Asymptotic Analysis of a Fully Sequential Selection Procedure for Steady-State Simulation

Barry L. Nelson
Department of Industrial Engineering &
Management Sciences
Northwestern University

Supported by NSF Grant DMI-9622065

Steady-State Simulation

A goal of **steady-state simulation** is to compare the limiting behaviors of several ergodic stochastic processes. Finding the service system that minimizes long-run average customer delay in queue is an example.

Because the simulation cannot be initialized in "steady state," point estimators based on finite samples are typically biased.

To minimize bias, the standard experiment design calls for a single, long replication (perhaps with some data deletion at the beginning).

Even in the best case this provides dependent and nonnormal (but stationary) data for statistical inference.

Notation

$X_{i1}, X_{i2}, X_{i3}, \ldots$ output from a single replication of system $i = 1, 2, \ldots, k$

μ_i is the steady-state mean performance of system i

Assume $\mu_k \geq \mu_{k-1} \geq \cdots \geq \mu_1$ so that system k is the best (although we don't really know its identity).

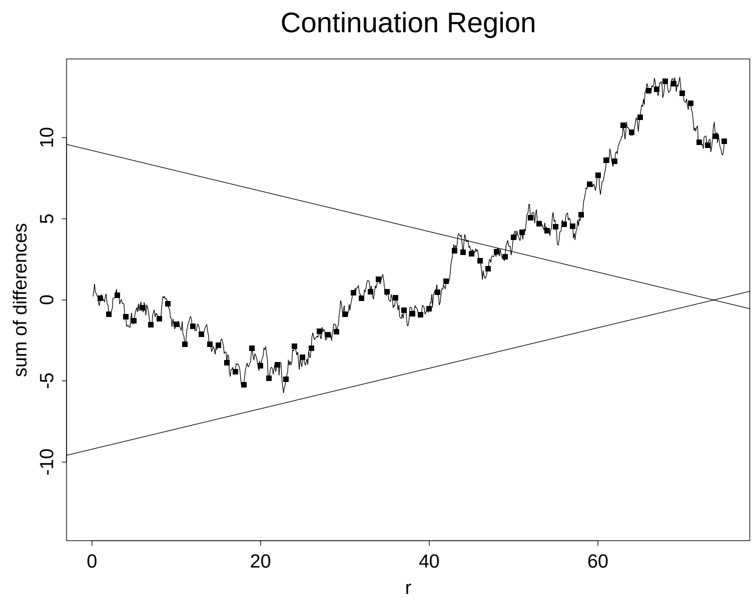
δ is a practically significant difference worth detecting.

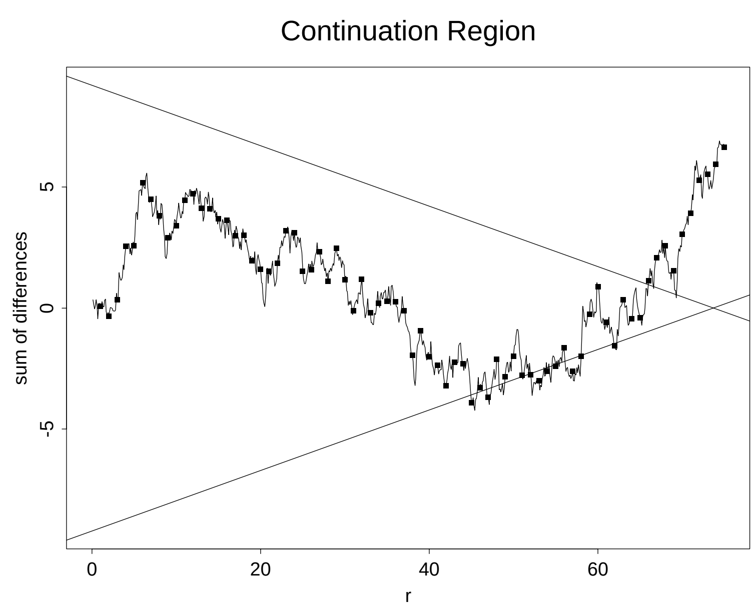
Our procedures focus on the sum of differences through sampling stage r

$$\sum_{j=1}^{r} (X_{kj} - X_{ij})$$

and an approximating Brownian motion process $\mathcal{B}(r, \Delta), r \geq 0$.

Selection Events





Background

Fabian (1974) showed how to specify the continuation region to control the probability of an incorrect selection for $\mathcal{B}(r, \Delta)$. The probability of an incorrect selection is *no greater* if we observe $\mathcal{B}(r, \Delta)$ only at integer times.

Hartmann (1988, 1991) used this result to derive fully sequential procedures for i.i.d. normal data with equal variances (known or unknown).

We extended this work to allow unequal variances and common random numbers.

This talk is about extending such procedures to stationary stochastic processes.

Extension to Stationary Processes

We can standardize the sum of differences in such a way that it is well approximated by $\mathcal{B}(r, \Delta)$ as the sample size gets large.

$$C_{ki}(t) \equiv \frac{\sum_{j=1}^{\lfloor nt \rfloor} \Big((X_{kj} - X_{ij}) - nt(\mu_k - \mu_i) \Big)}{v_{ki}\sqrt{n}}$$

for $0 \le t \le 1$, where

$$v_{ki}^2 \equiv \lim_{n \to \infty} n \, \mathsf{Var} \left[\bar{X}_k(n) - \bar{X}_i(n) \right]$$

<u>FCLT</u>: As $n \to \infty$, $C_{ki}(t) \xrightarrow{D} B(t), 0 \le t \le 1$.

Since v_{ki}^2 is unknown, we employ a strongly consistent estimator—based on "batching"—that is updated as we continue sampling.

We can show that $\lim_{\delta \to 0} \Pr\{\text{CS}\} \ge 1 - \alpha$ if the FCLT holds.

The Basic Procedure

Setup: Select $1-\alpha$, δ, initial sample size n_0 and batch size m_0; choose a batching sequence m_r.

Initialization: Let $I = \{1, 2, \ldots, k\}$ and determine h

Obtain $X_{ij}, j = 1, 2, \ldots, n_0$, $i = 1, 2, \ldots, k$

Set $r = n_0$ and $m_r = m_0$

Update: If m_r has changed, then for all $i \neq \ell$ compute estimates of the asymptotic variance $m_r V_{i\ell}^2(r)$

Let $N_{i\ell}(r) = \left\lfloor \frac{h^2 m_r V_{i\ell}^2(r)}{\delta^2} \right\rfloor$

$N_i(r) = \max_{\ell \neq i} N_{i\ell}(r)$

If $r > \max_i N_i(r)$, stop and select best. Otherwise go to **Screening**.

Screening: Set $I^{\text{old}} = I$. Let

$$\begin{aligned} I &= \Big\{ i : i \in I^{\text{old}} \text{ and} \\ &\quad \bar{X}_i(r) \geq \bar{X}_\ell(r) - W_{i\ell}(r), \forall \ell \in I^{\text{old}}, \ell \neq i \Big\} \end{aligned}$$

where

$$W_{i\ell}(r) = \max\left\{0, \frac{\delta}{2r}\left(\frac{h^2 m_r V_{i\ell}^2(r)}{\delta^2} - r\right)\right\}$$

Stopping Rule: If $|I| = 1$, then stop. Or if $r = \max_i N_i + 1$, then stop.

Otherwise, take $X_{i,r+1}$ for $i \in I$, set $r = r + 1$ and go to **Screening**.

Empirical Results

M/M/1 processes with $k = 5$, $n_0 = 24000$ and $\rho = 0.9$

Sample average of total basic observations ($\times 10^5$)

batch size	$\mathcal{R}+$	$\mathcal{KN}+$	$\mathcal{KN}++$
12000	109.54	83.66	9.10
8000	52.46	36.10	8.11
6000	30.86	24.26	7.44
4800	26.40	19.09	7.01
4000	22.95	16.27	6.81
3000	20.29	12.94	6.29
2400	18.74	11.20	6.01
2000	17.52	10.15	6.00
1600	16.64	9.02	5.51
1000	13.95	6.98	4.91

Estimated PCS (0.95 nominal)

batch size	$\mathcal{R}+$	$\mathcal{KN}+$	$\mathcal{KN}++$
12000	0.984	0.989	0.949
8000	0.940	0.968	0.932
6000	0.891	0.943	0.914
4800	0.878	0.932	0.920
4000	0.864	0.927	0.894
3000	0.849	0.908	0.904
2400	0.831	0.898	0.900
2000	0.834	0.890	0.889
1600	0.835	0.878	0.882
1000	0.817	0.845	0.844