Multiple-Comparison Procedures for Steady-State Simulations

Marvin K. Nakayama
Department of Computer and Information Science
New Jersey Institute of Technology
Newark, NJ 07102, U.S.A.

Abstract

Suppose that there are $k \geq 2$ different systems (i.e., stochastic processes), where each system has an unknown steady-state mean performance and unknown asymptotic variance. We allow for the asymptotic variances to be unequal and for the distributions of the k systems to be different. We consider the problem of running independent, single-stage simulations to make multiple comparisons of the steady-state means of the different systems. We derive asymptotically valid (as the run lengths of the simulations of the systems tend to infinity) simultaneous confidence intervals for each of the following problems: all pairwise comparisons of means, all contrasts, multiple comparisons with a control, and multiple comparisons with the best. Our confidence intervals are based on standardized time series methods, and we establish the asymptotic validity of each under the sole assumption that the stochastic processes representing the simulation output of the different systems satisfy a functional central limit theorem. Although simulation is the context in this paper, the results naturally apply to (asymptotically) stationary time series.[1]

1 Introduction

Suppose that there are $k \geq 2$ different systems (i.e., stochastic processes) that we want to compare, where system i has (unknown) steady-state mean μ_i and (unknown) asymptotic variance σ_i^2. We allow for the asymptotic variances to be unequal and for the distributions of the k systems to be different. We consider the problem of running independent simulations to compare the steady-state

[1]This work is supported, in part, by the National Science Foundation under Grant No. DMI-9624469 and by the New Jersey Institute of Technology under Grant No. 421180.
AMS 1991 *subject classifications.* Primary 68U20, 62J15; secondary 65C05, 62M10, 60F17.
Key words and phrases. Stochastic simulation; time series; output analysis; functional central limit theorem; standardized time series; multiple comparisons.

means of the different systems. For example, the different systems may represent various service disciplines in a queueing system, and we are interested in comparing the steady-state throughputs of the systems. Although simulation is the context in this paper, the results naturally apply to (asymptotically) stationary time series.

In this paper, we present some single-stage simulation procedures for constructing simultaneous confidence intervals for each of the following multiple-comparisons problems: (1) all pairwise comparisons $\mu_i - \mu_j$, $i < j$; (2) all contrasts $c_1\mu_1 + c_2\mu_2 + \cdots + c_k\mu_k$, where the constants $(c_1, c_2, \ldots, c_k) \in \Re^k$ satisfy $c_1 + c_2 + \cdots + c_k = 0$; (3) multiple comparisons with a control $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$, where system k is considered to be the control; and (4) multiple comparisons with the best (MCB), $\mu_i - \max_{j \neq i} \mu_j$, $i = 1, 2, \ldots, k$. Our confidence intervals are constucted using a standardized time series method, and they are shorter than those based on the Bonferroni inequality. We prove that our confidence intervals are asymptotically valid (as $n \to \infty$, with the run length of each of the systems equal to n); i.e., for each problem above, the joint probability that all of our confidence intervals simultaneous cover the true values is, in the limit, at least $1 - \alpha$, where α is prespecified by the user.

Most of the previous work on multiple-comparison procedures compared k normally distributed populations using i.i.d. sampling within each population. Tamhane (1977) studied the first two problems listed above, and Spøtvoll (1972), Dalal (1978), and Tamhane (1979) constructed confidence intervals for all linear combinations of means of normals. Hsu (1981,1984a,1984b) and Edwards and Hsu (1983) developed confidence intervals for multiple comparisons with the best. For an overview of these and other multiple-comparison procedures for i.i.d. random variables, see Hochberg and Tamhane (1987) and Miller (1981).

There has been some additional work on multiple-comparison procedures specifically developed for use in simulations. Nelson and Hsu (1993), Nelson (1993), and Yang and Nelson (1991) attack the problem of comparing normally distributed populations by using common random numbers to reduce the variance. Also, Yuan and Nelson (1993) consider MCB procedures for steady-state simulations under the assumption that the simulation output of each system can be modeled as an autoregressive process. Goldsman and Nelson (1990) empirically study a heuristic simulation method for steady-state MCB.

Our results extend the previous work by proving the asymptotic validity of multiple-comparison methods for the types of dependent, non-normally distributed output typically encountered in steady-state simulations. We establish our results under the sole assumption that the stochastic processes representing the simulation output of the different systems satisfy a functional central

limit theorem. This assumption is satisfied by virtually all stochastic processes arising in practice.

As previously mentioned, our confidence intervals are based on standardized time series methods. Schruben (1983) proposed this class of techniques for constructing confidence intervals for the steady-state mean of a stochastic process representing the simulation output of a single system. Glynn and Iglehart (1990) formalized and generalized the class of methods and studied some of its theoretical properties. The basic idea behind these approaches is to "cancel out" the asymptotic variance constant σ (in a manner akin to the t-statistic) rather than consistently estimate it. This is desirable because consistent estimation of σ can be difficult in practice. Specifically, certain methods for doing this (viz., the regenerative, autoregressive, and spectral methods) are computationally complicated and not robust. For further work on standardized time series, see Goldsman and Schruben (1984), Chen and Sargent (1987), Sargent, Kang, and Goldsman (1992), and Nakayama (1994). Finally, we mention that the method of batch means (also known as using subseries), a technique that has been studied extensively in the simulation literature (e.g., see Bratley, Fox, and Schrage 1987 and Schmeiser 1982) and in the statistics literature (e.g., see Carlstein 1986), is an example of a standardized time series methodology.

The rest of the paper is organized as follows. In Section 2, we develop the notation and state our functional-central-limit-theorem assumption. We also present the class of standardized time series methods in Section 2. Section 3 contains our multiple-comparison procedures. We give examples of standardized time series techniques in Section 4, and all of the proofs are collected in Section 5. Finally, we note that Damerdji and Nakayama (1996) develop some two-stage multiple-comparison procedures for steady-state simulations. Also, Nakayama (1996a) presents (without proof) MCB confidence intervals for single-stage steady-state simulations using the method of batch means. Nakayama (1996b) studies the case when there is correlation among the different systems induced by common random numbers.

2 Notation and Assumptions

Suppose that there are $k \geq 2$ systems, labeled $1, 2, \ldots, k$. For system $i = 1, 2, \ldots, k$, let $\mathbf{Y}_i = \{Y_i(t) : t \geq 0\} \in D_1[0, \infty)$ be a real-valued (measureable) stochastic process representing the simulation output of system i, where $D_1[0, \infty)$ is the space of right-continous real-valued functions on $[0, \infty)$ having left limits (see Ethier and Kurtz 1986 or Glynn 1990 for more details on the space $D_1[0, \infty)$.) Essentially all stochastic processes arising in practice have sample paths lying in $D_1[0, \infty)$. We can work with discrete-time processes $\{Y_{i,l} : l = 0, 1, 2, \ldots\}$ by taking $Y_i(t) = Y_{i,\lfloor t \rfloor}$, where $\lfloor \beta \rfloor$ denotes the greatest integer less than or equal to $\beta \in \Re$. Let $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k)$ and

$Y(t) = (Y_1(t), Y_2(t), \ldots, Y_k(t))$.

We assume that the processes $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k$ are mutually independent. (In practice, this means that for all i and j with $j \neq i$, the simulations of systems i and j are generated using non-overlapping streams of uniform random numbers.) We allow for the distributions of $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k$ to be different.

To establish our results, we need to restrict our attention to processes $\mathbf{Y}$ that satisfy a functional central limit theorem (FCLT). More formally, letting "$\Rightarrow$" denote weak convergence (see Billingsley 1968 for details), we assume the following:

A1 *There exist a finite diagonal matrix $\Sigma \in \Re^{k \times k}$ with diagonal elements σ_i, $i = 1, 2, \ldots, k$, such that $\sigma_i > 0$, $i = 1, 2, \ldots, k$, and a finite constant $\mu = (\mu_1, \mu_2, \ldots, \mu_k) \in \Re^k$ such that*

$$X_n \Rightarrow \Sigma B$$

as $n \to \infty$, where B is a standard k-dimensional Brownian motion, $X_n = (X_{1,n}, X_{2,n}, \ldots, X_{k,n})$, and

$$X_{i,n}(t) = n^{1/2} \left(\bar{Y}_{i,n}(t) - \mu_i t \right), \quad 0 \leq t \leq 1,$$

with

$$\bar{Y}_{i,n}(t) = \frac{1}{n} \int_0^{nt} Y_i(s) ds, \quad 0 \leq t \leq 1,$$

for $i = 1, 2, \ldots, k$.

Since we assumed that the $\mathbf{Y}_i$, $i = 1, 2, \ldots, k$, are mutually independent, the off-diagonal elements of the matrix Σ are all 0. Note that both X_n and $\bar{Y}_n = (\bar{Y}_{1,n}, \bar{Y}_{2,n}, \ldots, \bar{Y}_{k,n})$ lie in $C[0,1]$, the space of continuous $\Re^k$-valued functions on $[0,1]$; see Ethier and Kurtz (1986) or Glynn (1990) for further details on the space $C[0,1]$. Also, X_n is a rescaled, normalized, integrated version of the original process $\mathbf{Y}$, and the time parameter of X_n and $\bar{Y}_n$ are rescaled by n as compared to $\mathbf{Y}$.

Observe that Assumption A1 implies that for each i,

$$\frac{1}{t} \int_0^t Y_i(s) ds - \mu_i \;=\; \frac{1}{\sqrt{t}} X_{i,t}(1) \;\Rightarrow\; 0 \cdot \sigma_i B_i(1) \;=\; 0$$

as $t \to \infty$, and so the μ_i, $i = 1, 2, \ldots, k$, appearing in A1 are precisely the steady-state means of the process $\mathbf{Y}$. Also, A1 ensures that for each i,

$$\sqrt{n} \left[\frac{1}{n} \int_0^n Y_i(s) ds - \mu_i \right] \;=\; X_{i,n}(1) \;\Rightarrow\; \sigma_i B_i(1)$$

as $n \to \infty$. Recalling that $B_i(1)$ has a standard normal distribution, we see that σ_i is the asymptotic variance parameter of the process $\mathbf{Y}_i$. In addition, $B = (B_1, B_2, \ldots, B_k)$, where each B_i

is a standard 1-dimensional Brownian motion, and $B_1, B_2, \ldots, B_k$ are mutually independent. By Theorem 3.2 and pages 26–27 of Billingsley (1968), Assumption A1 is then equivalent to requiring that

$$X_{i,n} \Rightarrow \sigma_i B_i$$

as $n \to \infty$ for each $i = 1, 2, \ldots, k$.

Virtually all "real-world" stochastic systems having a steady state satisfy the FCLT in Assumption A1. For example, Assumption A1 holds if the process $\mathbf{Y}$ satisfies any of the following:

1. $\mathbf{Y}$ is regenerative and satisfies suitable moment conditions (see Glynn and Whitt 1987);

2. $\mathbf{Y}$ is a martingale process (see Chapter 7 of Ethier and Kurtz 1986);

3. $\mathbf{Y}$ satisfies appropriate mixing conditions (see Chapter 7 of Ethier and Kurtz 1986); or

4. The $\mathbf{Y}(t)$ are associated (see Newman and Wright 1981).

Now we describe the class of standardized time series methods (as applied to the output of a single system). The foundation of these techniques is a class of functions g defined by Glynn and Iglehart (1990). The basic idea is to divide the output of each system into a fixed number $m \geq 1$ of (non-overlapping) batches. The function g is then applied to a scaled, normalized, and integrated version of each process Y_i, namely $X_{i,n}$, and we can think of $g(X_{i,n})$, when appropriately scaled, as an "estimate" of the asymptotic variance constant σ_i. (Glynn and Whitt 1990 show that the method of batch means with a fixed number of batches, which is an example of a standardized time series method, cannot consistently estimate the asymptotic variance.)

More formally, let $C_1[0,1]$ be the space of $\Re$-valued continuous functions on $[0,1]$, and let B_1 be a standard 1-dimensional Brownian motion. Also, for a (measurable) function $h : C_1[0,1] \to S$ with S some metric space, let $D(h)$ be the set of discontinuities of h. Also define the function $e_1 \in C_1[0,1]$ to be $e_1(t) = t$. Then

A2 *The (measurable) function $g : C_1[0,1] \to \Re$ satisfies the following conditions:*

(i) $g(\alpha x) = \alpha g(x)$ *for* $\alpha > 0$, $\alpha \in \Re$, $x \in C_1[0,1]$.

(ii) $g(x - \beta e_1) = g(x)$ *for* $\beta \in \Re$ *and* $x \in C_1[0,1]$.

(iii) $P\{g(B_1) > 0\} = 1$.

(iv) $P\{B_1 \in D(g)\} = 0$.

Glynn and Iglehart (1990) defined $\mathcal{M}$ as the class of functions g satisfying Assumption A2. Condition (i) ensures that $g(X_{i,n})$ is a well-behaved "estimator" of the parameter σ_i in the sense that if we multiply all of the observations of $\mathbf{Y}_i$ by some constant α, then the new asymptotic variance parameter will be $\alpha\sigma_i$. This property will allow us to "cancel out" the asymptotic variance constant σ_i; for more details, see the proof of Theorem 1 in the Section 5. Condition (ii) guarantees that $g(X_{i,n})$ does not depend on the unknown parameter μ_i. Conditions (iii) and (iv) are technical assumptions required to invoke the continuous mapping principle (Proposition 1 in Section 5).

Observe that for each system $i = 1, 2, \ldots, k$, $X_{i,n} = n^{1/2}(\bar{Y}_{i,n} - \mu_i e_1)$. Thus, if $g \in \mathcal{M}$, then

$$g(X_{i,n}) = n^{1/2} g(\bar{Y}_{i,n} - \mu_i e_1) = n^{1/2} g(\bar{Y}_{i,n})$$

by Assumption A2(i) and (ii).

As noted by Glynn and Iglehart (1990), the method of batch means (with a fixed number of batches) is an example of a standardized time series methodology. Therefore, this technique has a corresponding function g. For further details on this and other functions g, see Section 4.

3 Our Multiple-Comparison Procedures

When presenting all of our procedures, we will use the following notation and assumptions. There are k systems, which are simulated independently. Prior to running the simulation, we specify the desired confidence level $1 - \alpha$. We run the simulation of each system i, $i = 1, 2, \ldots, k$, with a run length n, where n is large. For each system i, we analyze its simulation output as follows. We divide the simulation output of system i into $m_i \geq 1$ (non-overlapping) batches, each of length n/m_i, and apply a function g_i to the output, where g_i satisfies Assumption A2. Then, compute the estimate of the steady-state mean of system i as

$$\hat{\mu}_i(n) = \frac{1}{n} \int_0^n Y_i(s)\, ds$$

and

$$S_i^2(n) = g_i^2(X_{i,n}) m_i / n, \tag{1}$$

which, when divided by m_i, is an "estimate" of the variance of $\hat{\mu}_i(n)$. Explicit formulae for calculating $g_i(X_{i,n})$ for various functions g_i are given in Section 4.

3.1 All Pairwise Comparisons

First suppose that we would like to simultaneously make all pairwise comparisons of systems. Before presenting our confidence intervals, we first need the following definition. For a given

probability $1-2\gamma$ and a function g_i satisfying A2 using m_i batches, we define the constant $\nu_{i,\gamma} \geq 0$ such that

$$E\left[F_{\chi^2}\left(\nu_{i,\gamma}^2 g_i^2(B_i)\right)\right] = 1-2\gamma, \tag{2}$$

where F_{χ^2} denotes the distribution function of a χ^2 random variable with 1 degree of freedom. The continuity of F_{χ^2} and the bounded convergence theorem imply that $\nu_{i,\gamma}$ exists. In Section 4 we will describe how to select $\nu_{i,\gamma}$ for various functions g.

Now we simultaneously construct the (two-sided) confidence intervals

$$\begin{aligned} I_{i,j}(n) \;=\; & \left[\hat{\mu}_i(n) - \hat{\mu}_j(n) - \left(\frac{\nu_{i,\beta/2}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta/2}^2 S_j^2(n)}{m_j}\right)^{1/2}, \right. \\ & \left. \hat{\mu}_i(n) - \hat{\mu}_j(n) + \left(\frac{\nu_{i,\beta/2}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta/2}^2 S_j^2(n)}{m_j}\right)^{1/2}\right] \end{aligned} \tag{3}$$

for $\mu_i - \mu_j$, $1 \leq i < j \leq k$, where $\beta = 1-(1-\alpha)^{2/(k(k-1))}$ and the desired confidence level is $1-\alpha$. Then we have the following result, whose proof is given in Section 5.

Theorem 1 *Assume Assumption A1 holds and that for each $i = 1, 2, \ldots, k$, g_i satisfies Assumption A2 with $m_i \geq 1$ batches. Also, for a desired confidence level $1-\alpha$, let $\beta = 1-(1-\alpha)^{2/(k(k-1))}$. Then,*

$$\lim_{n\to\infty} P\left\{\mu_i - \mu_j \in I_{i,j}(n),\ \forall\ i<j\right\} \geq 1-\alpha.$$

Theorem 1 makes use of Šidák's (1967) inequality to bound below the probability of simultaneous coverage of the confidence intervals by a function of the individual coverage probabilities. This bound is sharper than the Bonferroni inequality. Thus, Theorem 1 yields confidence intervals that are shorter than those based on the Bonferroni inequality.

3.2 All Contrasts

The previous theorem considered the problem of making all pairwise comparisons. Now we examine constructing simultaneous confidence intervals for all contrasts $\sum_{i=1}^k c_i\mu_i$ with $c = (c_1, c_2, \ldots, c_k) \in C^k$, where $C^k = \{c \in \Re^k : \sum_{i=1}^k c_i = 0\}$ is the k-dimensional contrast space. For example, this is useful if we want to analyze weighted means such as $\mu_1 - (\mu_2+\mu_3)/2$.

To study the setting of all contrasts, we need the following lemma due to Tukey (1953); also see pages 81–82 of Hochberg and Tamhane (1987).

Lemma 1 *Let $x = (x_1, x_2, \ldots, x_k) \in \Re^k$ and let $\xi_{i,j}$, $1 \le i < j \le k$, be nonnegative real numbers. Then $|x_i - x_j| \le \xi_{i,j}$ for all $i < j$ if and only if*

$$\left|\sum_{i=1}^{k} c_i x_i\right| \le \frac{2}{\sum_{l=1}^{k} |c_l|} \sum_{i=1}^{k}\sum_{j=1}^{k} c_i^+ c_j^- \xi_{i,j} \qquad \forall\ c = (c_1, c_2, \ldots, c_k) \in C^k,$$

where $c_i^+ = \max(c_i, 0)$ and $c_j^- = -\min(c_j, 0)$.

Thus, for each $c = (c_1, c_2, \ldots, c_k) \in C^k$, we now define the confidence interval

$$\begin{aligned} I_c(n) \quad = \quad & \left[\sum_{i=1}^{k} c_i \hat{\mu}_i(n) - \frac{2}{\sum_{l=1}^{k} |c_l|} \sum_{i=1}^{k}\sum_{j=1}^{k} c_i^+ c_j^- \left(\frac{\nu^2_{i,\beta/2} S_i^2(n)}{m_i} + \frac{\nu^2_{j,\beta/2} S_j^2(n)}{m_j}\right)^{1/2}, \right. \\ & \left. \sum_{i=1}^{k} c_i \hat{\mu}_i(n) + \frac{2}{\sum_{l=1}^{k} |c_l|} \sum_{i=1}^{k}\sum_{j=1}^{k} c_i^+ c_j^- \left(\frac{\nu^2_{i,\beta/2} S_i^2(n)}{m_i} + \frac{\nu^2_{j,\beta/2} S_j^2(n)}{m_j}\right)^{1/2}\right] \end{aligned} \tag{4}$$

for $\sum_{i=1}^{k} c_i \mu_i$, where $\nu_{i,\beta/2}$ is as defined in (2). Then Theorem 1 and Lemma 1 immediately imply the following.

Theorem 2 *Assume Assumption A1 holds and that for each $i = 1, 2, \ldots, k$, g_i satisfies Assumption A2 with $m_i \ge 1$ batches. Also, for a desired confidence level $1-\alpha$, let $\beta = 1 - (1-\alpha)^{2/(k(k-1))}$. Then,*

$$\lim_{n\to\infty} P\left\{\sum_{i=1}^{k} c_i \mu_i \in I_c(n),\ \forall\ c = (c_1, c_2, \ldots, c_k) \in C^k\right\} \ge 1 - \alpha.$$

Theorem 2 is mainly intended for constructing only those confidence intervals with $c \in C^k$ that are of interest. Also, we note that contrasts only allow comparisons between means. However, in many instances, we may also like simultaneously to construct confidence intervals for the individual means. For example, we may want joint confidence intervals for $\mu_1 - \mu_2$ and μ_1. This may be accomplished by developing simultaneous confidence intervals for all linear combinations of means as done on pp. 183–186 of Hochberg and Tamhane (1987). However, we do not examine this further.

3.3 Multiple Comparisons with a Control

We now consider the problem of making multiple comparisons with a control. More specifically, suppose that system k is the control, and we want to compare all other systems $i \ne k$ simultaneously to the control. For example, system k may represent a system already in place, and systems $1, 2, \ldots, k-1$ are various alternatives with which we might replace system k if one of the other systems is better (as measured by the steady-state means).

To do this, we define simultaneous $(1-\alpha)$-level two-sided confidence intervals $I_{i,k}(n)$ as in (3) for $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$, with $\beta = 1 - (1-\alpha)^{1/(k-1)}$. Similarly, we define upper one-sided confidence intervals as

$$I_{u,i,k}(n) = \left(-\infty,\ \hat{\mu}_i(n) - \hat{\mu}_k(n) + \left(\frac{\nu_{i,\beta}^2 S_i^2(n)}{m_i} + \frac{\nu_{k,\beta}^2 S_k^2(n)}{m_k} \right)^{1/2} \right], \tag{5}$$

and lower one-sided confidence intervals as

$$I_{l,i,k}(n) = \left[\hat{\mu}_i(n) - \hat{\mu}_k(n) - \left(\frac{\nu_{i,\beta}^2 S_i^2(n)}{m_i} + \frac{\nu_{k,\beta}^2 S_k^2(n)}{m_k} \right)^{1/2},\ +\infty \right) \tag{6}$$

for $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$. (We use subscripts u and l on I to denote that the one-sided confidence interval are upper and lower, respectively.) Then we have the following result, whose proof is given in Section 5.

Theorem 3 *Assume Assumption A1 holds and that for each $i = 1, 2, \ldots, k$, g_i satisfies Assumption A2 with $m_i \geq 1$ batches. Also, for a desired confidence level $1-\alpha$, let $\beta = 1 - (1-\alpha)^{1/(k-1)}$. Then,*

(i) *for the simultaneous two-sided confidence intervals,*

$$\lim_{n\to\infty} P\left\{\mu_i - \mu_k \in I_{i,k}(n),\ i = 1, 2, \ldots, k-1\right\} \geq 1 - \alpha;$$

(ii) *for the simultaneous upper one-sided confidence intervals,*

$$\lim_{n\to\infty} P\left\{\mu_i - \mu_k \in I_{u,i,k}(n),\ i = 1, 2, \ldots, k-1\right\} \geq 1 - \alpha;$$

(iii) *for the simultaneous lower one-sided confidence intervals,*

$$\lim_{n\to\infty} P\left\{\mu_i - \mu_k \in I_{l,i,k}(n),\ i = 1, 2, \ldots, k-1\right\} \geq 1 - \alpha.$$

3.4 Multiple Comparisons with the Best

Now we construct simultaneous confidence intervals for $\mu_i - \max_{j\neq i} \mu_j$, $i = 1, 2, \ldots, k$. This is useful when we want to determine the system with the largest mean. Thus, define the confidence interval

$$\begin{aligned} I_{b,i}(n) = \Bigg[& -\left(\min_{j\in\mathcal{A}(n),\ j\neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) - \left(\frac{\nu_{i,\beta}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta}^2 S_j^2(n)}{m_j} \right)^{1/2} \right) \right)^{-}, \\ & \left(\min_{j\neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + \left(\frac{\nu_{i,\beta}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta}^2 S_j^2(n)}{m_j} \right)^{1/2} \right) \right)^{+} \Bigg] \end{aligned} \tag{7}$$

for $\mu_i - \max_{j\neq i}\mu_j$, $i = 1, 2, \ldots, k$, where $\nu_{i,\beta}$ is as defined in (2), $\beta = 1-(1-\alpha)^{1/(k-1)}$,

$$\mathcal{A}(n) = \left\{ i : \min_{j\neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + \left(\frac{\nu_{i,\beta}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta}^2 S_j^2(n)}{m_j} \right)^{1/2} \right) \geq 0 \right\},$$

and we recall that $(\gamma)^+ = \max(\gamma, 0)$ and $(\gamma)^- = -\min(\gamma, 0)$. In (7), we define $\min_{j\in\emptyset} x_j = 0$. (We use a subscript b on I to denote that the confidence interval is for multiple comparisons with the best.) Then we have the following result, whose proof is given in Section 5.

Theorem 4 *Assume Assumption A1 holds and that for each $i = 1, 2, \ldots, k$, g_i satisfies Assumption A2 with $m_i \geq 1$ batches. Also, for a desired confidence level $1-\alpha$, let $\beta = 1-(1-\alpha)^{1/(k-1)}$. Then,*

$$\lim_{n\to\infty} P\left\{ \mu_i - \max_{j\neq i}\mu_j \in I_{b,i}(n),\ i = 1, 2, \ldots, k \right\} \geq 1-\alpha.$$

Note that the MCB confidence intervals use a quantile point with $\beta = 1-(1-\alpha)^{1/(k-1)}$ rather than $\beta = 1-(1-\alpha)^{2/(k(k-1))}$ as in all pairwise comparisons. Thus, MCB intervals will typically be shorter than those arising from all pairwise comparisons.

4 Examples of Standardized Time Series

In this section we present various functions g satisfying Assumption A2. All of these examples are taken directly from Glynn and Iglehart (1990). Also, we describe how to determine the constant $\nu_{i,\gamma}$ given in (2).

Example 1. The first function g that we describe corresponds to the method of batch means. For this example we require that the number of batches m is at least 2. Define the function $g : C_1[0,1] \to \Re$ as

$$g(x) = \left[\frac{m}{m-1} \sum_{i=1}^{m} \left(\Delta_m x\left(\frac{i}{m}\right) - \frac{x(1)}{m} \right)^2 \right]^{1/2},$$

where $\Delta_d x(t) = x(t) - x(t-1/d)$ for $d \in \Re$ with $d > 0$. Thus, if we divide the output of system i into $m_i \geq 2$ (non-overlapping) batches and apply function g_i as above to $X_{i,n}$, we obtain

$$g_i(X_{i,n}) = \left(\frac{n}{m_i}\right)^{1/2} \left[\frac{1}{m_i - 1} \sum_{j=1}^{m_i} \left(Z_{i,j}(n) - \frac{1}{m_i} \sum_{l=1}^{m_i} Z_{i,l}(n) \right)^2 \right]^{1/2},$$

where n is the run length of the simulation of system i, which is proportional to n, and

$$Z_{i,j}(n) = \frac{1}{n/m_i} \int_{(j-1)n/m_i}^{jn/m_i} Y_i(s)\ ds, \quad j \geq 1,$$

which is the sample mean of the jth (non-overlapping) batch of size n/m_i of system i. Note that $S_i^2(n) = g_i^2(X_{i,n})m_i/n$ is equal to the sample variance of the m_i batch means of system i. Also, $(m_i - 1)g_i^2(B_i)$ has a χ^2 distribution with $m_i - 1$ degrees of freedom.

When g_i is defined this way, the constant $\nu_{i,\gamma}$ in (2) is given by $\nu_{i,\gamma} = t_{m_i-1,\gamma}$, where $t_{m,\gamma}$ is the upper γ point of a Student's t-distribution with m degrees of freedom; i.e., $P\{t_m \geq t_{m,\gamma}\} = \gamma$, where t_m is a random variable having a t-distribution with m degrees of freedom. ∎

Example 2. Our next function g gives rise to the standardized sum method developed by Schruben (1983). Let $m \geq 1$. Then define the function $g : C_1[0,1] \to \Re$ as

$$g(x) = \left[\sum_{j=0}^{m-1}\left(\int_0^1 x\left(\frac{j+t}{m}\right)dt - \frac{1}{2}\left(x\left(\frac{j+1}{m}\right) + x\left(\frac{j}{m}\right)\right)\right)^2\right]^{1/2}.$$

Thus, if we divide the output of system i into $m_i \geq 1$ (non-overlapping) batches and apply function g_i as above to $X_{i,n}$, we obtain

$$g_i(X_{i,n}) = \left[\frac{1}{n}\sum_{j=0}^{m_i-1}\left(\int_0^1 \int_{jn/m_i}^{(j+t)n/m_i} Y_i(s)\,ds\,dt - \frac{1}{2}\int_{jn/m_i}^{(j+1)n/m_i} Y_i(s)\,ds\right)^2\right]^{1/2}.$$

Glynn and Iglehart (1990) show that $12m_i g_i^2(B_i)$ has a χ^2-distribution with m_i degrees of freedom. When g_i is defined this way, the point $\nu_{i,\gamma}$ in (2) is given by $\nu_{i,\gamma} = (12)^{1/2}t_{m_i,\gamma}$. ∎

Example 3. The next function g corresponds to the standardized maximum intervals method described in Schruben (1983). Let $m \geq 1$. Also, for $x \in C_1[0,1]$, we define t^* $(= t^*(x)) = \inf\{t \geq 0 : x(t) = M^*\}$ and M^* $(= M^*(x)) = \max\{x(t) : 0 \leq t \leq 1\}$. Also, define the functions $\Gamma : C_1[0,1] \to C_1[0,1]$ and $\Lambda_j : C_1[0,1] \to C_1[0,1]$ for $j = 0, 1, \ldots, m-1$ as

$$\begin{aligned}\Gamma(x) &= x - e_1 x(1),\\ \Lambda_j(x) &= x\left(\frac{j+e_1}{m}\right) - x\left(\frac{j}{m}\right),\end{aligned}$$

where $e_1(t) = t$. Then let

$$g(x) = \left[\sum_{j=0}^{m-1}\frac{1}{t_j^*(1-t_j^*)}\left(x\left(\frac{j+t_j^*}{m}\right) - x\left(\frac{j}{m}\right) - t_j^*\left(x\left(\frac{j+1}{m}\right) - x\left(\frac{j}{m}\right)\right)\right)^2\right]^{1/2},$$

where $t_j^* = \inf\{t \geq 0 : (\Gamma \circ \Lambda_j)(x)(t) = M_j^*\}$ and $M_j^* = \max\{(\Gamma \circ \Lambda_j)(x)(t) : 0 \leq t \leq 1\}$. Thus, if we divide the output of system i into $m_i \geq 1$ (non-overlapping) batches and apply function g_i as

above to $X_{i,n}$, we obtain

$$g_i(X_{i,n}) = \left[\frac{1}{n} \sum_{j=0}^{m_i-1} \frac{1}{t_j^*(1-t_j^*)} \left(\int_{jn/m_i}^{(j+t_j^*)n/m_i} Y_i(s)\, ds - t_j^* \int_{jn/m_i}^{(j+1)n/m_i} Y_i(s)\, ds \right)^2 \right]^{1/2},$$

where $t_j^* = \inf\{t \geq 0 : (\Gamma \circ \Lambda_j)(\bar{Y}_{i,n})(t) = M_j^*\}$ and $M_j^* = \max\{(\Gamma \circ \Lambda_j)(\bar{Y}_{i,n})(t) : 0 \leq t \leq 1\}$. Schruben (1982) showed that $m_i g_i^2(B_i)$ has a χ^2-distribution with $3m_i$ degrees of freedom. When g_i is defined this way, the point $\nu_{i,\gamma}$ in (2) is given by $\nu_{i,\gamma} = (1/3)^{1/2} t_{3m_i,\gamma}$. ∎

Other standardized time series methods include the Cramér-von Mises method (Goldsman, Kang, and Seila, 1993) and the L_p-norm methods (Tokol, Goldsman, Ockerman, and Swain, 1996).

5 Proofs

Here we will provide the proofs for Theorems 1, 3, and 4 from Section 3. To establish our results, we will repeatedly apply the following proposition, which is known as the continuous mapping principle. (See Theorem 5.1 of Billingsley 1968 or Glynn 1990 for the proof.)

Proposition 1 *Suppose $X_n, X \in C[0,1]$ are random elements such that $X_n \Rightarrow X$ as $n \to \infty$. Consider a (measurable) function $h : C[0,1] \to S$, S a metric space, and let $D(h)$ be the set of discontinuities of h. If $P\{X \in D(h)\} = 0$, then $h(X_n) \Rightarrow h(X)$ as $n \to \infty$.*

We now present some preparatory lemmas that will be useful for proving Theorem 1. The first is due to Šidák (1967).

Lemma 2 *Let $W = (W_1, W_2, \ldots, W_d)$, $d > 1$, be a multivariate normal with mean vector 0 and arbitrary covariance matrix $\Theta = (\Theta_{i,j} : i, j = 1, 2, \ldots, d)$. Then,*

$$P\{|W_i| \leq \beta_i,\ i = 1, 2, \ldots, d\} \geq \prod_{i=1}^{d} P\{|W_i| \leq \beta_i\}$$

for $(\beta_1, \beta_2, \ldots, \beta_{k-1}) \in \Re^d$.

The next lemma was established by Tamhane (1977).

Lemma 3 *Let $W_1, W_2, \ldots, W_d$ be mutually independent real-valued random variables, and let $\Psi = (\Psi_1, \Psi_2, \ldots, \Psi_p) : \Re^d \to \Re^p$, be a nonnegative function, where for each $i = 1, 2, \ldots, p$, Ψ_i is nondecreasing in each of its arguments. Then,*

$$E\left[\prod_{j=1}^{p} \Psi_j(W_1, W_2, \ldots, W_d) \right] \geq \prod_{j=1}^{p} E\left[\Psi_j(W_1, W_2, \ldots, W_d) \right].$$

The following is a generalization of a result of Banerjee (1961).

Lemma 4 *Let F_{χ^2} denote the distribution function of a χ^2 random variable with 1 degree of freedom. Also, let $\lambda_1, \lambda_2, \ldots, \lambda_d$ be real-valued constants such that $\lambda_i \geq 0$, $i = 1, 2, \ldots, d$, and $\lambda_1 + \lambda_2 + \cdots + \lambda_d = 1$. Then*

$$E\left[F_{\chi^2}\left(\sum_{i=1}^{d} \lambda_i \nu_{i,\gamma}^2 g_i^2(B_i)\right)\right] \geq 1 - 2\gamma.$$

Proof. Note that

$$F_{\chi^2}(x) = 2\int_0^{\sqrt{x}} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

for $x \geq 0$ and 0 otherwise. By taking the second derivative of $F_{\chi^2}(x)$ with respect to x, we can show that it is a concave function. Thus,

$$E\left[F_{\chi^2}\left(\sum_{i=1}^{d} \lambda_i \nu_{i,\gamma}^2 g_i^2(B_i)\right)\right] \;\geq\; \sum_{i=1}^{d} \lambda_i \; E\left[F_{\chi^2}\left(\nu_{i,\gamma}^2 g_i^2(B_i)\right)\right] \;=\; \sum_{i=1}^{d} \lambda_i (1-2\gamma) \;=\; 1 - 2\gamma$$

by the definition of $\nu_{i,\gamma}$ in (2). ∎

We are now in a position to prove Theorem 1 of Section 3.

Proof of Theorem 1. Note that

$$\begin{aligned}
&P\{\mu_i - \mu_j \in I_{i,j}(n),\ \forall\ i < j\} \\
&\quad = \quad P\left\{|(\hat{\mu}_i(n) - \mu_i) - (\hat{\mu}_j(n) - \mu_j)| \leq \left(\frac{\nu_{i,\beta/2}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta/2}^2 S_j^2(n)}{m_j}\right)^{1/2},\ \forall\ i < j\right\} \\
&\quad = \quad P\left\{|(\hat{\mu}_i(n) - \mu_i) - (\hat{\mu}_j(n) - \mu_j)| \leq \left(\frac{\nu_{i,\beta/2}^2 m_i g_i^2(X_{i,n})}{m_i n} + \frac{\nu_{j,\beta/2}^2 m_j g_j^2(X_{j,n})}{m_j n}\right)^{1/2},\ \forall\ i < j\right\}
\end{aligned}$$

since $S_i^2(n) = g_i^2(X_{i,n}) m_i / n$ for all i by (1). Thus,

$$\begin{aligned}
&P\{\mu_i - \mu_j \in I_{l,i,j}(n),\ \forall\ i < j\} \\
&\quad = \quad P\left\{\left|n^{1/2}\left(\hat{\mu}_i(n) - \mu_i\right) - n^{1/2}\left(\hat{\mu}_j(n) - \mu_j\right)\right| - \left(\nu_{i,\beta/2}^2 g_i^2(X_{i,n}) + \nu_{j,\beta/2}^2 g_j^2(X_{j,n})\right)^{1/2} \leq 0,\ \forall\ i < j\right\} \\
&\quad = \quad P\left\{|X_{i,n}(1) - X_{j,n}(1)| - \left(\nu_{i,\beta/2}^2 g_i^2(X_{i,n}) + \nu_{j,\beta/2}^2 g_j^2(X_{j,n})\right)^{1/2} \leq 0,\ \forall\ i < j\right\} \\
&\quad = \quad P\left\{u(X_n) \leq \mathbf{0}\right\},
\end{aligned}$$

where the function $u = (u_{i,j} : 1 \leq i < j \leq k) : C[0,1] \to \Re^{k(k-1)/2}$ is defined as

$$u_{i,j}(x) = |x_i(1) - x_j(1)| - \left(\nu_{i,\beta/2}^2 g_i^2(x_i) + \nu_{j,\beta/2}^2 g_j^2(x_j)\right)^{1/2}$$

for $i < j$, and $\mathbf{0} \in \Re^{k(k-1)/2}$ is the $(k(k-1)/2)$-dimensional vector with all components 0. By Assumption A2(iv) and by the fact that the projection mapping is continuous, u is continuous at B with probability 1, and so the continuous mapping principle ensures that $u(X_n) \Rightarrow u(\Sigma B)$ as $n \to \infty$.

Now we show that $u(\Sigma B)$ has a continuous distribution function. Let $H_{i,j}$ be the distribution function of $u_{i,j}(\Sigma B)$. Note that

$$\begin{aligned} u_{i,j}(\Sigma B) &= |\sigma_i B_i(1) - \sigma_j B_j(1)| - \left(\nu_{i,\beta/2}^2 g_i^2(\sigma_i B_i) + \nu_{j,\beta/2}^2 g_j^2(\sigma_j B_j)\right)^{1/2} \\ &= |\sigma_i B_i(1) - \sigma_j B_j(1)| - \left(\nu_{i,\beta/2}^2 \sigma_i^2 g_i^2(B_i) + \nu_{j,\beta/2}^2 \sigma_j^2 g_j^2(B_j)\right)^{1/2} \end{aligned}$$

by Assumption A2(i). Recall that standard 1-dimensional Brownian motion has normal increments. In addition, $B_1, B_2, \ldots, B_k$ are mutually independent, which implies that $g_1(B_1), g_2(B_2), \ldots, g_k(B_k)$ are also mutually independent. Moreover, $(\sigma_i B_i(1) - \sigma_j B_j(1))/(\sigma_i^2 + \sigma_j^2)^{1/2}$ has a standard normal distribution since $0 < \sigma_l < \infty$ for all l. Now let $\bar{\Phi}$ denote the distribution function of the absolute value of a standard normal random variable, and let G_i be the distribution function of $g_i(B_i)$. Then, for $\eta \in \Re$,

$$\begin{aligned} H_{i,j}(\eta) &= P\left\{ \left| \frac{\sigma_i B_i(1) - \sigma_j B_j(1)}{(\sigma_i^2 + \sigma_j^2)^{1/2}} \right| \leq \frac{\eta}{(\sigma_i^2 + \sigma_j^2)^{1/2}} + \left(\frac{\nu_{i,\beta/2}^2 \sigma_i^2 g_i^2(B_i) + \nu_{j,\beta/2}^2 \sigma_j^2 g_j^2(B_j)}{\sigma_i^2 + \sigma_j^2} \right)^{1/2} \right\} \\ &= \int_0^\infty \int_0^\infty \bar{\Phi}\left(\frac{\eta}{(\sigma_i^2 + \sigma_j^2)^{1/2}} + \left(\frac{\nu_{i,\beta/2}^2 \sigma_i^2 x^2 + \nu_{j,\beta/2}^2 \sigma_j^2 y^2}{\sigma_i^2 + \sigma_j^2} \right)^{1/2} \right) G_i(dx)\, G_j(dy) \end{aligned}$$

since each $B_i(1)$ is independent of $g_i(B_i)$, as was shown by Glynn and Iglehart (1990). Now the continuity of $\bar{\Phi}$ and the bounded convergence theorem imply that $H_{i,j}$ is continuous for all $i < j$.

Since $u(\Sigma B)$ has a continuous distribution function, Theorem 2.1 of Billingsley (1968) implies that

$$P\{u(X_n) \leq \mathbf{0}\} \to P\{u(\Sigma B) \leq \mathbf{0}\}$$

as $n \to \infty$. Now we will show that

$$P\{u(\Sigma B) \leq \mathbf{0}\} \geq 1 - \alpha.$$

Let $\mathcal{F}_g$ denote the σ-field generated by $(g_1(B_1), g_2(B_2), \ldots, g_k(B_k))$. Note that

$$\begin{aligned} & P\{u(\Sigma B) \leq \mathbf{0}\} \\ &= E\left[P\left\{ \left| \frac{\sigma_i B_i(1) - \sigma_j B_j(1)}{(\sigma_i^2 + \sigma_j^2)^{1/2}} \right| \leq \left(\frac{\nu_{i,\beta/2}^2 \sigma_i^2 g_i^2(B_i) + \nu_{j,\beta/2}^2 \sigma_j^2 g_j^2(B_j)}{\sigma_i^2 + \sigma_j^2} \right)^{1/2}, \ \forall\ i < j \ \middle|\ \mathcal{F}_g \right\} \right] \\ &\geq E\left[\prod_{i<j} \bar{\Phi}\left(\left(\frac{\nu_{i,\beta/2}^2 \sigma_i^2 g_i^2(B_i) + \nu_{j,\beta/2}^2 \sigma_j^2 g_j^2(B_j)}{\sigma_i^2 + \sigma_j^2} \right)^{1/2} \right) \right] \end{aligned}$$

by Lemma 2. Now define

$$\Psi_{i,j}\left(g_1^2(B_1), g_2^2(B_2), \ldots, g_k^2(B_k)\right) = \bar{\Phi}\left(\left(\frac{\nu_{i,\beta/2}^2\sigma_i^2 g_i^2(B_i) + \nu_{j,\beta/2}^2\sigma_j^2 g_j^2(B_j)}{\sigma_i^2+\sigma_j^2}\right)^{1/2}\right)$$

for all $i<j$, and observe that each $\Psi_{i,j}$ is a nondecreasing function of $g_1^2(B_1), g_2^2(B_2), \ldots, g_k^2(B_k)$. Hence, applying Lemma 3, we get

$$\begin{aligned} P\{u(\Sigma B)\le \mathbf{0}\} &\ge \prod_{i<j} E\left[\bar{\Phi}\left(\left(\frac{\nu_{i,\beta/2}^2\sigma_i^2 g_i^2(B_i) + \nu_{j,\beta/2}^2\sigma_j^2 g_j^2(B_j)}{\sigma_i^2+\sigma_j^2}\right)^{1/2}\right)\right] \\ &= \prod_{i<j} E\left[F_{\chi^2}\left(\frac{\sigma_i^2}{\sigma_i^2+\sigma_j^2}\nu_{i,\beta/2}^2 g_i^2(B_i) + \frac{\sigma_j^2}{\sigma_i^2+\sigma_j^2}\nu_{j,\beta/2}^2 g_j^2(B_j)\right)\right], \end{aligned}$$

where F_{χ^2} denotes the distribution function of a χ^2 random variable with 1 degree of freedom. Thus, Lemma 4 implies that

$$\begin{aligned} P\{u(\Sigma B)\le \mathbf{0}\} &\ge \prod_{i<j} E\left[F_{\chi^2}\left(\frac{\sigma_i^2}{\sigma_i^2+\sigma_j^2}\nu_{i,\beta/2}^2 g_i^2(B_i) + \frac{\sigma_j^2}{\sigma_i^2+\sigma_j^2}\nu_{j,\beta/2}^2 g_j^2(B_j)\right)\right] \\ &\ge \prod_{i<j}(1-\beta) \;=\; (1-\beta)^{k(k-1)/2} \;=\; 1-\alpha \end{aligned}$$

since $\beta = 1-(1-\alpha)^{2/(k(k-1))}$, which completes the proof. ∎

To prove Theorem 3, we need the following result due to Slepian (1962).

Lemma 5 *Let $W=(W_1,W_2,\ldots,W_d)$, $d>1$, be a multivariate normal with mean vector 0 and covariance matrix $\Theta=(\Theta_{i,j}: i,j=1,2,\ldots,d)$. Also, let $V=(V_1,V_2,\ldots,V_d)$ be a multivariate normal with mean vector 0 and covariance matrix $\Xi=(\Xi_{i,j}: i,j=1,2,\ldots,d)$. If $\Theta_{i,j}\ge\Xi_{i,j}$ for all $1\le i<j\le k$, then*

$$P\{W_i\le\beta_i,\ i=1,2,\ldots,d\} \ge P\{V_i\le\beta_i,\ i=1,2,\ldots,d\}$$

for $(\beta_1,\beta_2,\ldots,\beta_{k-1})\in\Re^d$.

Proof of Theorem 3. The proof of part (i) is the same as that of Theorem 1. Also, parts (ii) and (iii) can be established using arguments similar to those employed in the proof of Theorem 1 except we need to rely on Lemma 5 instead of Lemma 2. ∎

Now we prove Theorem 4.

Proof of Theorem 4. First, define

$$D_{i,j}(n) = \left(\frac{\nu_{i,\beta}^2 S_i^2(n)}{m_i} + \frac{\nu_{j,\beta}^2 S_j^2(n)}{m_j} \right)^{1/2}.$$

Also, define $(1), (2), \ldots, (k)$ such that $\mu_{(1)} \leq \mu_{(2)} \leq \cdots \leq \mu_{(k)}$; i.e., system (j) has the j-th smallest steady-state mean. Then, define the events

$$\begin{aligned}
E(n) &= \left\{ \mu_i - \mu_{(k)} \geq \hat{\mu}_i(n) - \hat{\mu}_{(k)}(n) - D_{i,(k)}(n),\ \forall\ i \neq (k) \right\}, \\
E_1(n) &= \left\{ \mu_i - \max_{j \neq i} \mu_j \leq \left[\min_{j \neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + D_{j,i}(n) \right) \right]^+,\ \forall\ i \right\}, \\
E_2(n) &= \left\{ \mu_i - \max_{j \neq i} \mu_j \geq - \left[\min_{j \in \mathcal{A}(n),\ j \neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + D_{i,j}(n) \right) \right]^-,\ \forall\ i \right\}.
\end{aligned}$$

Note that $E(n)$ is the event that the lower one-sided confidence intervals for multiple comparisons with a control, with the control being system (k), contain all of the true differences $\mu_i - \mu_{(k)}$. Thus, we have that $\lim_{n \to \infty} P(E(n)) \geq 1 - \alpha$ by Theorem 3(iii). Now following an argument developed by Edwards and Hsu (1983), we show that $E(n) \subset E_1(n) \bigcap E_2(n)$ for all n, which will establish the result.

First we prove that $E(n) \subset E_1(n)$:

$$\begin{aligned}
E(n) &\subset \left\{ \mu_{(k)} - \mu_{(k-1)} \leq \hat{\mu}_{(k)}(n) - \hat{\mu}_j(n) + D_{j,(k)}(n),\ \forall\ j \neq (k) \right\} \\
&\subset \left\{ \mu_i - \mu_{(k-1)} \leq \left[\min_{j \neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + D_{j,i}(n) \right) \right]^+,\ \forall\ i \right\} \\
&\subset \left\{ \mu_i - \max_{j \neq i} \mu_j \leq \left[\min_{j \neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + D_{j,i}(n) \right) \right]^+,\ \forall\ i \right\},
\end{aligned}$$

where the penultimate step follows since $\mu_i - \mu_{(k-1)} \leq 0$ for all $i \neq (k)$ and $[\ \cdot\ ]^+ \geq 0$.

Now we show $E(n) \subset E_2(n)$. First note that on the event $E(n)$, we have that $(k) \in \mathcal{A}(n)$ since $E(n) = \{\hat{\mu}_{(k)}(n) - \hat{\mu}_j(n) + D_{j,(k)}(n) \geq \mu_{(k)} - \mu_j,\ \forall\ j \neq (k)\}$ and $\mu_{(k)} - \mu_j \geq 0$ for all j. Hence,

$$\begin{aligned}
E(n) &\subset \left\{ \mu_i - \mu_{(k)} \geq \min_{j \in \mathcal{A}(n),\ j \neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) - D_{i,j}(n) \right),\ \forall\ i \neq (k) \right\} \\
&\subset \left\{ \mu_i - \max_{j \neq i} \mu_j \geq - \left[\min_{j \in \mathcal{A}(n),\ j \neq i} \left(\hat{\mu}_i(n) - \hat{\mu}_j(n) - D_{i,j}(n) \right) \right]^-,\ \forall\ i \right\},
\end{aligned}$$

where the last step follows since $\mu_{(k)} - \max_{j \neq (k)} \mu_j \geq 0$ and $-[\ \cdot\ ]^- \leq 0$. Hence, $E(n) \subset E_1(n) \bigcap E_2(n)$, and the proof is complete. ∎

Acknowledgments

The author would like to thank the referees for their comments, which improved the quality of the paper.

References

Banerjee, S. K. 1961. On confidence interval for two-means problem based on separate estimates of variances and tabulated values of t-table. *Sankyā A* 23:359–378.

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

Bratley, P., B. L. Fox, and L. E. Schrage. 1987. *A Guide to Simulation, Second Edition*. New York: Springer-Verlag.

Carlstein, E. 1986. The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Annals of Statistics* 14:1171–1179.

Chen, B.-C. and R. G. Sargent. 1987. Using standardized time series to estimate confidence intervals for the difference between two stationary stochastic processes. *Operations Research* 35:428–436.

Dalal, S. R. 1978. Simultaneous confidence interval procedures for univariate and multivariate Behrens-Fisher type problems. *Biometrika* 65:221–224.

Damerdji, H. and Nakayama, M. K. 1996. Two-stage multiple-comparison procedures for steady-state simulations. In preparation.

Edwards, D. G. and J. C. Hsu. 1983. Multiple comparisons with the best treatment *Journal of the American Statistical Association* 78:965–971.

Ethier, S. N. and T. G. Kurtz. 1986. *Markov Processes: Characterization and Convergence*. New York: John Wiley.

Glynn, P. W. 1990. Diffusion approximations. Chapter 4 of *Handbooks in Operations Research and Management Science, Vol. 2, Stochastic Models*, ed. D. Heyman and M. Sobel. Elsevier Science Publishers B. V. (North-Holland).

Glynn, P. W. and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 14:1–16.

Glynn, P. W. and W. Whitt. 1987. Sufficient conditions for functional-limit-theorem versions of $L = \lambda W$. *Queueing Systems: Theory and Applications* 1:279–287.

Glynn, P. W. and W. Whitt. 1990. Estimating the asymptotic variance with batch means. *Operations Research Letters* 10:431–435.

Goldsman, D., K. Kang, and A. F. Seila. 1993. Cramér-von Mises variance estimators for simulations. Technical report, School of Industrial Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia.

Goldsman, D. and L. W. Schruben. 1984. Asymptotic properties of some confidence interval estimators for simulation output. *Management Science* 30:1217–1225.

Goldsman, L. and B. L. Nelson. 1990. Batch-size effects on simulation optimization using multiple comparisons with the best. In *Proceedings of the 1990 Winter Simulation Conference*, ed. O. Balci, R. P. Sadowski, and R. E. Nance. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers. 288–293.

Hochberg, Y. and A. C. Tamhane. 1987. *Multiple Comparison Methods*. New York: John Wiley.

Hsu, J. C. 1981. Simultaneous confidence intervals for all distances from the "best". *Annals of Statistics* 9:1026–1034.

Hsu, J. C. 1984a. Ranking and selection and multiple comparisons with the best. In *Design of Experiments: Ranking and Selection (Essays in Honor of Robert E. Bechhofer)*, ed. T. J. Santner and A. C. Tamhane, 23–33. New York: Marcel Dekker.

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

Miller, R. G. 1981. *Simultaneous Statistical Inference, 2nd Edition*. New York: Springer-Verlag.

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

Nakayama, M. K. 1996a. Multiple comparisons with the best in steady-state simulations. *Proceedings of the Second International Workshop on Mathematical Methods in Stochastic Simulation and Experimental Design*. (S. M. Ermakov and V. B. Melas, eds.), St. Petersburg, Russia, 230–235.

Nakayama, M. K. 1996b. Multiple comparisons with the best using common random numbers in steady-state simulations. Research Report CIS-96-22, Computer and Information Science Department, New Jersey Institute of Technology.

Nelson, B. L. 1993. Robust multiple comparisons under common random numbers. *ACM Transactions on Modeling and Computer Simulation* 3:225–243.

Nelson, B. L. and J. C. Hsu. 1993. Control-variate models of common random numbers for multiple comparison with the best. *Management Science* 39:989–1001.

Newman, C. M. and A. L. Wright. 1981. An invariance priciple for certain dependent sequences. *Annals of Probability* 9:671–675.

Sargent, R. G., K. Kang, and D. Goldsman. 1992. An investigation of finite sample behavior of confidence interval estimation procedures. *Operations Research* 40:898–913.

Schmeiser, B. 1982. Batch size effects in the analysis of simulation output. *Operations Research* 30:556–568.

Schruben, L. W. 1982. Detecting initialization bias in simulation output. *Operations Research* 30:569–590.

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

Šidák, Z. 1967. Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association* 62:626–633.

Slepian, D. 1962. The one-sided barrier problem for Gaussian noise. *Bell System Technical Journal* 41:463–501.

Spøtvoll, E. 1972. Joint confidence intervals for all linear combinations in ANOVA with unknown variances. *Biometrika* 59:684–685.

Tamhane, A. C. 1977. Multiple comparisons in model I one-way ANOVA with unequal variances. *Communications in Statistics* A6:15–32.

Tamhane, A. C. 1979. A comparison of procedures for multiple comparisons of means with unequal variances. *Journal of the American Statistical Association* 74:471–480.

Tokol, G., D. Goldsman, D. J. Ockerman, and J. J. Swain. 1996. Standardized time series L_p-norm variance estimators for simulations. *Management Science*, to appear.

Tukey, J. W. 1953. *The Problem of Multiple Comparisons*, mimeographed monograph.

Yang, W. and B. L. Nelson. 1991. Using common random numbers and control variates in multiple-comparison procedures. *Operations Research* 39:583–591.

Yuan, M. and B. L. Nelson. 1993. Multiple comparisons with the best for steady-state simulation. *ACM Transactions on Modeling and Computer Simulation* 3:66–79.