Notes: Conditions for the Applicability of the Regenerative Method

Peter W. Glynn • Donald L. Iglehart
Department of Operations Research, Stanford University, Stanford, California 94305

The regenerative method for estimating steady-state parameters is one of the basic methods in simulation output analysis. This method depends on central limit theorems for regenerative processes and weakly consistent estimates for the variance constants arising in the central limit theorems. A weak sufficient condition for both the central limit theorems and consistent estimates is given. Previous authors have implicitly made stronger moment assumptions which have led to strongly consistent variance estimates, more than is needed for the regenerative method to hold. The relationship between conditions for the validity of the regenerative method and those for the validity of standardized time series methods is also discussed.
(*Consistent Variance Estimators; Regenerative Method; Regenerative Processes; Simulation; Standardized Time Series; Steady-state Estimation*)

1. Introduction

The regenerative method (RM) for estimating steady-state parameters via simulation has been widely studied; cf. Bratley et al. (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. 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 standardized times series methods (STSMs).

Let $\mathbf{X} = \{X(t): t \geq 0\}$ 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

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

Often we abbreviate $Y_n(f)$ as Y_n, when no ambiguity arises. Throughout this paper we assume that the regenerative process $\mathbf{X}$ 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. *For* $\mathbf{X}$, *a regenerative process with* $E\tau_1 < \infty$, *there exists a finite constant* α *such that*

$$E[Y_1 - \alpha\tau_1] = 0 \quad \textit{and} \quad 0 < E[(Y_1 - \alpha\tau_1)^2] < \infty.$$

Let $Z_n \equiv Y_n - \alpha\tau_n$, $n \geq 1$. The regenerative property of $\mathbf{X}$ 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

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

and $\alpha_n \equiv \bar{Y}_n/\bar{\tau}_n$, where $\bar{Y}_n = n^{-1} \sum_{k=1}^n Y_k$ and $\bar{\tau}_n = n^{-1} \times \sum_{k=1}^n \tau_k$. 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 conditions α_n is

0025-1909/93/3909/1108$01.25

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.'s) for both $\alpha(t)$ and α_n:

(1.1) THEOREM. *If Condition A holds, then*

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

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

where $\Rightarrow$ *denotes weak convergence*, $N(0, 1)$ *is a normal random variable with mean* 0 *and 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 construct 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

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

$$\left[\alpha_n - \frac{z(\delta)v_n^{1/2}}{n^{1/2}}, \alpha_n + \frac{z(\delta)v_n^{1/2}}{n^{1/2}}\right] \tag{1.4}$$

are asymptotic $100(1-\delta)\%$ confidence intervals for α, where $z(\delta)$ is defined by

$$P[N(0, 1) \le z(\delta)] = 1 - \delta/2.$$

In §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. Here is an example.

(1.5) EXAMPLE. Here we develop an example for which Condition A holds, but $EY_1^2(f) = \infty$ and $E\tau_1^2 = \infty$. Let $\{\tau_n: n \ge 1\}$ be a sequence of i.i.d. cycle lengths with $\tau_n \ge 1$, a.s., $E\tau_1 = \mu < \infty$, and $E\tau_1^{3/2} = \infty$. Let $\{N_n(0, 1): n \ge 1\}$ be an i.i.d. sequence of mean 0, variance 1 normal random variables which are independent of the τ_n's. Now we define a discrete-time regenerative process $\{X_n: n \ge 0\}$ with state space $S = \mathbb{R}$ by

$$X_{T(n)} = \tau_{n+1} + N_{n+1}(0, 1), \quad n \ge 0, \quad \text{and}$$

$$X_j = 0, \quad j \notin \{T(0), T(1), \cdots\}.$$

As usual a continuous-time regenerative process $\mathbf{X} = \{X(t): t > 0\}$ can be formed by setting $X(t) = X_{\lfloor t \rfloor}$. The f function here is the identity function. For process $\mathbf{X}$, $Y_n \equiv \tau_n + N_n(0, 1)$, $\alpha = 1$, and $Z_n = N_n(0, 1)$. Thus $EZ_n = 0$, $EZ_n^2 = 1$, $E\tau_n^2 = \infty$, and $EY_n^2 = \infty$. □

2. Consistency of Regenerative Variance Estimators

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

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

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

where $N(t) = \sup\{n \ge -1: T(n) \le t\}$. When $N(t) \le 0$, we set $v(t) = 0$. 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*

$$v(t) \Rightarrow \sigma_1^2 \quad \text{as} \quad t \to \infty, \quad \text{and} \tag{2.2}$$

$$v_n \Rightarrow \sigma_2^2 \quad \text{as} \quad n \to \infty. \tag{2.3}$$

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

$$\begin{aligned} v(t) &= t^{-1} \sum_{i=1}^{N(t)} [Y_i - \alpha\tau_i - (\alpha(t) - \alpha)\tau_i]^2 \\ &= t^{-1} \sum_{i=1}^{N(t)} Z_1^2 - 2t^{-1} \sum_{i=1}^{N(t)} Z_i\tau_i(\alpha(t) - \alpha) \\ &\quad + t^{-1}(\alpha(t) - \alpha)^2 \sum_{i=1}^{N(t)} \tau_i^2. \end{aligned} \tag{2.4}$$

The first term on the right-hand side (r.h.s.) of (2.4) converges (as $t \to \infty$) 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

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

in order to conclude that the second term converges weakly to 0. Let $Q_n \equiv n^{-1} \sum_{i=1}^{n} Z_i^2$ for $n \geq 1$. Since $EZ_1^2 < \infty$, we have by the strong law of large numbers

$$\lim_{n\to\infty} Q_n = \lim_{n\to\infty} \left(\frac{n-1}{n}\right) Q_{n-1} = EZ_1^2 \quad \text{a.s.}$$

which implies that

$$\lim_{n\to\infty} Z_n^2/n = \lim_{n\to\infty} \left[Q_n - \left(\frac{n-1}{n}\right) Q_{n-1} \right] = 0 \quad \text{a.s.}$$

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

$$\left| t^{-3/2} \sum_{i=1}^{N(t)} Z_i \tau_i \right| \leq t^{-3/2} \sum_{i=1}^{N(t)} |Z_i \tau_i|$$

$$\leq \left(t^{-1/2} \max_{1\leq k\leq N(t)} |Z_k| \right) \left(t^{-1} \sum_{i=1}^{N(t)} \tau_i \right).$$

As shown above the first term on the r.h.s. converges to zero a.s., while the second term converges to one a.s. by the strong laws used previously. For the third term on the r.h.s. of (2.4), we note that

$$t(\alpha(t) - \alpha)^2 \Rightarrow \sigma_1^2 N(0, 1)^2$$

from (1.2). Thus it suffices to show that

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

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

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

establishing (2.5). □

A number of previous authors have assumed our Condition A plus (implicitly) that $EY_1^2 < \infty$ and $E\tau_1^2 < \infty$; cf. Bratley et al. (1987, p. 118–119), Crane and Lemoine (1977, p. 42), Law and Kelton (1991, p. 559), and Shedler (1987, p. 29). From these assumptions they show the strong consistency of v_n. Strong consistency is more than is needed for the regenerative method to hold.

The argument developed above also established 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 $\{(Y_i, \tau_i): i \geq 1\}$ are i.i.d., $E|\tau_1| < \infty$, 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 Equation (1.4).

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), is standardized time series methods (STSMs). 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 STSMs 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 $\mathbf{X} = (X(t): t \geq 0)$ is the regenerative process introduced in §1 and f is measurable. 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 obtained by Glynn and Whitt (1991). To state this result we define the sequence of i.i.d. r.v.'s

$$W_n(f) = \sup_{0\leq s\leq \tau_n} \left| \int_0^s f[X(T(n-1) + u)]du \right|, \quad n \leq 1.$$

The result of Glynn and Whitt is

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

$$E[(Y_1(f-\alpha))^2] < \infty \quad \text{and} \tag{3.2}$$

$$n^2 P[W_1(f-\alpha) > n] \to 0 \quad \text{as} \quad n \to \infty, \tag{3.3}$$

where $U_n(t) = n^{1/2}[X_n(t) - \alpha t]$ and B is standard Brownian motion. In case conditions (3.2) and (3.3) hold,

$$\alpha = EY_1(f)/E\tau_1 \quad \text{and} \quad \sigma^2 = E[(Y_1(f-\alpha))^2]/E\tau_1.$$

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. The next example shows that the regenerative c.l.t. (Theorem 1.1) can hold without the f.c.l.t. holding; cf. Bratley et al. (1987, p. 121).

(3.4) EXAMPLE. This example is a continuation of Example (1.5). Recall that $\alpha = 1$ and f is the identity function. It is easy to check that

$$W_1(f-\alpha) = \max\{|\tau_1 + N_1(0,1) - 1|, |N_1(0,1)|\}.$$

Therefore

$$E[W_1(f-\alpha)]^{3/2} \geq E|\tau_1 + N_1(-1,1)|^{3/2} = \infty,$$

since $E\tau_1^{3/2} = \infty$ by assumption. This implies that

$$\sum_{n=1}^{\infty} n^{1/2} P\{W_1(f-\alpha) > n\} = \infty$$

and (3.3) is violated. □

Thus the RM is valid for cases in which the f.c.l.t. leading to STSMs is not. In this sense the RM has a larger domain of applicability than do the STSMs. Of course, it is also the case that the f.c.l.t. (and hence STSMs) can hold for a nonregenerative 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 STSMs hold contains the other.

4. Conclusions

The RM is one of the basic methods used for estimating steady-state parameters of a stochastic system from a simulation. This paper presents the weakest known condition for the validity of the RM. Thus the RM is now available for regenerative processes not previously covered. The other basic approach to simulation output analysis are STSMs. Also shown in this paper is an example where the RM is valid and STSMs are not. The converse is also true. As for further research, the principal question is establishing whether or not Condition A is necessary for the RM.[1]

[1] We are grateful for the suggestions provided by the Department Editor and an Associate Editor for improving the paper. 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.

References

Billingsley, P., *Convergence of Probability Measures*, John Wiley, New York, 1968.

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

Chung, K. S., *Markov Chains with Stationary Transition Probabilities*, Springer-Verlag, New York, 1960.

Chung, K. L., *A Course in Probability Theory*, Harcourt, Brace & World, New York, 1968.

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

Glynn, P. W. and D. L. Iglehart, "Simulation Output Analysis Using Standardized Time Series," *Math. Oper. Res.*, 15 (1990), 1–16.

—— and W. Whitt, Limit Theorems for Cumulative Processes," Technical Report No. 72, Department of Operations Research, Stanford University, Stanford, CA, 1991 (to appear in *Stoch. Proc. and Their Applic.*).

Law, A. M. and W. D. Kelton, *Simulation Modeling and Analysis* (2nd ed.), McGraw-Hill, New York, 1991.

Schruben, L., "Confidence Interval Estimation Using Standardized Time Series," *Oper. Res.*, 31 (1983), 1090–1108.

Schedler, G. S., *Regeneration and Networks of Queues*, Springer-Verlag, New York, 1987.

All notes are refereed. Accepted by James R. Wilson; former Departmental Editor; received September 20, 1991. This paper has been with the authors $4\frac{1}{2}$ months for 2 revisions.