

## MULTIPLE-COMPARISON PROCEDURES FOR STEADY-STATE SIMULATIONS<sup>1</sup>

BY MARVIN K. NAKAYAMA

*New Jersey Institute of Technology*

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 of this paper, the results naturally apply to (asymptotically) stationary time series.

**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  $\mu_i$  and (unknown) asymptotic variance  $\sigma_i^2$ . 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 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 of 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-comparison 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  $(c_1, c_2, \dots, c_k) \in \Re^k$  satisfy  $c_1 + c_2 + \cdots + c_k = 0$ ; (3) multiple comparisons with a control  $\mu_i - \mu_k$ ,  $i = 1, 2, \dots, k - 1$ , where system  $k$  is considered to be the control;

---

Received February 1995; revised June 1997.

<sup>1</sup>Work is supported, in part, by the National Science Foundation under Grant DMI-96-24469 and by the New Jersey Institute of Technology under Grant 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.

and (4) multiple comparisons with the best (MCB),  $\mu_i - \max_{j \neq i} \mu_j$ ,  $i = 1, 2, \dots, k$ . Our confidence intervals are constructed 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  $n \rightarrow \infty$ , with the run length of each of the systems equal to  $n$ ); that is, for each problem above, the joint probability that all of our confidence intervals simultaneously cover the true values is, in the limit, at least  $1 - \alpha$ , where  $\alpha$  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, b) 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) attacked the problem of comparing normally distributed populations by using common random numbers to reduce the variance. Also, Yuan and Nelson (1993) considered 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 studied 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, nonnormally 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  $\sigma$  (in a manner akin to the  $t$ -statistic) rather than to estimate it consistently. This is desirable because consistent estimation of  $\sigma$  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 Damerджи and Nakayama (1996) developed some two-stage multiple-comparison procedures for steady-state simulations. Also, Nakayama (1996a) presented (without proof) MCB confidence intervals for single-stage steady-state simulations using the method of batch means. Nakayama (1996b) studied 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  $1, 2, \dots, k$ . For system  $i = 1, 2, \dots, k$ , let  $\mathbf{Y}_i = \{Y_i(t): t \geq 0\} \in D_1[0, \infty)$  be a real-valued (measurable) stochastic process representing the simulation output of system  $i$ , where  $D_1[0, \infty)$  is the space of right-continuous 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  $\{Y_{i,l}: l = 0, 1, 2, \dots\}$  by taking  $Y_i(t) = Y_{i, \lfloor t \rfloor}$ , where  $\lfloor \beta \rfloor$  denotes the greatest integer less than or equal to  $\beta \in \mathfrak{R}$ . Let  $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_k)$  and  $Y(t) = (Y_1(t), Y_2(t), \dots, Y_k(t))$ .

We assume that the processes  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_k$  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 nonoverlapping streams of uniform random numbers.) We allow for the distributions of  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_k$  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.

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

$$X_n \Rightarrow \Sigma B$$

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

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

with

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

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

Since we assumed that the  $\mathbf{Y}_i$ ,  $i = 1, 2, \dots, k$ , are mutually independent, the off-diagonal elements of the matrix  $\Sigma$  are all 0. Note that both  $X_n$  and  $\bar{Y}_n = (\bar{Y}_{1,n}, \bar{Y}_{2,n}, \dots, \bar{Y}_{k,n})$  lie in  $C[0, 1]$ , the space of continuous  $\mathfrak{R}^k$ -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 parameters 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  $t \rightarrow \infty$ , and so the  $\mu_i$ ,  $i = 1, 2, \dots, k$ , appearing in Assumption A1 are precisely the steady-state means of the process  $\mathbf{Y}$ . Also, Assumption 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  $n \rightarrow \infty$ . Recalling that  $B_i(1)$  has a standard normal distribution, we see that  $\sigma_i$  is the asymptotic variance parameter of the process  $\mathbf{Y}_i$ . In addition,  $B = (B_1, B_2, \dots, B_k)$ , where each  $B_i$  is a standard one-dimensional Brownian motion and  $B_1, B_2, \dots, B_k$  are mutually independent. By Billingsley [(1968), Theorem 3.2 and pages 26–27] Assumption A1 is then equivalent to requiring that

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

as  $n \rightarrow \infty$  for each  $i = 1, 2, \dots, k$ .

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 Ethier and Kurtz (1986), Chapter 7];
3.  $\mathbf{Y}$  satisfies appropriate mixing conditions [see Ethier and Kurtz (1986), Chapter 7]; 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 (nonoverlap-

ping) 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  $\sigma_i$ . [Glynn and Whitt (1990) showed 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  $\mathfrak{R}$ -valued continuous functions on  $[0, 1]$  and let  $B_1$  be a standard one-dimensional Brownian motion. Also, for a (measurable) function  $h: C_1[0, 1] \rightarrow S$  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 the following assumption holds.

**ASSUMPTION A2.** The (measurable) function  $g: C_1[0, 1] \rightarrow \mathfrak{R}$  satisfies the following conditions:

- (i)  $g(\alpha x) = \alpha g(x)$  for  $\alpha > 0$ ,  $\alpha \in \mathfrak{R}$ ,  $x \in C_1[0, 1]$ ;
- (ii)  $g(x - \beta e_1) = g(x)$  for  $\beta \in \mathfrak{R}$  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  $\sigma_i$  in the sense that if we multiply all of the observations of  $Y_i$  by some constant  $\alpha$ , then the new asymptotic variance parameter will be  $\alpha\sigma_i$ . This property will allow us to “cancel out” the asymptotic variance constant  $\sigma_i$ ; for more details, see the proof of Theorem 1 in Section 5. Condition (ii) guarantees that  $g(X_{i,n})$  does not depend on the unknown parameter  $\mu_i$ . Conditions (iii) and (iv) are technical assumptions required to invoke the continuous mapping principle (Proposition 1 in Section 5).

Observe that for each system  $i = 1, 2, \dots, k$ ,  $X_{i,n} = n^{1/2}(\bar{Y}_{i,n} - \mu_i e_1)$ . 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  $i$ ,  $i = 1, 2, \dots, k$ , with a run length  $n$ , where  $n$  is large. For each system  $i$ , we analyze its simulation output as follows. We divide the simula-

tion output of system  $i$  into  $m_i \geq 1$  (nonoverlapping) batches, each of length  $n/m_i$ , and apply a function  $g_i$  to the output, where  $g_i$  satisfies Assumption A2. Then, we 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

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

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 Assumption A2 using  $m_i$  batches, we define the constant  $\nu_{i,\gamma} \geq 0$  such that

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

where  $F_{\chi^2}$  denotes the distribution function of a  $\chi^2$  random variable with 1 degree of freedom. The continuity of  $F_{\chi^2}$  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

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

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  $i = 1, 2, \dots, k$ ,  $g_i$  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*

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

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 the construction of simultaneous confidence intervals for all contrasts  $\sum_{i=1}^k c_i \mu_i$  with  $c = (c_1, c_2, \dots, c_k) \in C^k$ , where  $C^k = \{c \in \mathfrak{R}^k: \sum_{i=1}^k c_i = 0\}$  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 Hochberg and Tamhane [(1987), pages 81–82].

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

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

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

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

$$I_c(n) = \left[ \sum_{i=1}^k c_i \hat{\mu}_i(n) - \frac{2}{\sum_{l=1}^k |c_l|} \sum_{i=1}^k \sum_{j=1}^k c_i^+ c_j^- \left( \frac{\nu_{i, \beta/2}^2 S_i^2(n)}{m_i} + \frac{\nu_{j, \beta/2}^2 S_j^2(n)}{m_j} \right)^{1/2}, \right. \\ \left. \sum_{i=1}^k c_i \hat{\mu}_i(n) + \frac{2}{\sum_{l=1}^k |c_l|} \sum_{i=1}^k \sum_{j=1}^k c_i^+ c_j^- \left( \frac{\nu_{i, \beta/2}^2 S_i^2(n)}{m_i} + \frac{\nu_{j, \beta/2}^2 S_j^2(n)}{m_j} \right)^{1/2} \right]$$

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  $i = 1, 2, \dots, k$ ,  $g_i$  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*

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

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 to construct confidence intervals simultaneously for the individual means. For example, we may want joint confidence intervals for  $\mu_1 - \mu_2$  and  $\mu_1$ . This may be accomplished by developing simultaneous confidence intervals for all linear combinations of means as done by Hochberg and Tamhane [(1987), pages 183–186]. 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  $i \neq k$  simultaneously to the control. For example, system  $k$  may represent a system already in place, and systems  $1, 2, \dots, k - 1$  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  $\mu_i - \mu_k$ ,  $i = 1, 2, \dots, k - 1$ , with  $\beta = 1 - (1 - \alpha)^{1/(k-1)}$ . Similarly, we define upper one-sided confidence intervals as

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

and lower one-sided confidence intervals as

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

for  $\mu_i - \mu_k$ ,  $i = 1, 2, \dots, k - 1$ . (We use subscripts  $u$  and  $l$  on  $I$  to denote that the one-sided confidence intervals 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  $i = 1, 2, \dots, k$ ,  $g_i$  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,*

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

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

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

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

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

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



confidence interval

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

for  $\mu_i - \max_{j \neq i} \mu_j$ ,  $i = 1, 2, \dots, k$ , 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  $i = 1, 2, \dots, k$ ,  $g_i$  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*

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

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  $g: C_1[0, 1] \rightarrow \mathfrak{R}$  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  $d \in \mathfrak{R}$  with  $d > 0$ . Thus, if we divide the output of system  $i$  into  $m_i \geq 2$  (nonoverlapping) 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  $j$ th (nonoverlapping) 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  $\chi^2$  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  $\gamma$  point of a Student's  $t$ -distribution with  $m$  degrees of freedom; that is,  $P\{t_m \geq t_{m,\gamma}\} = \gamma$ , 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  $g: C_1[0, 1] \rightarrow \mathfrak{R}$  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  $m_i \geq 1$  (nonoverlapping) 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) showed that  $12m_i g_i^2(B_i)$  has a  $\chi^2$  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  $\Gamma: C_1[0, 1] \rightarrow C_1[0, 1]$  and  $\Lambda_j: C_1[0, 1] \rightarrow C_1[0, 1]$  for  $j = 0, 1, \dots, m - 1$  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

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

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  $m_i \geq 1$  (nonoverlapping) batches and apply function  $g_i$  as above to  $X_{i,n}$ , we obtain

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

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  $\chi^2$  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 Billingsley (1968), Theorem 5.1, 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  $n \rightarrow \infty$ . Consider a (measurable) function  $h: C[0, 1] \rightarrow S$ ,  $S$  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  $n \rightarrow \infty$ .*

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

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

1, 2, \dots, d). Then

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

for  $(\beta_1, \beta_2, \dots, \beta_{d-1}) \in \mathfrak{R}^d$ .

The next lemma was established by Tamhane (1977).

LEMMA 3. Let  $W_1, W_2, \dots, W_d$  be mutually independent real-valued random variables and let  $\Psi = (\Psi_1, \Psi_2, \dots, \Psi_p): \mathfrak{R}^d \rightarrow \mathfrak{R}^p$  be a nonnegative function, where for each  $i = 1, 2, \dots, p$ ,  $\Psi_i$  is nondecreasing in each of its arguments. Then

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

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

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

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

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,

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

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

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

PROOF OF THEOREM 1. Note that

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

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

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

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

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

for  $i < j$ , and  $\mathbf{0} \in \mathfrak{R}^{k(k-1)/2}$  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  $n \rightarrow \infty$ .

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

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

by Assumption A2(i). Recall that standard one-dimensional Brownian motion has normal increments. In addition,  $B_1, B_2, \dots, B_k$  are mutually indepen-

dent, which implies that  $g_1(B_1), g_2(B_2), \dots, g_k(B_k)$  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  $\eta \in \mathfrak{R}$ ,

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

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

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

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

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

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

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

by Lemma 2. Now define

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

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

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

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

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

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

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

LEMMA 5. Let  $W = (W_1, W_2, \dots, W_d)$ ,  $d > 1$ , be multivariate normal with mean vector  $\mathbf{0}$  and covariance matrix  $\Theta = (\Theta_{i,j}; i, j = 1, 2, \dots, d)$ . Also, let  $V = (V_1, V_2, \dots, V_d)$  be multivariate normal with mean vector  $\mathbf{0}$  and covariance matrix  $\Xi = (\Xi_{i,j}; i, j = 1, 2, \dots, d)$ . If  $\Theta_{i,j} \geq \Xi_{i,j}$  for all  $1 \leq i < j \leq k$ , then

$$P\{W_i \leq \beta_i, i = 1, 2, \dots, d\} \geq P\{V_i \leq \beta_i, i = 1, 2, \dots, d\}$$

for  $(\beta_1, \beta_2, \dots, \beta_{k-1}) \in \mathfrak{R}^d$ .

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.  $\square$

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), ..., (k) such that  $\mu_{(1)} \leq \mu_{(2)} \leq \dots \leq \mu_{(k)}$ ; that is, system (j) has the jth smallest steady-state mean. Then, define the events

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

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  $\lim_{n \rightarrow \infty} P(E(n)) \geq 1 - \alpha$  by Theorem 3(iii). Now following an argument developed by Edwards and Hsu (1983), we show that  $E(n) \subset E_1(n) \cap E_2(n)$  for all  $n$ , which will establish the result.

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

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

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,

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

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



**Acknowledgment.** 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 Ser. A* **23** 359–378.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.
- BRATLEY, P., FOX, B. L. and SCHRAGE, L. E. (1987). *A Guide to Simulation*, 2nd ed. Springer, New York.
- CARLSTEIN, E. (1986). The use of subseries values for estimating the variance of a general statistic from a stationary sequence. *Ann. Statist.* **14** 1171–1179.
- CHEN, B.-C. and SARGENT, R. G. (1987). Using standardized time series to estimate confidence intervals for the difference between two stationary stochastic processes. *Oper. Res.* **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. NJIT Research Report CIS-96-30, Newark, NJ.
- EDWARDS, D. G. and HSU, J. C. (1983). Multiple comparisons with the best treatment *J. Amer. Statist. Assoc.* **78** 965–971.
- ETHIER, S. N. and KURTZ, T. G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.
- GAMES, P. A. (1977). An improved table for simultaneous control of  $g$  contrasts. *J. Amer. Statist. Assoc.* **72** 531–534.
- GLYNN, P. W. (1990). Diffusion approximations. In *Handbooks in Operations Research and Management Science* (D. Heyman and M. Sobel, eds.) **2** 145–198. Elsevier Science Publishers, New York.
- GLYNN, P. W. and IGLEHART, D. L. (1990). Simulation output analysis using standardized time series. *Math. Oper. Res.* **14** 1–16.
- GLYNN, P. W. and WHITT, W. (1987). Sufficient conditions for functional-limit-theorem versions of  $L = \lambda W$ . *Queueing Systems: Theory and Applications* **1** 279–287.
- GLYNN, P. W. and WHITT, W. (1990). Estimating the asymptotic variance with batch means. *Oper. Res. Lett.* **10** 431–435.
- GOLDSMAN, D., KANG, K. and SEILA, A. F. (1993). Cramér–von Mises variance estimators for simulations. Technical report, School of Industrial Systems Engineering, Georgia Inst. Technology, Atlanta.
- GOLDSMAN, D. and SCHRUBEN, L. W. (1984). Asymptotic properties of some confidence interval estimators for simulation output. *Management Sci.* **30** 1217–1225.
- GOLDSMAN, L. and NELSON, B. L. (1990). Batch-size effects on simulation optimization using multiple comparisons with the best. In *Proceedings of the 1990 Winter Simulation Conference* (O. Balci, R. P. Sadowski and R. E. Nance, eds.) 288–293. Institute of Electrical and Electronics Engineers, Piscataway, NJ.
- HOCHBERG, Y. and TAMHANE, A. C. (1987). *Multiple Comparison Methods*. Wiley, New York.
- HSU, J. C. (1981). Simultaneous confidence intervals for all distances from the “best.” *Ann. Statist.* **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)* (T. J. Santner and A. C. Tamhane, eds.) 23–33. Dekker, New York.
- HSU, J. C. (1984b). Constrained simultaneous confidence intervals for multiple comparisons with the best. *Ann. Statist.* **12** 1136–1144.
- MILLER, R. G. (1981). *Simultaneous Statistical Inference*, 2nd ed. Springer, New York.
- NAKAYAMA, M. K. (1994). Two-stage stopping procedures based on standardized time series. *Management Sci.* **40** 1189–1206.

- NAKAYAMA, M. K. (1996a). Multiple comparisons with the best in steady-state simulations. In *Proceedings of the Second International Workshop on Mathematical Methods in Stochastic Simulation and Experimental Design* (S. M. Ermakov and V. B. Melas, eds.) 230–235. Publishing House of Saint Petersburg Univ., Saint Petersburg, Russia.
- 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 Dept., New Jersey Inst. Technology.
- NELSON, B. L. (1993). Robust multiple comparisons under common random numbers. *ACM Trans. Modeling Computer Simulation* **3** 225–243.
- NELSON, B. L. and HSU, J. C. (1993). Control-variate models of common random numbers for multiple comparison with the best. *Management Sci.* **39** 989–1001.
- NEWMAN, C. M. and WRIGHT, A. L. (1981). An invariance principle for certain dependent sequences. *Ann. Probab.* **9** 671–675.
- SARGENT, R. G., KANG, K. and GOLDSMAN, D. (1992). An investigation of finite sample behavior of confidence interval estimation procedures. *Oper. Res.* **40** 898–913.
- SCHMEISER, B. (1982). Batch size effects in the analysis of simulation output. *Oper. Res.* **30** 556–568.
- SCHRUBEN, L. W. (1982). Detecting initialization bias in simulation output. *Oper. Res.* **30** 569–590.
- SCHRUBEN, L. W. (1983). Confidence interval estimation using standardized time series. *Oper. Res.* **31** 1090–1108.
- ŠIDÁK, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *J. Amer. Statist. Assoc.* **62** 626–633.
- SLEPIAN, D. (1962). The one-sided barrier problem for Gaussian noise. *Bell Syst. Tech. J.* **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. *Comm. Statist. Theory Methods* **6** 15–32.
- TAMHANE, A. C. (1979). A comparison of procedures for multiple comparisons of means with unequal variances. *J. Amer. Statist. Assoc.* **74** 471–480.
- TOKOL, G., GOLDSMAN, D., OCKERMAN, D. J. and SWAIN, J. J. (1996). Standardized time series  $L_p$ -norm variance estimators for simulations. *Management Sci.* To appear.
- TUKEY, J. W. (1953). The problem of multiple comparisons. Mimeographed monograph.
- YANG, W. and NELSON, B. L. (1991). Using common random numbers and control variates in multiple-comparison procedures. *Oper. Res.* **39** 583–591.
- YUAN, M. and NELSON, B. L. (1993). Multiple comparisons with the best for steady-state simulation. *ACM Trans. Modeling Computer Simulat.* **3** 66–79.

DEPARTMENT OF COMPUTER AND  
INFORMATION SCIENCE  
NEW JERSEY INSTITUTE OF TECHNOLOGY  
NEWARK, NEW JERSEY 07102  
E-MAIL: marvin@cis.njit.edu