and sufficient conditions for a f.c.l.t. to hold for X_n were recently obtained by Glynn and Whitt (1991). To state this result we define the sequence of i.i.d. r.v.'s

$$W_n(f) = \sup_{0 \leq s \leq \tau_n} |\int_0^s f[X(T_{n-1}+u)du|, \qquad n \geq 1.$$

The result of Glynn and Whitt is

(3.1) **Theorem.** *There exist finite-valued deterministic constants α and σ such that $Y_n \Rightarrow \sigma B$ in $C[0,\infty)$ if and only if*

(3.2) $$E[(Y_1(f-\alpha))^2] < \infty$$

and

(3.3) $$n^2 P[W_1(f-\alpha) > n] \to 0 \qquad n \to \infty,$$

where $Y_n(t) = n^{\frac{1}{2}}[X_n(t) - \alpha t]$ and B is standard Brownian motion. In case conditions (3.2) and (3.3) hold, $\alpha = EY_1(f)/E\tau_1$ and $\sigma^2 = E[(Y_1(f-\alpha))^2]/E\tau_1$.

Note that $Y_1(f-\alpha) = Z_1$. Observe that condition (3.3) is an additional condition needed for the f.c.l.t. that is not needed for the c.l.t. Here is an example of a regenerative process that satisfies the c.l.t. but not the f.c.l.t.

Let $(U_n : n \geq 0)$ and $(V_n : n \geq 0)$ be two sequences of real-valued r.v.'s satisfying the following conditions:

(3.4) $\{(U_n, V_n) : n \geq 0\}$ is a sequence of independent pairs,

(3.5) $(U_n : n \geq 0)$ is an identically distributed sequence with $EU_1 = 0$ and $EU_1^2 \equiv \sigma^2 < \infty$; and

(3.6) $(V_n : n \geq 0)$ is an identically distributed sequence with $E|V_1| < \infty$ and $E|V_1|^{3/2} = \infty$.

Now let $T(0) = 0$ and $T(n) = 2n$ for $n \geq 1$; all regenerative cycles are of length 2. For $n \geq 0$, set $X_{2n} = U_n + V_n$, $X_{2n+1} = -V_n$, and $f(x) = x$. Finally, set $X(t) = X_{\lfloor t \rfloor}$ for $t \geq 0$. Observe that $X = (X(t): t \geq 0)$ is a regenerative process with $\alpha = 0$, $Y_n = U_{n-1}$ $Z_n = U_{n-1}$ and $W_n = \max\{|U_{n-1} + V_{n-1}|, |U_{n-1} + 2V_{n-1}|\}$. If we can exhibit U_0 and V_0 such that $EW_1^{3/2} = \infty$, then $\sum_{n=1}^{\infty} n^{1/2} P(W_1 > n) = \infty$ and (3.3) will be violated.

Note that $EW_1^2 \geq E|U_0 + V_0|^2$. Now take U_0 and V_0 independent with $P\{U_0 = +1\} = P\{U_0 = -1\} = \frac{1}{2}$ and $P\{V_0 = n^{1/2}\} = c/n^{3/2}$ for $n = 1, 2, \ldots$ and a positive constant c. It is easy to check that (3.3) is not valid.

The example above shows that the regenerative c.l.t. can hold without the f.c.l.t. holding; this result is mentioned in Bratley, Fox, and Schrage (1987), p. 121. Thus the RM is valid for cases in which the f.l.c.t leading to the STSM is not. In this sense the RM has a larger domain of applicability than does the STSM.

Of course, it is also the case that the f.c.l.t (and hence the STSM) can hold for a non-regenerative stochastic process for which the RM is no longer valid. Hence, neither the class of stochastic processes for which the RM holds nor the class for which the STSM holds contains the other.

REFERENCES

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

[2] Bratley, P., B.L. Fox, and L.E. Schrage. (1987). *A Guide to Simulation*, Second Edition. Springer-Verlag, New York.

[3] Chung, K.S. (1960). *Markov Chains with Stationary Transition Probabilities.* Springer-Verlag, New York.

[4] Chung, K.L. (1968). *A Course in Probability Theory.* Harcourt, Brace & World, New York.

[5] Crane, M.A. and A.J. Lemoine. (1977). *An Introduction to the Regenerative Method for Simulation Analysis.* Springer-Verlag, New York.

[6] Glynn, P.W., and D.L. Iglehart. (1990). Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15, 1-16.

[7] Glynn, P.W. and W. Whitt. (1991). *Limit theorems for regenerative processes: necessary and sufficient conditions.* Forthcoming.

[8] Law, A.M. and W.D. Kelton. (1991). *Simulation Modeling and Analysis.* McGraw-Hill, New York.

[9] Schruben, L. (1983). Confidence internal estimation using standardized time series. *Operations Research* 31, 1090-1108.

[10] Shedler, G.S. (1987). *Regeneration and Networks of Queues.* Springer-Verlag, New York.

OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
September 1991

3. REPORT TYPE AND DATES COVERED
Technical

4. TITLE AND SUBTITLE
Conditions for the Applicability of the Regenerative Method

5. FUNDING NUMBERS

6. AUTHOR(S)
Peter W. Glynn and Donald L. Iglehart

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Operations Research - 4022
Terman Engineering Center
Stanford University
Stanford, CA 94305-4022

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
(please see next page)

14. SUBJECT TERMS
Consistent variance estimators, regenerative method, regenerative processes, simulation, standardized time series, steady-state estimation.

15. NUMBER OF PAGES
8

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev 2-89)
Prescribed by ANSI Std Z39-18
298-102

Abstract.

The regenerative method for estimating steady-state parameters is one of the basic methods in the simulation literature. This method depends on central limit theorems for regenerative processes and weakly consistent estimates for the variance constants arising in the central limit theorem. A weak sufficient (and probably necessary) condition for both the central limit theorems and consistent estimates is given. Several references have claimed (without proof) strong consistency of the variance estimates under our condition. This claim seems to us unjustified and we hope this note helps clarify the situation. Also discussed is the relationship between conditions for the validity of the regenerative method and those for the validity of the standardized time series methods.