Multiple Comparisons with the Best Using Common Random Numbers 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. We consider the problem of running a single-stage simulation using common random numbers to construct simultaneous confidence intervals for $\mu_i - \max_{j \neq i} \mu_j$, $i = 1, 2, \ldots, k$. This is known as multiple comparisons with the best (MCB). Under an assumption that the stochastic processes representing the simulation output of the different systems satisfy a functional central limit theorem, we prove that our confidence intervals are asymptotically valid (as the run lengths of the simulations of each system tends to infinity). We develop algorithms for two different cases: when the asymptotic covariance matrix has sphericity, and when the covariance matrix is arbitrary.

AMS 1991 *subject classifications.* Primary 68U20, 62J15; secondary 65C05, 62M10, 60F17.

Key words and phrases. Stochastic simulation; output analysis; functional central limit theorem; common random numbers; multiple comparisons.

1 Introduction

Suppose that there are $k \geq 2$ different systems (i.e., stochastic processes) having (unknown) steady-state means $\mu_i,\ i = 1, 2, \ldots, k$. We want to compare the systems based on their steady-state means, where we assume that larger means are better. (The situation when smaller means are more desirable can be similarly treated.) For example, the different systems may represent various designs for a manufacturing system, where the designs have different buffer allocations, and we want to compare the systems relative to their steady-state throughputs. We estimate the steady-state means by simulation using a variance-reduction technique known as common random numbers (CRN); e.g., see Section 2.1 of Bratley, Fox, and Schrage (1987). In many situations, CRN is known to significantly increase the statistical efficiency of a simulation comparing different systems.

We will develop simultaneous confidence intervals for $\mu_i - \max_{j \neq i} \mu_j,\ i = 1, 2, \ldots, k$. This problem is known as multiple comparisons with the best (MCB). Note that if $\mu_i - \max_{j \neq i} \mu_j > 0$, then system i is the best. On the other hand, if $\mu_i - \max_{j \neq i} \mu_j < 0$ and $\mu_i - \max_{j \neq i} \mu_j > -\epsilon$ for some $\epsilon > 0$, then system i is not the best but it is within ϵ of the best. MCB methods have the advantage over making all pairwise comparisons $\mu_i - \mu_j$, $i < j$, since only $k - 1$ comparisons are needed for MCB, as opposed to $k(k-1)/2$ for the all-pairwise case. This leads to sharper confidence intervals.

Hsu (1981) first developed a single-stage MCB procedure to compare the means of independent normal populations, where independent sampling is used within each population. Other work on single-stage MCB procedures for comparing independent normally-distributed populations includes Hsu (1984ab) and Edwards and Hsu (1983). Two-stage procedures for normally-distributed independent populations are examined by Matejcik and Nelson (1995). Nelson and Hsu (1993), Yang and Nelson (1991), and Nelson (1993) developed MCB methods to be used when the different normally-distributed populations are no longer independent but instead there is an induced correlation (i.e., common random numbers) among the populations. Nelson and Matejcik (1995) studied two-stage MCB procedures using common random numbers for normally-distributed populations. See Hochberg and Tamhane (1987) and Hsu (1996) for details of these and other MCB methods.

There also has been work on MCB procedures for steady-state simulations when the different systems are independent. (In practice, the independence is achieved by using non-overlapping streams of random numbers to drive the simulations of the different systems.) Yuan and Nelson (1993) studied single-stage MCB methods for the steady-state means of autoregressive pro-

cesses. Nakayama (1996ab) developed asymptotically valid MCB intervals for single-stage steady-state simulations when the systems are mutually independent assuming that a functional central limit theorem holds. Some two-stage multiple-comparison procedures for steady-state simulations without CRN are analyzed by Damerdji and Nakayama (1996).

All of the previous work on MCB with common random numbers assumes that the observations within a population are normally distributed and independent. We now extend those results by establishing the asymptotic validity of these methods for steady-state simulations, where typically there are autocorrelations present in the output processes. Our techniques are based on the method of (non-overlapping) batch means, an approach frequently applied in simulation practice (e.g., see Section 3.3.1 of Bratley, Fox, and Schrage 1987) and statistics (e.g., see Carlstein 1986).

To establish the asymptotic validity of our methods, we will assume that the systems jointly satisfy a functional central limit theorem. We will develop techniques for two different situations. The first method will be asymptotically valid when the asymptotic covariance matrix in the functional central limit theorem satisfies the property of sphericity (e.g., see pp. 208–210 of Hochberg and Tamhane 1987). Sphericity is often used to model the dependence among repeated measurements on a single subject in clinical trials (see pp. 208–209 of Hochberg and Tamhane 1987), and it has also been applied in the simulation literature as a way of modeling the effects of CRN (Nelson 1993, Nelson and Matejcik 1995). For cases when a sphericity assumption cannot be justified, we develop an alternative method based on the Bonferroni inequality.

The rest of the paper is organized as follows. In Section 2, we develop the notation and state our functional-central-limit-theorem assumption. Section 3 contains our MCB procedure under the sphericity assumption, and our Bonferroni method is presented in Section 4. Some empirical results are given in Section 5. All of the proofs are collected in Section 6.

2 Notation and Assumptions

We want to compare the steady-state behavior of $k \geq 2$ 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 (e.g., see Ethier and Kurtz 1986 or Glynn 1990). Essentially all stochastic processes used in applications possess sample paths in $D_1[0, \infty)$. Discrete-time processes $\{Y_{i,l} : l = 0, 1, 2, \ldots\}$ can be handled by taking $Y_i(t) = Y_{i,\lfloor t \rfloor}$, where $\lfloor \beta \rfloor$ denotes the

greatest integer less than or equal to $\beta \in \Re$.

We assume that the processes are dependent, where the dependence arises from using common random numbers. This entails using the same stream of random numbers to drive the simulations of the different systems; see Section 2.1 of Bratley, Fox, and Schrage (1987) for details on CRN. Let $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k)^T$ and $Y(t) = (Y_1(t), Y_2(t), \ldots, Y_k(t))^T$, where the superscript T denotes the transpose. (Throughout the paper all vectors are assumed to be column vectors.) We then have that $\mathbf{Y} \in D[0,\infty)$, where $D[0,\infty)$ is the space of right-continous $\Re^k$-valued functions on $[0,\infty)$ having left limits.

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

A1 *There exist a (unknown) nonsingular matrix* $\Gamma = (\Gamma_{i,j} : 1 \leq i,j \leq k) \in \Re^{k \times k}$ *and a (unknown) constant* $\mu = (\mu_1, \mu_2, \ldots, \mu_k)^T \in \Re^k$ *such that* $X_n \Rightarrow \Gamma B$ *as* $n \to \infty$, *where* B *is a standard* k*-dimensional Brownian motion,* $X_n = (X_{1,n}, X_{2,n}, \ldots, X_{k,n})^T$, *and* $X_{i,n} = \{X_{i,n}(s) : 0 \leq s \leq 1\}$ *is defined as* $X_{i,n}(s) = n^{1/2}\left(\bar{Y}_{i,n}(s) - \mu_i s\right)$ *with* $\bar{Y}_{i,n} = \{\bar{Y}_{i,n}(s) : 0 \leq s \leq 1\}$ *defined as* $\bar{Y}_{i,n}(s) = \frac{1}{n}\int_0^{ns} Y_i(t)\,dt$ *for* $i = 1, 2, \ldots, k$.

Both X_n and $\bar{Y}_n = (\bar{Y}_{1,n}, \bar{Y}_{2,n}, \ldots, \bar{Y}_{k,n})^T$ 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 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 $\mathbf{Y}$.

Assumption A1 and the converging-together lemma guarantee that for each i,

$$\frac{1}{n}\int_0^n Y_i(t)\,dt - \mu_i \;=\; \frac{1}{\sqrt{n}} X_{i,n}(1) \;\Rightarrow\; 0 \cdot (\Gamma B)_i(1) \;=\; 0$$

as $n \to \infty$, where $(C)_i$ denotes the i-th element of a vector C. Hence, the μ_i appearing in A1 are precisely the steady-state means of the process $\mathbf{Y}$. Also, recall that by definition, a standard k-dimensional Brownian motion satisfies $B = (B_1, B_2, \ldots, B_k)^T$, where each B_i is a standard 1-dimensional Brownian motion with $B_1, B_2, \ldots, B_k$ mutually independent.

Several authors have used a FCLT assumption in analyses of steady-state simulations; e.g., see Glynn and Iglehart (1990) and Nakayama (1994,1996ab). Also, functional central limit theorems have been established for a rich class of stochastic processes, including Markov processes in discrete and continuous time, stationary processes satisfying appropriate mixing conditions, and generalized

semi-Markov processes; see Glynn and Iglehart (1990), Chapter 7 of Ethier and Kurtz (1986), Chapter 11 of Stroock and Varadhan (1979), and Haas (1997) for further details. (The last three references specifically address the multivariate case, which is appropriate in our setting.)

We will develop procedures under two different assumptions on the asymptotic covariance matrix $\Sigma = \Gamma\Gamma^T$ of $\mathbf{Y}$. First, we will assume that Σ has a special structure known as sphericity, and then we will allow Σ to be arbitrary. In both cases, we do not assume that Γ (or μ) is known.

3 MCB Procedure Under Sphericity

We first assume that the covariance matrix $\Sigma = (\Sigma_{i,j} : i, j = 1, 2, \ldots, k)$ has the following form known as sphericity, which was proposed by Huynh and Feldt (1970):

A2 $\Sigma_{i,i} = 2\psi_i + \tau^2$ *for* $i = 1, 2, \ldots, k$ *and* $\Sigma_{i,j} = \psi_i + \psi_j$ *for* $i \neq j$, *where* $\tau > 0$ *and* ψ_i, $i = 1, 2, \ldots, k$, *are (unknown) constants such that* $\tau^2 > (k \sum_{i=1}^k \psi_i^2)^{1/2} - \sum_{i=1}^k \psi_i$ *(which ensures that* Σ *is positive definite).*

Grieve and Ag (1984) developed an empirical test to determine if a sample covariance matrix has sphericity. Also, sphericity generalizes compound symmetry (i.e., $\Sigma_{i,i} = 1$ for $i = 1, 2, \ldots, k$, and $\Sigma_{i,j} = \rho$ for $i \neq j$), which was used by Schruben and Margolin (1978), Nozari et al. (1987), and Tew and Wilson (1993) to model the effects of common random numbers. Finally, for his MCB procedures using CRN with normally-distributed populations, Nelson (1993) also used a sphericity assumption, and he showed empirically that the coverage of his resulting confidence intervals was robust as long as all of the covariances are positive. Thus, from a practical viewpoint, it seems that Assumption A2 is not overly restrictive.

The following MCB procedure, which generalizes a method developed in Nelson (1993), is to be used under A2.

Procedure S:

1. Specify the number of systems k, the desired confidence level $1 - \alpha$, the number of batches $m \geq 2$, and the run length n to be used for all systems, where n is large.

2. Simulate the k systems using common random numbers for run length n.

3. Compute

$$S_s^2(n) = \frac{1}{(k-1)(m-1)} \sum_{i=1}^{k} \sum_{j=1}^{m} \left(Z_{i,j}(n) - \hat{\mu}_i(n) - \frac{1}{k} \sum_{l=1}^{k} Z_{l,j}(n) + \frac{1}{k} \sum_{l=1}^{k} \hat{\mu}_l(n) \right)^2,$$

where

$$Z_{i,j}(n) = \frac{1}{n/m} \int_{(j-1)n/m}^{jn/m} Y_i(t)\, dt,\ j \geq 1,$$

is the j-th (non-overlapping) batch mean of size n/m for system i, and $\hat{\mu}_i(n) = \frac{1}{n} \int_0^n Y_i(t)\, dt$ is the sample mean of system i.

4. Define the MCB confidence intervals

$$I_{s,i}(n) = \left[\left(\hat{\mu}_i(n) - \max_{j \neq i} \hat{\mu}_j(n) - \nu S_s(n) \sqrt{\frac{2}{m}} \right)^-,\ \left(\hat{\mu}_i(n) - \max_{j \neq i} \hat{\mu}_j(n) + \nu S_s(n) \sqrt{\frac{2}{m}} \right)^+ \right] \tag{1}$$

for $\mu_i - \max_{j \neq i} \mu_j$, $i = 1, 2, \ldots, k$, where $x^- = \min(x, 0)$, $x^+ = \max(x, 0)$, and $\nu = \nu(\alpha, k, m)$ is the upper $(1-\alpha)$-quantile point of the maximum of a $(k-1)$-variate t distribution with $(k-1)(m-1)$ degrees of freedom and common correlation $1/2$; see pp. 374–375 and Table 4 of Hochberg and Tamhane (1987) for more details on the multivariate t distribution.

The following result shows that the MCB confidence intervals formed in the above method are asymptotically valid.

Theorem 1 *If Assumptions A1 and A2 hold, then* $\lim_{n \to \infty} P\left\{ \mu_i - \max_{j \neq i} \mu_j \in I_{s,i}(n),\ \forall\ i = 1, 2, \ldots, k \right\} \geq 1 - \alpha$.

4 Bonferroni-MCB Procedure

If the covariance matrix Σ does not satisfy Assumption A2 (or if we are not comfortable with this assumption), then we can instead use the following procedure, which is based on the Bonferroni inequality. It places no restrictions on the structure of Σ, and it is based on an idea of Clark and Yang (1986).

Procedure B:

1. Specify the number of systems k, the desired confidence level $1 - \alpha$, the number of batches $m \geq 2$, and the run length n to be used for all systems, where n is large.

2. Simulate the k systems using common random numbers for run length n.

3. Compute

$$S^2_{b,i,j}(n) = \frac{1}{m-1}\sum_{l=1}^{m}\left(Z_{i,l}(n) - Z_{j,l}(n) - (\hat{\mu}_i(n) - \hat{\mu}_j(n))\right)^2$$

for all $i \neq j$.

4. Define the MCB confidence intervals

$$I_{b,i}(n) = \left[\left(\min_{\substack{j\in\mathcal{C}(n)\\ j\neq i}}\left(\hat{\mu}_i(n) - \hat{\mu}_j(n) - \phi\frac{S_{b,i,j}(n)}{\sqrt{m}}\right)\right)^-,\ \left(\min_{j\neq i}\left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + \phi\frac{S_{b,i,j}(n)}{\sqrt{m}}\right)\right)^+\right] \tag{2}$$

for $\mu_i - \max_{j\neq i}\mu_j$, $i = 1, 2, \ldots, k$, where $\phi = \phi(\alpha, k, m)$ is the upper $(1-\alpha/(k-1))$-quantile point of a univariate t distribution with $m-1$ degrees of freedom, and

$$\mathcal{C}(n) = \left\{i : \min_{j\neq i}\left(\hat{\mu}_i(n) - \hat{\mu}_j(n) + \phi\frac{S_{b,i,j}(n)}{\sqrt{m}}\right) \geq 0\right\}.$$

In (2), we define $\min_{j\in\emptyset} x_j = 0$.

Theorem 2 *If Assumption A1 holds, then* $\lim_{n\to\infty} P\left\{\mu_i - \max_{j\neq i}\mu_j \in I_{b,i}(n),\ \forall\ i = 1, 2, \ldots, k\right\} \geq 1-\alpha$.

As the number of systems k gets large, the conservativeness of the Bonferroni inequality overwhelms the benefits of common random numbers. Thus, Procedure B should only be used when k is small; i.e., $k \leq 5$.

5 Empirical Results

We now present some results from simulating four different M/M/1 queueing systems, labeled 1, ..., 4. The mean interarrival time of customers in each of the systems is 1, and the mean service times for systems 1–4 are 0.3, 0.4, 0.45, and 0.5, respectively. Let μ_i be the steady-state expected number of customers in system i, $i = 1, \ldots, 4$, and we compared the different systems in terms of this performance measure by constructing our MCB confidence intervals in (1) and (2). For other empirical work on MCB using batch means (but without common random numbers), see Goldsman and Nelson (1990).

We implemented common random numbers in two different ways. First, we used CRN for everything (i.e., both the interarrival times and the service times). Next, only the interarrival times for the different systems were generated using CRN; the service times among the different systems are independent.

	Observed coverages when using CRN for	
Run length	everything	only interarrival times
10^2	0.600	0.777
10^3	0.728	0.840
10^4	0.821	0.851
10^5	0.826	0.875

Table 1: Observed coverages for Procedure S

	Observed coverages when using CRN for	
Run length	everything	only interarrival times
10^2	0.692	0.892
10^3	0.854	0.917
10^4	0.926	0.901
10^5	0.942	0.919

Table 2: Observed coverages for Procedure B

We performed coverage experiments by running 2000 independent replications, where in each replication nominal 90% MCB confidence intervals are constructed using 5 batches for each system. We performed this experiment for four different run lengths n ranging from 10^2 to 10^5.

Tables 1 and 2 contain the results when applying Procedures S and B, respectively. The second column gives the observed coverage when CRN was used for everything, and the third column is for when CRN was used only for the interarrival times. Note that for small run lengths, most of the observed coverages for both Procedures S and B are significantly below the nominal level. For the largest run lengths, the observed coverage for Procedure S is still somewhat below the nominal level, but Procedure B's is above 90%.

In general the observed coverages for Procedure S (and also for Procedure B) are closer to the nominal level for each run length when CRN is not used for everything. Thus, we conclude that in practice, the performance of both procedures may depend on the way that CRN is implemented.

6 Proofs

Lemma 1 *Suppose Assumption 1 holds, and define the function* $g_s : C[0,1] \to \Re$ *as*

$$g_s(x) = \left[\frac{m}{(k-1)(m-1)} \sum_{i=1}^{k} \sum_{j=1}^{m} \left(\Delta_m x_i \left(\frac{j}{m} \right) - \frac{x_i(1)}{m} - \frac{1}{k} \sum_{l=1}^{k} \Delta_m x_l \left(\frac{j}{m} \right) + \sum_{l=1}^{k} \frac{x_l(1)}{km} \right)^2 \right]^{1/2},$$

where $\Delta_d x_i(t) = x_i(t) - x_i(t - 1/d)$ *for* $d \in \Re$ *with* $d > 0$. *If* Σ *satisfies Assumption A2, then* $g_s^2(\Gamma B)$ *has the same distribution as* $\tau^2 \chi^2_{(k-1)(m-1)}/((k-1)(m-1))$ *and is independent of* $\Gamma B(1)$, *where* $\chi^2_{(k-1)(m-1)}$ *is a* χ^2 *random variable with* $(k-1)(m-1)$ *degrees of freedom.*

Proof. Consider any $x \in C[0,1]$ satisfying $x(0) = 0$. (When applying the function g_s to X_n or B, this corresponds to requiring that $X_n(0) = 0$ or $B(0) = 0$, which hold by definition.) We start by examining the summands in the definition of $g_s(x)$. The second term of g_s satisfies $x_i(1)/m = (1/m)\sum_{l=1}^{m} \Delta_m x_i\,(l/m)$, which is the sample mean of the $\Delta_m x_i(l/m)$, $l = 1, 2, \ldots, m$. Similarly, the fourth term satisfies $\sum_{l=1}^{k} x_l(1)/(km) = (1/km)\sum_{l=1}^{k}\sum_{p=1}^{m} \Delta_m x_l\,(p/m)$, which is the sample mean of the $\Delta_m x_l(p/m)$, $l = 1, 2, \ldots, k$, $p = 1, 2, \ldots, m$.

Recall that standard 1-dimensional Brownian motion has independent and normally distributed increments. Thus, $B(1)$ has a standard k-dimensional normal distribution, and so $\Gamma B(1)$ is normally distributed with mean 0 and covariance matrix Σ. Also, after bringing the m from outside the summations inside in the definition of g_s, we see that $m^{1/2}(\Gamma B(i/m) - \Gamma B((i-1)/m))$ has a normal distribution with mean 0 and covariance matrix Σ. The result then follows from Theorem 1 of Huynh and Feldt (1970) (also see Hochberg and Tamhane 1987, pp. 210–211). ∎

Proof of Theorem 1. For our given collection of k systems, we define $(1), (2), \ldots, (k)$ such that $\mu_{(1)} \le \mu_{(2)} \le \cdots \le \mu_{(k)}$; i.e., system (j) has the j-th smallest steady-state mean. Now consider the event

$$A(n) = \{\mu_i - \mu_{(k)} \ge \hat{\mu}_i(n) - \hat{\mu}_{(k)}(n) - H(n),\ \forall\ i \ne (k)\},$$

where $H(n) = \nu S_s(n)\sqrt{\frac{2}{m}}$. By slightly modifying an argument used in Hsu (1984b) (also see Hochberg and Tamhane 1987, pp. 150–151) we can show that

$$A(n) \subset \{\mu_i - \max_{j \ne i} \mu_j \in I_{s,i}(n)\} \tag{3}$$

for all n.

We now prove that

$$\lim_{n\to\infty} P\{A(n)\} = 1 - \alpha. \tag{4}$$

It is straightforward to show that $g_s(X_n) = S_s(n)\sqrt{n/m}$, and so

$$P\{A(n)\} \quad = \quad P\left\{(\hat{\mu}_i(n) - \mu_i) - (\hat{\mu}_{(k)}(n) - \mu_{(k)}) \le \nu g_s(X_n)\sqrt{\frac{2}{n}},\ \forall i \ne (k)\right\}$$

$$
\begin{aligned}
&= P\left\{\sqrt{n}\left(\hat{\mu}_i(n)-\mu_i\right)-\sqrt{n}\left(\hat{\mu}_{(k)}(n)-\mu_{(k)}\right)-\sqrt{2}\nu g_s(X_n)\le 0,\ \forall\ i\neq(k)\right\}\\
&= P\left\{u(X_n)\le \mathbf{0}\right\},
\end{aligned}
$$

where the function $u : C[0,1] \to \Re^{k-1}$ is defined as $u(x) = (x_i(1) - x_{(k)}(1) - \sqrt{2}\nu g_s(x) : i \neq (k))$, and $\mathbf{0} \in \Re^{k-1}$ is the $(k-1)$-dimensional vector with all components 0. By the fact that the projection mapping is continuous and since Γ is non-singular, u is continuous at ΓB with probability 1, and so the continuous mapping principle (see Theorem 5.1 of Billingsley 1968 or Glynn 1990) ensures that $u(X_n) \Rightarrow u(\Gamma B)$ as $n \to \infty$.

Now we will prove that

$$
P\left\{u(X_n)\le \mathbf{0}\right\} \to P\left\{u(\Gamma B)\le \mathbf{0}\right\} \text{ as } n\to\infty. \tag{5}
$$

By Theorem 2.1 of Billingsley (1968), it is sufficient to show that $u(\Gamma B)$ has a continuous distribution function. Note that $\Gamma B(1)$ has a k-dimensional normal distribution with mean 0 and covariance matrix $\Sigma = \Gamma\Gamma^T$. Now define the function $v : C[0,1] \to \Re^{k-1}$ as $v(x) = (x_i(1) - x_{(k)}(1) : i \neq (k))$. Then, it is straightforward to verify that under Assumption A2, $v(\Gamma B)$ has a $(k-1)$-dimensional normal distribution with mean 0 and covariance matrix $\Xi = (\Xi_{i,j} : i,j \neq (k))$, where $\Xi_{i,i} = 2\tau^2$ and $\Xi_{i,j} = \tau^2$ for $i \neq j$. Lemma 1 implies that $\Gamma B(1)$ is independent of $g_s(\Gamma B)$. Thus, since the distribution functions of the normal and χ^2 distributions are continuous, it follows that $u(\Gamma B)$ has a continuous distribution function, and so (5) holds.

Now we will show that $P\left\{u(\Gamma B)\le \mathbf{0}\right\} = 1-\alpha$. Let Φ denote a standard 1-dimensional normal distribution function, and let F denote the distribution function of $g_s(\Gamma B)/\tau$. By Lemma 1 and by the form of the covariance matrix Ξ of $v(\Gamma B)$, we get that

$$
\begin{aligned}
P\left\{u(\Gamma B)\le \mathbf{0}\right\} &= P\left\{\max_{i\neq(k)}\left(\frac{(\Gamma B)_i(1)-(\Gamma B)_{(k)}(1)}{\sqrt{2}\tau}\right)\le\frac{\sqrt{2}\nu g_s(\Gamma B)}{\sqrt{2}\tau}\right\}\\
&= \int_0^\infty\int_{-\infty}^\infty\prod_{l=1}^{k-1}\Phi\left(\frac{(\sqrt{1/2})z+\nu w}{\sqrt{1/2}}\right)d\Phi(z)\,dF(w) \;=\; 1-\alpha,
\end{aligned}
$$

where the last two equalities follow from equations (1.1a) and (1.2a), respectively, on pp. 374–375 of Hochberg and Tamhane (1987). Hence, we have established (4). Putting this together with (3) completes the proof. ∎

To prove Theorem 2, we will need the following result.

Lemma 2 *Suppose Assumption 1 holds, and define the function* $g_b = (g_{b,i,j} : i \neq j) : C[0,1] \to \Re^{k(k-1)}$ *as*

$$g_{b,i,j}(x) = \left[\frac{m}{m-1} \sum_{l=1}^{m} \left(\Delta_m x_i \left(\frac{l}{m} \right) - \Delta_m x_j \left(\frac{l}{m} \right) - \left(\frac{x_i(1)}{m} - \frac{x_j(1)}{m} \right) \right)^2 \right]^{1/2}$$

for $i \neq j$. *Then for each* $i \neq j$, $(m-1)g_{b,i,j}^2(\Gamma B)/\Lambda_{i,j}$ *has a* χ^2 *distribution with* $m-1$ *degrees of freedom and is independent of* $\Gamma B(1)$, *where* $\Lambda_{i,j} = \Sigma_{i,i} + \Sigma_{j,j} - 2\Sigma_{i,j}$.

Proof. Consider any $x \in C[0,1]$ satisfying $x(0) = 0$. Note that

$$\frac{x_i(1)}{m} - \frac{x_j(1)}{m} = \frac{1}{m} \sum_{l=1}^{m} \left(\Delta_m x_i \left(\frac{l}{m} \right) - \Delta_m x_j \left(\frac{l}{m} \right) \right),$$

which is the sample mean of the $\Delta_m x_i(l/m) - \Delta_m x_j(l/m)$, $l = 1, 2, \ldots, m$. Also, $m^{1/2}(\Gamma B(l/m) - \Gamma B((l-1)/m))$ has a k-variate normal distribution with mean 0 and covariance matrix Σ, which does not necessarily satisfy Assumption 2. Therefore, $Q_{i,j,l} \equiv m^{1/2} \left(\Delta_m(\Gamma B)_i(l/m) - \Delta_m(\Gamma B)_j(l/m) \right)$ is normally distributed with mean 0 and variance $\Lambda_{i,j} = \Sigma_{i,i} + \Sigma_{j,j} - 2\Sigma_{i,j}$. Also, since Brownian motion has independent increments, the $Q_{i,j,l}$, $l = 1, 2, \ldots, m$, are mutually independent. Thus, $(m-1)g_{b,i,j}^2(\Gamma B)/\Lambda_{i,j}$ has a χ^2 distribution with $m-1$ degrees of freedom.

Now we need to show that $(m-1)g_{b,i,j}^2(\Gamma B)/\Lambda_{i,j}$ is independent of $\Gamma B(1)$. Define the function $b = (b_{i,j} : i \neq j) : C[0,1] \to \Re^{k(k-1)}$ as

$$b_{i,j}(x) = \left[\frac{m}{m-1} \sum_{l=1}^{m} \left(\Delta_m x_i \left(\frac{l}{m} \right) - \Delta_m x_j \left(\frac{l}{m} \right) \right)^2 \right]^{1/2}$$

for $i \neq j$. Also, define the mapping $\Theta : C[0,1] \to C[0,1]$ to be $(\Theta x)(t) = x(t) - tx(1)$. It can be shown that $g_b = b \circ \Theta$. Since $B_1, B_2, \ldots, B_k$ are mutually independent, we can easily modify the proof on page 84 of Billingsley (1968) to show that process $\{B(t) - tB(1) : 0 \leq t \leq 1\}$ is independent of $B(1)$. Thus, $\{\Gamma B(t) - t\Gamma B(1) : 0 \leq t \leq 1\}$ is independent of $\Gamma B(1)$. This implies $\Theta(\Gamma B)$ is independent of $\Gamma B(1)$, and so $g_b(\Gamma B) = (b \circ \Theta)(\Gamma B)$ is independent of $\Gamma B(1)$. ∎

Proof of Theorem 2. Consider the event

$$E(n) = \{\mu_i - \mu_{(k)} \geq \hat{\mu}_i(n) - \hat{\mu}_{(k)}(n) - \phi S_{b,i,(k)}(n),\ \forall\ i \neq (k)\}.$$

By slightly modifying the proof of Theorem 4.1.2 in Hsu (1996), we can prove that $E(n) \subset \{\mu_i - \max_{j \neq i} \mu_j \in I_{b,i}(n)\}$ for all n.

We now show that

$$\lim_{n\to\infty} P\{E(n)\} \geq 1-\alpha \tag{6}$$

Note that $g_{b,i,j}(X_n) = S_{b,i,j}(n)\sqrt{n/m}$, and so using an argument similar to that applied in the proof of Theorem 1, we can establish that

$$P\{E(n)\} \to P\left\{(\Gamma B)_i(1) - (\Gamma B)_{(k)}(1) \leq \phi g_{b,i,(k)}(\Gamma B),\ \forall\ i \neq (k)\right\} \tag{7}$$

as $n \to \infty$.

Now we will show that the right-hand side of (7) is at least $1-\alpha$. Bonferroni's inequality implies that

$$\begin{aligned} & P\left\{(\Gamma B)_i(1) - (\Gamma B)_{(k)}(1) \leq \phi g_{b,i,(k)}(\Gamma B),\ \forall\ i \neq (k)\right\} \\ & \geq\ 1 - \sum_{i\neq(k)} P\left\{\frac{(\Gamma B)_i(1) - (\Gamma B)_{(k)}(1)}{g_{b,i,(k)}(\Gamma B)} > \phi\right\} \ =\ 1-\alpha, \end{aligned}$$

where the last step follows from Lemma 2 and the fact that $(\Gamma B)_i(1) - (\Gamma B)_{(k)}(1)$ is normally distributed with mean 0 and variance $\Lambda_{i,j}$. Hence, (6) holds, and the proof is complete. ∎

Acknowledgments

The author would like to thank the referees for providing comments that improved the quality of the paper. Also, 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.

References

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

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

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.

Clark, G. M. and W. N. Yang (1986). A Bonferroni selection procedure when using common random numbers with unknown variances. *Proceedings of the 1986 Winter Simulation Conference*, 313–315.

Damerdji, H. and M. K. Nakayama (1996). Two-stage multiple-comparison procedures for steady-state simulations. Research report CIS-96-30, Computer and Information Science Department, New Jersey Institute of Technology, Newark, NJ.

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*. John Wiley, New York.

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.

Goldsman, L. and B. L. Nelson (1990). Batch-size effects on simulation optimization using multiple comparisons with the best. *Proceedings of the 1990 Winter Simulation Conference*, 288–293.

Grieve, A. P. and C.-G. Ag (1984). Tests of sphericity of normal distributions and the analysis of repeated measures design. *Psychometrika* 49, 257–267.

Haas, P. (1997). On simulation output analysis for generalized semi-Markov processes. Preprint.

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

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: T. J. Santner and A. C. Tamhane, Eds., *Design of Experiments: Ranking and Selection (Essays in Honor of Robert E. Bechhofer)*, Marcel Dekker, New York, 23–33.

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

Hsu, J. C. (1996). *Multiple Comparisons: Theory and Methods*. Chapman & Hall, London.

Huynh, H. and L. S. Feldt (1970). Conditions under which mean square ratios in repeated measurement designs have exact F-distributions. *Journal of the American Statistical Association* 65, 1582–1589.

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

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

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

Nakayama, M. K. (1996b). Multiple-comparison procedures for steady-state simulations. *Annals of Statistics*, to appear.

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.

Nelson, B. L. and F. J. Matejcik (1995). Using Common Random Numbers for Indifference-Zone Selection and Multiple Comparisons in Simulation. *Management Science* 41, 1935–1945.

Nozari, A., S. Arnold, and C. Pegden (1987). Statistical analysis for use with the Schruben and Margolin correlation induction strategy. *Operations Research* 35, 127–139.

Schruben, L. W. and B. Margolin (1978). Pseudorandom number assignment in statistically designed simulation and distribution sampling experiments. *Journal of the American Statistical Association* 73, 504–525.

Stroock, D. W. and S. R. S. Varadhan (1979). *Multidimensional Diffusion Processes*. Springer-Verlag, New York.

Tew, J. D. and J. R. Wilson (1994). Estimating simulation metamodels using combined correlation-based variance reduction techniques. *IIE Transactions* 26, 2–16.

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.