Proceedings of the 2008 Winter Simulation Conference
S. J. Mason, R. R. Hill, L. Mönch, O. Rose, T. Jefferson, J. W. Fowler eds.

RUN-LENGTH VARIABILITY OF TWO-STAGE MULTIPLE COMPARISONS WITH THE BEST FOR STEADY-STATE SIMULATIONS AND ITS IMPLICATIONS FOR CHOOSING FIRST-STAGE RUN LENGTHS

Marvin K. Nakayama

Computer Science Department
New Jersey Institute of Technology
Newark, N.J. 07102, U.S.A.

ABSTRACT

We analyze the asymptotic behavior of two-stage procedures for multiple comparisons with the best (MCB) for comparing the steady-state means of alternative systems using simulation. The two procedures we consider differ in how they estimate the variance parameters of the alternatives in the first stage. One procedure uses a consistent estimator, and the other employs an estimator based on one of Schruben's standardized time series (STS) methods. While both methods lead to mean total run lengths that are of the same asymptotic order of magnitude, the limiting variability of the run lengths is strictly smaller for the method based on a consistent variance estimator. We also provide some analysis showing how to choose the first-stage run length.

1 INTRODUCTION

One problem often faced by simulation analysts is comparing alternative systems relative to a steady-state performance measure. For example, there may be 10 alternative designs for a manufacturing system, and we want to compare the designs relative to their steady-state throughputs.

To address this problem, we examine procedures for *multiple comparisons with the best* (MCB, Hsu 1984). Specifically, suppose there are k alternative systems, where system i has steady-state mean θ_i, and assume that larger means are better. MCB constructs simultaneous confidence intervals for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \ldots, k$. Note that if $\theta_i - \max_{j \neq i} \theta_j > 0$ for some i, then system i is the best. On the other hand, if $\theta_i - \max_{j \neq i} \theta_j < 0$, then system i is not the best, but this quantity indicates how close system i is to the best. This information could be useful when ultimately deciding which design to implement when secondary considerations (e.g., ease of maintenance) are taken into account.

We examine two-stage MCB procedures to construct simultaneous confidence intervals having absolute-width parameter $\delta > 0$, which is pre-specified by the user; i.e., the user wants the confidence intervals to have half-width at most δ. In the context of comparing means of normally distributed populations, Matejcik and Nelson (1995) show that two-stage MCB is closely related to an indifference-zone selection procedure (Bechhofer 1954). The latter seeks to correctly identify the best system with at least a pre-specified probability when the best system is at least δ better than the next best, where δ is pre-specified by the user.

Much of the previous work on MCB and selection procedures focuses on comparing means of normally distributed populations with independent and identically distributed (i.i.d.) sampling used within each population (Hochberg and Tamhane 1987, Swisher, Jacobson, and Yucesan 2003, Kim and Nelson 2006b). However, steady-state means, as we consider, are averages (or time averages) of stochastic processes, which typically exhibit autocorrelations and have non-normal output. Damerdji and Nakayama (1999), Kim and Nelson (2006a) and Nakayama (2008a) develop and analyze procedures for comparing steady-state means, and a focus of these papers is establishing asymptotic validity as $\delta \to 0$, where δ represents the desired width parameter of the MCB intervals or the indifference-zone parameter.

We compare two two-stage MCB procedures for steady-state means. In both methods, we simulate a first stage whose size is proportional to $1/\delta^2$, from which an estimate of the variance parameter of each system is computed. In the first method, which we call MCB-CVE and is from Nakayama (2008a), we use a consistent estimator of the variance. Such estimators can be constructed using, e.g., the regenerative, autoregressive or spectral method under various conditions; see Law (2006) for overviews of these methods. The second method, which we call MCB-STS and is a slight modification of a procedure in Damerdji and Nakayama (1999), applies one of Schruben's (1983) standardized time series (STS) estimators of the variance.

We compare the total run lengths of MCB-CVE and MCB-STS in terms of their means and variances. We find that the mean run lengths are of the same asymptotic order,

978-1-4244-2708-6/08/$25.00

but the limiting variability for MCB-CVE is strictly smaller than for MCB-STS.

Also, for MCB-CVE we analyze the effect of the first-stage run length, which the user specifies, on the mean and variability of the total run length. We use this to determine the "best" first-stage run length by solving a bicriteria optimization problem, in which we first minimize the expected total run length, and then of those optimal first-stage sizes, we break the tie by choosing the one that minimizes the variance of the total run length.

The rest of the paper is organized as follows. Section 2 describes the mathematical framework we adopt, and we present the two MCB procedures MCB-CVE and MCB-STS in Section 3. Section 4 contains a comparison of their total run lengths, and we provide analysis on choosing the first-stage run length of MCB-CVE in Section 5. We summarize our findings in Section 6. All proofs are given in Nakayama (2008b).

2 MATHEMATICAL FRAMEWORK

Suppose there are k systems to compare. Let $X_i = [X_i(t) : t \geq 0]$ be a stochastic process on a state space S_i denoting the evolution over time of system i. We assume that $X_i \in D_{S_i}[0,\infty)$, where $D_{S_i}[0,\infty)$ is the space of S_i-valued functions on $[0,\infty)$ that are right continuous and have left limits (Billingsley 1999). For each i, let $f_i : S_i \to \Re$ be a "reward function." We assume the following holds for some finite parameter θ_i:

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f_i(X_i(s))\,ds = \theta_i \quad \text{a.s.},$$

where a.s. stands for "almost surely." Thus, θ_i is the steady-state mean reward of system i, and our goal is to compare the k alternatives in terms of $\theta_1,\ldots,\theta_k$, where we assume larger θ_i is better.

For each system i, a natural estimator of θ_i based on a simulation of length t of process X_i is

$$\widehat{\theta}_i(t) = \frac{1}{t}\int_0^t f_i(X_i(s))\,ds.$$

We assume that the estimation process $\widehat{\theta}_i = [\widehat{\theta}_i(t) : t \geq 0]$ satisfies a functional central limit theorem (FCLT), which we now describe. Let $\Rightarrow$ denote weak convergence (Billingsley 1999). For each system i and each $n > 0$, define the process $U_{i,n} = [U_{i,n}(t) : t \geq 0]$ with $U_{i,n}(t) = n^{1/2}t(\widehat{\theta}_i(nt) - \theta_i)$.

Assumption 1 *The k systems are simulated independently, and for each system i, there exists a positive finite constant σ_i such that*

$$U_{i,n} \Rightarrow \sigma_i B_i \tag{1}$$

as $n\to\infty$, where $B_i = [B_i(t) : t \geq 0]$ is a standard Brownian motion.

Whitt (2002), Section 4.4, discusses examples of processes satisfying the FCLT in Assumption 1, which is slightly stronger than an ordinary central limit theorem (CLT). For example, the FCLT holds (under various conditions) for Markov chains, associated processes, martingales and stochastic processes satisfying mixing conditions, which are a form of asymptotic independence.

We now consider the special case when each process X_i is regenerative (Crane and Iglehart 1975). Let $A_{i,j}$, $j \geq 0$, be the sequence of regeneration epochs of system i, with $0 \leq A_{i,0} < A_{i,1} < A_{i,2} < \cdots$. For $j = 1,2,\ldots$, let $\tau_{i,j} = A_{i,j} - A_{i,j-1}$ be the length of the jth cycle of system i. Also, define $Y_{i,j} = \int_{A_{i,j-1}}^{A_{i,j}} f_i(X_i(s))\,ds$ to be the cumulative reward over the jth cycle of system i. Assume that $E[\tau_{i,1}] < \infty$ and that there exists a finite constant θ_i such that $E[Y_{i,1} - \theta_i\tau_{i,1}] = 0$ and $E[(Y_{i,1} - \theta_i\tau_{i,1})^2] < \infty$. Also, for a function $h : S_i \to \Re$ and $j \geq 1$, let $W_{i,j}(h) = \sup_{0\leq s\leq \tau_{i,j}} \left|\int_0^s h(X_i(A_{i,j-1}+u))du\right|$, and assume $r^2 P\{W_{i,1}(f_i - \theta_i) > r\} \to 0$ as $r \to \infty$, where $f_i - \theta_i$ denotes the function whose value for $x \in S_i$ is $f_i(x) - \theta_i$. Then $\theta_i = E[Y_{i,1}]/E[\tau_{i,1}]$ and the FCLT in (1) holds with $\sigma_i^2 = E[(Y_{i,1} - \theta_i\tau_{i,1})^2]/E[\tau_{i,1}]$; see Glynn and Whitt (1993).

Typically, the parameter σ_i in Assumption 1 is unknown and needs to be estimated. Below we discuss two broad categories of techniques to accomplish this.

2.1 Consistent Variance Estimation

First we consider consistent estimation of σ_i^2. Let $V_i = [V_i(t) : t \geq 0]$ be an estimation process of σ_i^2, where $V_i(t)$ is the estimator of σ_i^2 from a simulation up to time t. We assume that V_i is consistent in the following sense:

Assumption 2 $V_i(t) \Rightarrow \sigma_i^2$ *as $t \to \infty$.*

We can construct an estimator V_i satisfying Assumption 2 under various conditions on the original process X_i and the reward function f_i. For example, consider again our previous example when each X_i is regenerative. Define $N_i(t) = \sup\{j \geq 0 : A_{i,j} \leq t\}$, which is the number of regenerative cycles that process i completes by time t. Then the variance estimator

$$V_i(t) = \frac{1}{t}\sum_{j=1}^{N_i(t)} \left[Y_{i,j} - \widehat{\theta}_i(t)\tau_{i,j}\right]^2$$

satisfies Assumption 2 under the conditions we provided earlier; see Glynn and Iglehart (1993) for details.

Other examples of variance estimators satisfying Assumption 2 (under various conditions) include spectral estimators (Damerdji 1991), autoregressive estimators (Fishman 1978, p. 252), and various batch means and batched area estimators in which the number of batches $m \to \infty$ at an appropriate rate as the run length increases (Damerdji 1994).

Law (2006) and Bratley, Fox, and Schrage (1987) provide overviews of these techniques.

2.2 Variance Estimation Using STS

An alternative approach to consistently estimating σ_i^2 is to use one of Schruben's (1983) standardized time series methods. To do this, for each system i and $n > 0$, we define the integrated process $\bar{X}_{i,n} = [\bar{X}_{i,n}(t) : 0 \le t \le 1] \in C[0,1]$, where

$$\bar{X}_{i,n}(t) = n\,\widehat{\theta}_i(nt) = \frac{1}{n}\int_0^{nt} f_i(X_i(s))\,ds$$

and $C[0,1]$ is the space of continuous real-valued functions on $[0,1]$. Suppose that we have run a simulation up to time t. We then implement an STS method by applying to the integrated process $\bar{X}_{i,t}$ a function $g : C[0,1] \to \Re$ satisfying conditions given below, and we take

$$V_i'(t) = t g^2(\bar{X}_{i,t}) \tag{2}$$

as our STS estimator of σ_i^2 from the simulation of system i up to time t. Also, we can show that

$$V_i'(t) \Rightarrow g^2(\sigma_i B_i) = \sigma_i^2 g^2(B_i) \tag{3}$$

as $t \to \infty$, where B_i is a standard Brownian motion on $C[0,1]$.

We now provide the conditions the STS function g must satisfy. Let B be a standard Brownian motion on $C[0,1]$, and we assume the following hold:

C1. $g(\beta x) = \beta g(x)$ for all $\beta \in \Re$ with $\beta > 0$ and $x \in C[0,1]$;
C2. $g(x - \beta e) = g(x)$ for all $\beta \in \Re$ and $x \in C[0,1]$, where $e \in C[0,1]$ with $e(t) = t$;
C3. $P\{g(B) > 0\} = 1$;
C4. $P\{B \in D(g)\} = 0$, where $D(g)$ denotes the set of discontinuities of g;
C5. $g(B)$ has a density function f_g with respect to Lebesgue measure and $f_g(\beta) > 0$ for all $\beta \in \Re$ with $\beta > 0$.

Glynn and Iglehart (1990) give Conditions C1–C4, which we explain below, as the definition of STS functions, and Nakayama (1994) and Damerdji and Nakayama (1999) added a weaker version of C5 to handle two-stage STS procedures. Instead of C5, Damerdji and Nakayama (1999) assume that $P\{g(B) \in A\} = 0$ whenever A is a countable set. However, all known STS functions g satisfy both this condition and C5, and we believe (but do not have a formal proof) that Conditions C1–C4 imply C5. For example, most known STS methods result in $g^2(B)$ having a chi-squared distribution, so C5 holds. Calvin and Nakayama (2006) consider a maximum estimator, for which $g(B)$ has a Weibull distribution instead of the square root of a chi-square, but this also satisfies C5.

We now explain conditions C1–C4. Condition C1 states that if we scale a process by a constant, then evaluating the function g at the scaled process gives the same value as applying g to the original process and scaling the result by the same constant. Note that we used C1 in the second step of (3). C2 ensures that translating a process by a constant does not alter the value of g. The parameter σ_i is positive, and C3 guarantees that asymptotically the STS estimator of σ_i is a.s. positive. C4 is a technical condition required for the proofs.

Glynn and Iglehart (1990) provide examples of specific g functions corresponding to various STS methods, including g functions for batch means and Schruben's (1983) area estimator. For example, for batch means with $m \ge 2$ batches, the STS function applied to $x \in C[0,1]$ is

$$g_{\rm bm}(x) = \left[\frac{m}{m-1}\sum_{j=1}^{m}\left(x\left(\frac{j}{m}\right) - x\left(\frac{j-1}{m}\right) - \frac{x(1)}{m}\right)^2\right]^{1/2}$$

and

$$V_i'(t) = \frac{t}{m-1}\sum_{j=1}^{m}\left(\frac{1}{t/m}\int_{(j-1)t/m}^{jt/m} f_i(X_i(s))\,ds - \widehat{\theta}_i(t)\right)^2.$$

When applied to a standard Brownian motion B on the unit interval, we can show that $g_{\rm bm}^2(B)$ is distributed as $\chi_{m-1}^2/(m-1)$, where χ_{m-1}^2 is a chi-squared random variable with $m-1$ degrees of freedom, so $E[g_{\rm bm}^2(B)] = 1$. Hence, for batch means, since $g_{\rm bm}(B_i)$ is not deterministically equal to 1, V_i' is not consistent by (3), so Assumption 2 does not hold. In fact, Glynn and Whitt (1991) show that when the number m of batches is held fixed as the run length t grows, it is impossible to combine the batch means in some way to obtain a consistent variance estimator.

In this paper, we will allow for any STS function g (possibly with batching), which we assume satisfies the following:

Assumption 3 *The STS function g satisfies $E[g^2(B)] = 1$, where B is a standard Brownian motion.*

The assumption that $E[g^2(B)] = 1$, which we previously saw holds for batch means, is not restrictive as long as $E[g^2(B)] < \infty$. It ensures that the limiting variance estimator is unbiased by (3).

3 TWO-STAGE MCB PROCEDURES

Our goal is to construct (simultaneous) MCB intervals for $\theta_i - \max_{j \neq i} \theta_j$, $i = 1, \ldots, k$, where the intervals have a pre-specified absolute-width parameter $\delta > 0$.

3.1 Procedure Using Consistent Variance Estimators

An approach to do this, given in Nakayama (2008a), is the below two-stage procedure, which is related to a method of Rinott (1978). The first stage estimates the unknown variance parameter σ_i^2 using the consistent variance estimator V_i, which is used to determine the total run length required.

Procedure MCB-CVE

1. Specify the confidence level $1-\alpha$, and the desired absolute-width parameter δ of the MCB confidence intervals. Also, for each system i, specify the first-stage run length

$$T_{0,i} = \zeta_i/\delta^2, \text{ where } \zeta_i > 0 \text{ is any constant.} \quad (4)$$

   The ζ_i may be unequal for different systems.

2. For each system $i = 1, \ldots, k$, simulate for a run length of $T_{0,i}$, where the k systems are simulated independently.
3. For each system i, compute the total run length required as

$$T_i(\delta) = \max\left(T_{0,i}, \frac{\gamma^2 V_i(T_{0,i})}{\delta^2}\right), \quad (5)$$

   where the constant $\gamma \equiv \gamma(k, 1-\alpha) = \sqrt{2} z_{(1-\alpha)^{1/(k-1)}}$, with z_β satisfying $P\{N(0,1) \leq z_\beta\} = \beta$ for $0 < \beta < 1$ and $N(0,1)$ a standard (mean 0 and variance 1) normal random variable, and V_i is any estimator satisfying Assumption 2.

4. For each system i, continue to simulate from time $T_{0,i}$ to $T_i(\delta)$, where the k systems are simulated independently, and form the point estimator $\widetilde{\theta}_i(\delta) = \widehat{\theta}_i(T_i(\delta))$ of θ_i.
5. Use the absolute-width parameter δ to construct simultaneous MCB confidence intervals

$$I_i(\delta) = \left[-\left(\widetilde{\theta}_i(\delta) - \max_{j\neq i} \widetilde{\theta}_j(\delta) - \delta\right)^-, \left(\widetilde{\theta}_i(\delta) - \max_{j\neq i} \widetilde{\theta}_j(\delta) + \delta\right)^+\right],$$

   $i = 1, \ldots, k$, for $\theta_i - \max_{j\neq i} \theta_j$, $i = 1, \ldots, k$, respectively, where $-(\beta)^- = \min(\beta, 0)$ and $(\beta)^+ = \max(\beta, 0)$.

Nakayama (2008a) establishes the asymptotic validity (as $\delta \to 0$) of the resulting MCB intervals; i.e.,

$$\lim_{\delta\to 0} P\left\{\theta_i - \max_{j\neq i} \theta_j \in I_i(\delta), i = 1, \ldots, k\right\} > 1-\alpha,$$

so the asymptotic joint coverage of the MCB intervals is greater than the nominal level.

3.2 Procedure Using STS Variance Estimators

We now modify Procedure MCB-CVE to estimate σ_i^2 using a standardized time series estimator V_i' defined in (2) instead of the consistent variance estimator V_i. Our two-stage MCB procedure based on STS (which we call MCB-STS) is the same as Procedure MCB-CVE except the total run length for system i in (5) is changed to

$$T_i'(\delta) = \max\left(T_{0,i}, \frac{\gamma'^2 V_i'(T_{0,i})}{\delta^2}\right), \quad (6)$$

where γ' is a constant that we will define shortly and $T_{0,i}$ is the first-stage run length for system i; see (4). Also, in steps 4 and 5, we simulate system i up to time $T_i'(\delta)$ rather than to $T_i(\delta)$, and the point estimator is now $\widetilde{\theta}_i(\delta) = \widehat{\theta}_i(T_i'(\delta))$.

We now define the constant γ' in (6). Suppose that we want to construct STS MCB intervals having asymptotic joint confidence level at least $1-\alpha$. Then $\gamma' = \gamma'(k, 1-\alpha, g)$ in (6) is chosen to satisfy

$$E\left[\prod_{i=1}^{k-1} \Phi\left(\frac{\gamma'}{\left[(1/g^2(B_i)) + (1/g^2(B_k))\right]^{1/2}}\right)\right] = 1-\alpha, \quad (7)$$

where $B_1, \ldots, B_k$ are i.i.d. standard Brownian motions on the unit interval.

When g is the batch means function g_{bm} with $m \geq 2$ batches, the parameter $\gamma'(k, 1-\alpha, g_{\text{bm}})$ in (7) is Rinott's (1978) constant from his two-stage selection procedure for comparing the means of independent normal populations when the first-stage sample size for each population is m. Bechhofer, Santner, and Goldsman (1995) provide tables of values for γ' for different values of k, m and $1-\alpha$.

Damerdji and Nakayama (1999) present a variant of MCB-STS that determines the total number of batches to simulate rather than the total run length, as we do now, and establish its asymptotic validity (as $\delta \to 0$) when the first-stage run length is $T_0 = \zeta/\delta^2$ for some constant ζ which is the same for all systems i. We now also allow different first-stage run lengths for each system i, given by $T_{0,i} = \zeta_i/\delta^2$ in (4), where $\zeta_i > 0$ is any constant. We can slightly modify the proof in Damerdji and Nakayama (1999) to establish the asymptotic validity of MCB-STS.

4 COMPARISON OF RUN LENGTHS OF MCB-CVE AND MCB-STS

We now compare MCB-CVE and MCB-STS in terms of asymptotic properties of their total run lengths $T_i(\delta)$ and $T_i'(\delta)$ for each system i. To do this, we will also study the *potential* total run lengths, which we define as the second terms in the maximums in (5) and (6). Specifically, these are

$$\bar{T}_i(\delta) = \frac{\gamma^2 V_i(T_{0,i})}{\delta^2} \quad \text{and} \quad \bar{T}_i'(\delta) = \frac{\gamma'^2 V_i'(T_{0,i})}{\delta^2} \tag{8}$$

for MCB-CVE and MCB-STS, respectively. Thus,

$$T_i(\delta) = \max(T_{0,i}, \bar{T}_i(\delta)) \quad \text{and} \quad T_i'(\delta) = \max(T_{0,i}, \bar{T}_i'(\delta)). \tag{9}$$

4.1 Expected Run Lengths

We first examine the limiting means of the run lengths of the two methods. The proof of the following is given in Nakayama (2008a).

Theorem 1 *Suppose Assumptions 1–3 hold. Then the following hold for each system i as* $\delta \to 0$*:*

(i) *For MCB-CVE,*

$$\delta^2 T_i(\delta) \Rightarrow \tau_i, \tag{10}$$
$$\delta^2 \bar{T}_i(\delta) \Rightarrow \gamma^2\sigma_i^2, \tag{11}$$

where

$$\tau_i = \max(\zeta_i, \gamma^2\sigma_i^2), \tag{12}$$

which is a finite positive constant. If $\{V_i(t) : t > 0\}$ *is uniformly integrable, then*

$$\delta^2 E[T_i(\delta)] \to \tau_i, \tag{13}$$
$$\delta^2 E[\bar{T}_i(\delta)] \to \gamma^2\sigma_i^2. \tag{14}$$

(ii) *For MCB-STS,*

$$\delta^2 T_i'(\delta) \Rightarrow \tau_i', \tag{15}$$
$$\delta^2 \bar{T}_i'(\delta) \Rightarrow \gamma'^2\sigma_i^2 g^2(B_i),$$

where

$$\tau_i' = \max[\zeta_i, \gamma'^2\sigma_i^2 g^2(B_i)] \tag{16}$$

and τ_i' *is nondegenerate with* $\tau_i' \geq \zeta_i$ *a.s. If* $\{V_i'(t) : t > 0\}$ *is uniformly integrable, then*

$$\delta^2 E[T_i'(\delta)] \to E[\tau_i'], \tag{17}$$
$$\delta^2 E[\bar{T}_i'(\delta)] \to \gamma'^2\sigma_i^2,$$

where $\zeta_i < E[\tau_i'] < \infty$*. Moreover,* $E[\tau_i'] > \tau_i$*.*

Thus, we see that for both MCB-CVE and MCB-STS, the total run length and its expectation for each system i are all asymptotically of order $1/\delta^2$; see (10), (13), (15), and (17). The same also hold for the potential total run lengths of each method. However, even though MCB-CVE and MCB-STS have expected total run lengths that are of the same order of magnitude, the last part of Theorem 1(ii) shows that MCB-CVE asymptotically has strictly smaller expected total run lengths.

4.2 Variability of Run Lengths

We now examine the variability of the total run lengths $T_i(\delta)$ and $T_i'(\delta)$. For MCB-CVE the limit τ_i in (10) is deterministic. For MCB-STS with a fixed number $m \geq 1$ of batches, the limit τ_i' in (15) is nondegenerate. Hence, we see that STS-MCB has asymptotically more variable total run lengths than MCB-CVE.

Another way to quantify this is by comparing the limiting variances (appropriately normalized) of $T_i(\delta)$ and $T_i'(\delta)$. We first give the asymptotic variability of both the total run length and the potential total run length for MCB-CVE.

Theorem 2 *Suppose that Assumptions 1 and 2 hold. Also, assume that* $\sqrt{V_i}$ *satisfies the following CLT:*

$$\sqrt{t}(\sqrt{V_i(t)} - \sigma_i) \Rightarrow N(0, \psi_i^2) \tag{18}$$

as $t \to \infty$ *for some finite constant* $\psi_i > 0$*, and assume that* $\{t(V_i(t) - \sigma_i^2)^2 : t > 0\}$ *is uniformly integrable. Then the following hold for each system i as* $\delta \to 0$*:*

(i) *the potential total run length* $\bar{T}_i(\delta)$ *satisfies*

$$\delta(\bar{T}_i(\delta) - E[\bar{T}_i(\delta)]) \Rightarrow N(0, \kappa_i^2), \tag{19}$$
$$\delta^2 Var(\bar{T}_i(\delta)) \to \kappa_i^2, \tag{20}$$

where $\kappa_i = 2\gamma^2\sigma_i\psi_i/\sqrt{\zeta_i}$*, so* $0 < \kappa_i < \infty$*;*

(ii) *if* $\zeta_i < \gamma^2\sigma_i^2$*, then*

$$\delta(T_i(\delta) - E[T_i(\delta)]) \Rightarrow N(0, \kappa_i^2),$$
$$\delta^2 Var(T_i(\delta)) \to \kappa_i^2; \tag{21}$$

(iii) *if* $\zeta_i = \gamma^2\sigma_i^2$*, then*

$$\delta(T_i(\delta) - E[T_i(\delta)]) \Rightarrow Y,$$

$$\delta^2 Var(T_i(\delta)) \quad \rightarrow \quad 2(1-\pi^{-1})\gamma^2\psi_i^2, \quad (22)$$

where Y has distribution function H with

$$H(y) = \begin{cases} 0 & \text{for } y < y_0, \\ \Phi((y-y_0)/(2\gamma\psi_i)) & \text{for } y \geq y_0, \end{cases} \quad (23)$$

and $y_0 = 2\gamma\psi_i/\sqrt{2\pi}$;

(iv) *if $\zeta_i > \gamma^2\sigma_i^2$, then*

$$\delta(T_i(\delta) - E[T_i(\delta)]) \quad \Rightarrow \quad 0,$$
$$\delta^2 Var(T_i(\delta)) \quad \rightarrow \quad 0. \quad (24)$$

When $V_i(t)$ is the regenerative variance estimator, Proposition 4.34 of Glynn and Iglehart (1990) provides moment conditions to ensure (18) holds.

We now want to interpret the results in Theorem 2. If all the σ_i were known, rather than use a two-stage MCB procedure, we would instead employ a single-stage procedure with run length $\gamma^2\sigma_i^2/\delta^2$ for each system i. In our two-stage procedure MCB-CVE, the first-stage run length is $T_{0,i} = \zeta_i/\delta^2$ from (4), where the constant ζ_i is specified by the user, and parts (ii)–(iv) of Theorem 2 consider a partition of the possible values of ζ_i into three sets: less than, equal to, and greater than $\gamma^2\sigma_i^2$. These results show that the choice of the value for ζ_i has a significant impact on the variability of the total run length $T_i(\delta)$.

To understand this effect, because of (9), it will be helpful to first get a handle on the potential total run length $\bar{T}_i(\delta)$ in (8). Note that (11) implies

$$\bar{T}_i(\delta) \approx \gamma^2\sigma_i^2/\delta^2 \text{ for small } \delta. \quad (25)$$

The CLT in (19) refines this by showing that when δ is small, $\bar{T}_i(\delta)$ is roughly normally distributed about its mean $E[\bar{T}_i(\delta)]$, which is approximately $\gamma^2\sigma_i^2/\delta^2$ for small δ by (14).

Now consider the case when $\zeta_i < \gamma^2\sigma_i^2$, as in Theorem 2(ii). Since (25) implies $\bar{T}_i(\delta) \approx \gamma^2\sigma_i^2/\delta^2 > \zeta_i/\delta^2 = T_{0,i}$, the second stage is almost always needed in the limit by (9). Thus, the first-stage run length is "too short" by itself, and the total run length $T_i(\delta)$ is almost always equal to the potential run length $\bar{T}_i(\delta)$. Hence, the asymptotic variability of $T_i(\delta)$ is that of $\bar{T}_i(\delta)$; see (20) and (21).

Theorem 2(iii) examines when $\zeta_i = \gamma^2\sigma_i^2$. Recall that for small δ, the potential total run length $\bar{T}_i(\delta)$ has roughly a normal distribution centered at $E[\bar{T}_i(\delta)] \approx \gamma^2\sigma_i^2/\delta^2$, which is exactly the first-stage length $T_{0,i}$ in this case. Thus, the complementary events $\{\bar{T}_i(\delta) > T_{0,i}\}$ and $\{\bar{T}_i(\delta) \leq T_{0,i}\}$ each have approximately probability $1/2$ for small δ. By (9), the total run length $T_i(\delta) = \bar{T}_i(\delta)$ when $\bar{T}_i(\delta) > T_{0,i}$, so a second stage is required, and it results in variability in the total run length because of the randomness in $\bar{T}_i(\delta)$. When $\bar{T}_i(\delta) \leq T_{0,i}$, the second stage is not needed, in which case the total run length $T_i(\delta)$ exhibits no variability since it is just $T_{0,i}$, which is deterministic. Consequently, for small δ, the total run length $T_i(\delta)$ has approximately a distribution $\hat{H}$ that has a point mass at $T_{0,i}$ with probability $1/2$ (arising from when $\bar{T}_i(\delta) \leq T_{0,i}$) and the rest is the positive part of a normal (corresponding to when $\bar{T}_i(\delta) > T_{0,i}$). Shifting the distribution $\hat{H}$ by its expectation leads to the distribution H in (23). It can then be shown the variance of the total run length satisfies (22).

When $\zeta_i > \gamma^2\sigma_i^2$, as in Theorem 2(iv), the first stage is "too long," so the second stage is almost never needed. Thus, most of the time, the total run length is just $T_{0,i}$, which is deterministic, so there is little variability; see (24).

We now examine the variability of the run length for MCB-STS.

Theorem 3 *Suppose that Assumption 1 and 3 hold and that $\{V_i'^2(t) : t > 0\}$ is uniformly integrable. Then for each system i, the following hold as $\delta \to 0$:*

(i) *$\delta^4 Var(\bar{T}_i'(\delta)) \to Var(\bar{\tau}_i')$, where $\bar{\tau}_i' = \gamma'^2\sigma_i^2 g^2(B_i)$ and $0 < Var(\bar{\tau}_i') < \infty$;*

(ii) *$\delta^4 Var(T_i'(\delta)) \to Var(\tau_i')$, where τ_i' is defined in (16) and $0 < Var(\tau_i') < \infty$.*

We now compare the results of Theorems 2 and 3. In terms of asymptotic variance of total run lengths, we see that the variability of MCB-CVE run lengths are $O(1/\delta^2)$ by (21), (22) and (24), whereas the variability of MCB-STS run lengths are of order $1/\delta^4$ by Theorem 3(ii). Thus, MCB-CVE run lengths are asymptotically less variable than MCB-STS run lengths. The reason for the difference originates from the fact that MCB-CVE uses consistent estimators of σ_i^2 (Assumption 2), whereas the variance estimators in MCB-STS are not consistent (for a fixed number of batches); see Section 2.2.

5 CHOOSING A FIRST-STAGE RUN LENGTH FOR MCB-CVE

We now apply the results from the previous section to investigate the choice of ζ_i for MCB-CVE. Recall the first-stage run length is $T_{0,i} = \zeta_i/\delta^2$ from (4), where the constant $\zeta_i > 0$ is arbitrary and specified by the user, and Theorem 1 shows the asymptotic impact of the choice of ζ_i on the expected total run length $T_i(\delta)$ for MCB-CVE. Specifically, from (12) and (13) we see that $\delta^2 E[T_i(\delta)] \to \gamma^2\sigma_i^2$ if $\zeta_i \leq \gamma^2\sigma_i^2$, and $\delta^2 E[T_i(\delta)] \to \zeta_i$ if $\zeta_i > \gamma^2\sigma_i^2$. Thus, we have the following result.

Corollary 4 *For MCB-CVE, the values of $\zeta_i \in (0,\infty)$ that minimize $\lim_{\delta\to 0} \delta^2 E[T_i(\delta)]$ are $\zeta_i \in (0, \gamma^2\sigma_i^2]$.*

Corollary 4 suggests that the user should choose ζ_i such that $\zeta_i \leq \gamma^2\sigma_i^2$ to minimize the expected total run length. Of these values of ζ_i, we now want to determine which leads to the smallest asymptotic variance of the total run length. The next result follows from Theorem 2.

Corollary 5 *For MCB-CVE, the value of* $\zeta_i \in (0, \gamma^2\sigma_i^2]$ *that minimizes* $\lim_{\delta\to 0}\delta^2 Var(T_i(\delta))$ *is* $\zeta_i = \gamma^2\sigma_i^2$.

Thus, Corollaries 4 and 5 provide the optimal solution $\zeta_i = \gamma^2\sigma_i^2$ to a bicriteria optimization problem to find the best first-stage run length, by first choosing ζ_i to minimize the expected total run length, and next breaking the tie by minimizing the variance of the total run length. Of course, the user typically does not know the value of σ_i^2 (which is why a two-stage procedure is being applied in the first place), so the optimal choice of ζ_i cannot be directly implemented in practice since it depends on σ_i^2. However, it does provide interesting insight into how one should try to choose ζ_i. Moreover, if we have an approximation for σ_i^2, then we could use this to determine a good value for ζ_i. For example, Whitt (2006) and references therein provide such variance approximations for various queueing and loss models.

6 CONCLUSIONS

We compared the total run lengths for MCB-CVE and MCB-STS in terms of their means and variances. We found that although the two approaches have mean run lengths that are of the same order $1/\delta^2$, where δ is the absolute-width parameter of the MCB intervals, the variability of the MCB-STS run lengths are strictly greater than those of MCB-CVE. We also provided some analysis suggesting how to choose the first-stage run length for MCB-CVE.

REFERENCES

Bechhofer, R. E. 1954. A single-sample multiple-decision procedure for ranking means of normal populations with known variances. *Annals of Mathematical Statistics* 25:16–39.

Bechhofer, R. E., T. J. Santner, and D. M. Goldsman. 1995. *Design and analysis of experiments for statistical selection, screening, and multiple comparisons*. New York: John Wiley & Sons.

Billingsley, P. 1999. *Convergence of probability measures*. Second ed. New York: John Wiley & Sons.

Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A guide to simulation*. Second ed. New York: Springer-Verlag.

Calvin, J. M., and M. K. Nakayama. 2006. Permuted standardized time series for steady-state simulations. *Mathematics of Operations Research* 31:351–368.

Crane, M., and D. L. Iglehart. 1975. Simulating stable stochastic systems, III: Regenerative processes and discrete-event simulations. *Operations Research* 23:33–45.

Damerdji, H. 1991. Strong consistency and other properties of the spectral variance estimator. *Management Science* 37:1424–1440.

Damerdji, H. 1994. Strong consistency of the variance estimator in steady-state simulation output analysis. *Mathematics of Operations Research* 19:494–512.

Damerdji, H., and M. K. Nakayama. 1999. Two-stage multiple-comparison procedures for steady-state simulations. *ACM Transactions on Modeling and Computer Simulations* 9:1–30.

Fishman, G. S. 1978. *Principles of discrete event simulation*. New York: Wiley.

Glynn, P. W., and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15:1–16.

Glynn, P. W., and D. L. Iglehart. 1993. Conditions for the applicability of the regenerative method. *Management Science* 39:1108–1111.

Glynn, P. W., and W. Whitt. 1991. Estimating the asymptotic variance with batch means. *Operations Research Letters* 10:431–435.

Glynn, P. W., and W. Whitt. 1993. Limit theorems for cumulative processes. *Stochastic Processes and Their Applications* 47:299–314.

Hochberg, Y., and A. C. Tamhane. 1987. *Multiple comparison procedures*. New York: Wiley.

Hsu, J. C. 1984. Constrained simultaneous confidence intervals for multiple comparisons with the best. *Annals of Statistics* 12:1136–1144.

Kim, S.-H., and B. L. Nelson. 2006a. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. *Operations Research* 54:475–488.

Kim, S.-H., and B. L. Nelson. 2006b. Selecting the best system. In *Elsevier Handbooks in Operations Research and Management Science: Simulation*. Amsterdam: Elsevier.

Law, A. M. 2006. *Simulation Modeling and Analysis*. Fourth ed. New York: McGraw-Hill.

Matejcik, F. J., and B. L. Nelson. 1995. Two-stage multiple comparisons with the best for computer simulation. *Operations Research* 43:633–640.

Nakayama, M. K. 1994. Two-stage stopping procedures based on standardized time series. *Management Science* 40:1189–1206.

Nakayama, M. K. 2008a. Asymptotic analysis of two-stage selection and multiple-comparion procedures for simulations. Submitted.

Nakayama, M. K. 2008b. Two-stage multiple comparison procedures for steady-state simulations: Run-length variability and choosing first-stage run lengths. In preparation.

Rinott, Y. 1978. On two-stage selection procedures and related probability-inequalities. *Communications in Statistics—Theory and Methods* A7:799–811.

Schruben, L. W. 1983. Confidence interval estimation using standardized time series. *Operations Research* 31:1090–1108.

Swisher, J. R., S. H. Jacobson, and E. Yucesan. 2003. Discrete-event simulation optimization using ranking, selection, and multiple comparison procedures: A survey. *ACM Transactions on Modeling and Computer Simulation* 13:134–154.

Whitt, W. 2002. *Stochastic-Process Limits: An Introduction to Stochastic-Process Limits and Their Application to Queues*. New York: Springer-Verlag.

Whitt, W. 2006. Analysis for the design of simulation experiments. In *Handbooks in Operations Research and Management Science, Volume 13: Simulation*, ed. S. G. Henderson and B. L. Nelson, 381–413. Amsterdam: North Holland.

AUTHOR BIOGRAPHY

MARVIN K. NAKAYAMA is an associate professor in the Department of Computer Science at the New Jersey Institute of Technology. He received a Ph.D. in operations research from Stanford University. He won second prize in the 1992 George E. Nicholson Student Paper Competition sponsored by INFORMS and is a recipient of a CAREER Award from the National Science Foundation. He is the stochastic models area editor for *ACM Transactions on Modeling and Computer Simulation* and the simulation area editor for *INFORMS Journal on Computing*. His research interests include applied probability, statistics, simulation and modeling. His e-mail address is <marvin@njit.edu>, and his web page is <web.njit.edu/~marvin>.