Proceedings of the 1996 Winter Simulation Conference
ed. J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain

TWO-STAGE PROCEDURES FOR MULTIPLE COMPARISONS WITH A CONTROL IN STEADY-STATE SIMULATIONS

Halim Damerdji

Department of Industrial Engineering
North Carolina State University
Raleigh, NC 27695-7906, U.S.A.

Marvin K. Nakayama

Department of Computer and Information Science
New Jersey Institute of Technology
Newark, NJ 07102-1982, U.S.A.

ABSTRACT

Suppose that we have k different stochastic systems, where μ_i denotes the steady-state mean of system i. We assume that the system labeled k is a control and want to compare the performance of the other systems, labeled 1,2,...,$k-1$, relative to this control. This problem is known in the statistical literature as multiple comparisons with a control (MCC). Independent steady-state simulations will be performed to compare the systems to the control. Two-stage procedures, based on the method of batch means, are presented to construct simultaneous lower one-sided confidence intervals for $\mu_i - \mu_k$ $(i = 1, 2, \ldots, k)$, each having prespecified (absolute or relative) half-width δ. Under the assumption that the stochastic processes representing the evolution of the systems satisfy a functional central limit theorem, it can be shown that asymptotically (as $\delta \to 0$ with the size of the batches proportional to $1/\delta^2$), the joint probability that the confidence intervals simultaneously contain the $\mu_i - \mu_k$ $(i = 1, 2, \ldots, k-1)$ is at least $1 - \alpha$, where α is prespecified by the user.

1 INTRODUCTION

Suppose there are k different systems (i.e., stochastic processes), where system i has (unknown) steady-state mean μ_i and (unknown) asymptotic variance σ_i^2. We allow for the different systems to have different probability distributions governing them, and so the variances may be unequal. We assume that system k is a control and want to compare the mean performance of this system to that of each of the systems $i = 1, 2, \ldots, k-1$. For example, in a redesign setting, system k may be the existing system, and there are $(k-1)$ design alternatives under consideration. We assume here that larger performances are better than smaller. (The case when smaller performances are preferable can similarly be considered.)

Independent simulations of the various systems are run to construct simultaneous confidence intervals for $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$. This is known as *multiple comparisons with a control* (MCC). In this paper, we propose some two-stage procedures based on the method of batch means for constructing lower one-sided MCC confidence intervals having prespecified (absolute or relative) half-width δ. It can be shown (see Damerdji and Nakayama 1996) that asymptotically (as $\delta \to 0$ and the batch sizes are proportional to $1/\delta^2$) the probability that the $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$, simultaneously lie in the respective MCC confidence intervals is at least $1 - \alpha$ (which the user prespecifies).

The problem of MCC has been previously studied only for the case of i.i.d. normals. For example, see Tamhane (1977), Dudewicz and Ramberg (1972), Dudewicz, Ramberg, and Chen (1975), and Dudewicz and Dalal (1983). For a review of these and other procedures, the reader is referred to Chapter 7 of Hochberg and Tamhane (1987) and Goldsman and Nelson (1994). None of these papers covers relative-width confidence intervals, as we do here.

The rest of the paper has the following organization. In Section 2 we define the notation used and state an assumption on the processes being simulated. The procedures are presented in Section 3, while Section 4 contains a brief discussion on how to specify values for the parameters needed to run our procedures.

2 NOTATION AND ASSUMPTIONS

For system i $(i = 1, 2, \ldots, k)$, let $\mathbf{Y}_i = \{Y_i(t) : t \geq 0\} \in D[0, \infty)$ be a real-valued stochastic process representing the sample path of system i, where $D[0, \infty)$ is the space of right continuous functions on $[0, \infty)$ having left limits (for more details on $D[0, \infty)$, see Ethier and Kurtz 1986 or Glynn 1990). A large number of stochastic systems encountered in engi-

neering can be modeled as such. It is assumed here that $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k$ are mutually independent. (In a simulation, this means that for all i and j with $j \neq i$, the sample paths of systems i and j are generated using non-overlapping streams of uniform random numbers.) Also, let $\mathbf{Y} = (\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_k)$ and $Y(t) = (Y_1(t), Y_2(t), \ldots, Y_k(t))$.

To establish the results, we need to assume that the process $\mathbf{Y}$ satisfies a functional central limit theorem (FCLT). More specifically, let "$\Rightarrow$" denote weak convergence (see Billingsley 1968), and then assume the following:

A1 *There exist a nonsingular $k \times k$ matrix $\Sigma = (\sigma_{i,j} : i, j = 1, \ldots, k)$ with $0 < \sigma_{i,i} < \infty$, $i = 1, 2, \ldots, k$, and a constant $\mu = (\mu_1, \mu_2, \ldots, \mu_k) \in \Re^k$ such that*

$$X_\delta \Rightarrow \Sigma B$$

as $\delta \to 0$, where B is a k-dimensional standard Brownian motion, $X_\delta = (X_{1,\delta}, X_{2,\delta}, \ldots, X_{k,\delta})$, and

$$X_{i,\delta}(t) = \frac{1}{\delta}\left(\frac{\int_0^{t/\delta^2} Y_i(s)ds}{1/\delta^2} - \mu_i t\right), \quad t \geq 0,$$

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

Here, Σ is a diagonal matrix with ith diagonal entry $\sigma_i^2 \equiv \sigma_{i,i}$. For each $i = 1, 2, \ldots, k$, the constant μ_i appearing in A1 is precisely the steady-state mean of the process $\mathbf{Y}_i$. Also, σ_i^2 is the asymptotic variance of $\mathbf{Y}_i$.

A large number of stochastic systems that admit a steady state will also satisfy Assumption A1. For example, this assumption holds if the process $\mathbf{Y}$ satisfies any of the following: $\mathbf{Y}$ is regenerative and satisfies suitable moment conditions (see Glynn and Whitt 1987); $\mathbf{Y}$ is a martingale process (see Chapter 7 of Ethier and Kurtz 1986); $\mathbf{Y}$ satisfies appropriate mixing conditions (see Chapter 7 of Ethier and Kurtz 1986); or the $\mathbf{Y}(t)$ are associated (see Newman and Wright 1981). All these processes are weakly dependent, in the sense that events far apart in time are almost independent. Processes with long-range dependency may not satisfy the assumption.

3 THE MCC PROCEDURES

The basic idea of our two-stage MCC procedures is as follows. In the first stage we run independent simulations of the different systems. The method of batch means with $m \geq 2$ batches is applied to the output of each system, thus yielding an estimate of the variance of the first-stage sample mean. This is used to compute the total number of batches needed for each system, and so, additional simulation may need to be run in the second stage. (For each system, the total number of batches times the length of a batch will of course be the total simulation time that must be allocated to that system.) Statistics are then updated and the simultaneous MCC confidence intervals constructed. The two-stage batch means algorithm for constructing lower one-sided, absolute-half-width MCC confidence intervals is as follows:

Procedure 1

1. Specify the desired absolute half-width δ of each MCC confidence interval (where δ is small); the desired confidence level $1-\alpha$; and the number of initial batches $m \geq 2$. Let $\gamma \equiv \gamma(m, k, \alpha)$ solve Rinott's (1978) integral. (Wilcox 1984 presents tables for γ. Note that our notation differs from that used in Wilcox's tables. In particular, our m and γ correspond to his n_0 and h, respectively.)

2. First stage: Independently simulate systems $i = 1, 2, \ldots, k$, with run lengths $T_i = T_i(\delta)$, each proportional to $1/\delta^2$. For each system i, group the output into m (non-overlapping) batches, each of size T_i/m, and compute

$$Z_{i,j} = \frac{1}{T_i/m}\int_{(j-1)T_i/m}^{jT_i/m} Y_i(s)\, ds, \quad j \geq 1,$$

which is the sample mean of the jth batch.

3. For each system $i = 1, 2, \ldots, k$, compute the sample variance of the m batch means from the first stage as

$$S_i^2 = \frac{1}{m-1}\sum_{j=1}^{m}\left(Z_{i,j} - \frac{1}{m}\sum_{k=1}^{m} Z_{i,k}\right)^2.$$

4. For each system $i = 1, 2, \ldots, k$, compute the total number of batches to collect as

$$N_{a,i}(\delta) = \max\left\{m, \left\lceil\left(\frac{\gamma S_i}{\delta}\right)^2\right\rceil\right\},$$

where $\lceil x \rceil$ represents the smallest integer greater or equal to x.

5. Second stage: For each system i, if $N_{a,i}(\delta) \geq m+1$, simulate independently that system to collect the additional (non-overlapping) batches $m+1, \ldots, N_{a,i}(\delta)$, each of size T_i/m. Compute then the batch means $Z_{i,m+1}, \ldots, Z_{i,N_{a,i}(\delta)}$.

6. For each system $i = 1, 2, \ldots, k$, compute the overall sample mean as

$$\hat{\mu}_{a,i} = \frac{1}{N_{a,i}(\delta)} \sum_{j=1}^{N_{a,i}(\delta)} Z_{i,j}.$$

7. Simultaneously construct the absolute-precision, lower one-sided MCC confidence intervals

$$I_{a,i}(\delta) = [\hat{\mu}_{a,i} - \hat{\mu}_{a,k} - \delta, \ +\infty)$$

for $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$.

Damerdji and Nakayama (1996) establish the following result.

Theorem 1 *Assume that Assumption A1 holds. If Procedure 1 is used, then*

$$\lim_{\delta \to 0} P\{\mu_i - \mu_k \in I_{a,i}, \ i = 1, 2, \ldots, k-1\} > 1 - \alpha.$$

Theorem 1 establishes that our MCC confidence intervals are asymptotically valid (in the sense that the simultaneous coverage probability is at least $1 - \alpha$).

Procedure 1 above is for the case when we want MCC confidence intervals having a pre-specified absolute half-width. However, in certain settings, we may desire the interval for $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$, to have half-width that is, say 5%, of $\mu_i - \mu_k$. Thus, we now develop a two-stage procedure for relative-width, lower one-sided, MCC confidence intervals.

Procedure 2

1. Specify the desired relative half-width of each MCC confidence interval δ (where δ is small); the desired confidence level $1 - \alpha$; and the number of initial batches $m \geq 2$. Let $\gamma \equiv \gamma(m, k, \alpha)$ solve Rinott's (1978) integral.

2. In the first stage independently simulate systems $i = 1, 2, \ldots, k$, with run lengths $T_i = T_i(\delta)$, each proportional to $1/\delta^2$. For each system i, group the output into m (non-overlapping) batches, each of size T_i/m, and compute the first m batch means $Z_{i,1}, Z_{i,2}, \ldots, Z_{i,m}$.

3. For each system $i = 1, 2, \ldots, k$, compute the sample mean and the sample variance of the m batch means from the first stage as

$$\tilde{\mu}_i = \frac{1}{m} \sum_{j=1}^{m} Z_{i,k}$$

and

$$S_i^2 = \frac{1}{m-1} \sum_{j=1}^{m} (Z_{i,j} - \tilde{\mu}_i)^2,$$

respectively.

4. For each system $i = 1, 2, \ldots, k$, compute the total number of batches to collect as

$$N_{r,i}(\delta) = \max\left\{m, \left\lceil \left(\frac{\gamma S_i}{\delta \tilde{\mu}_k}\right)^2 \right\rceil\right\},$$

assuming $\tilde{\mu}_k \neq 0$.

5. For the second stage, independently simulate systems $i = 1, 2, \ldots, k$, if needed, to collect additional (non-overlapping) batches $m+1, m+2, \ldots, N_{r,i}(\delta)$, each of size T_i/m. Compute the batch means $Z_{i,m+1}, \ldots, Z_{i,N_{r,i}(\delta)}$.

6. For each system $i = 1, 2, \ldots, k$, compute the overall sample mean as

$$\hat{\mu}_{r,i} = \frac{1}{N_{r,i}(\delta)} \sum_{j=1}^{N_{r,i}(\delta)} Z_{i,j}.$$

7. Simultaneously construct the relative-precision, lower one-sided MCC confidence intervals

$$I_{r,i}(\delta) = [\hat{\mu}_{r,i} - \hat{\mu}_{r,k} - \delta|\hat{\mu}_{r,k}|, \ +\infty)$$

for $\mu_i - \mu_k$, $i = 1, 2, \ldots, k-1$.

The relative half-widths of the above intervals are relative not with respect to $|\hat{\mu}_{r,i} - \hat{\mu}_{r,k}|$ but rather $|\hat{\mu}_{r,k}|$. Damerdji and Nakayama (1996) establish the validity of the following theorem:

Theorem 2 *Assume that Assumption A1 holds and that $\mu_k \neq 0$. If Procedure 2 is used, then*

$$\lim_{\delta \to 0} P\{\mu_i - \mu_k \in I_{r,i}(\delta), \ i = 1, 2, \ldots, k-1\} > 1 - \alpha.$$

4 SPECIFYING VALUES FOR PARAMETERS

To use Procedures 1 and 2 in practice, the practitioner must specify values for several parameters. These include the desired (absolute or relative) half-width δ, the run length of the first stage T_i (which must be proportional to $1/\delta^2$) for each system, and the number of initial batches m.

Because of the similarity of Procedures 1 and 2 and the two-stage stopping procedures developed by Nakayama (1994) for the single-system setting, it is reasonable to assume that appropriate values for the parameters of Nakayama's (1994) algorithm are also valid for our new procedures. Nakayama (1994) suggests that one should choose $5 \leq m \leq 15$ and $\delta < 0.025$. However, as Nakayama (1994) notes, selecting a reasonable value for T_i given δ is a delicate matter. In the case when simulating queueing systems, though, Nakayama (1994) proposes using some of the results of Whitt (1989a,1989b); for more details, see Sections 5 and 6 of Nakayama (1994).

ACKNOWLEDGMENTS

The work of the second author was supported in part by NJIT Grant 421180.

REFERENCES

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

Damerdji, H. and M. K. Nakayama. 1996. Two-Stage Procedures for Multiple Comparisons with the Best in Steady-State Simulations. In preparation.

Dudewicz, E. J. and S. R. Dalal. 1983. Multiple comparisons with a control when variances are unknown and unequal. *American Journal of Mathematics and Management Sciences* 4:275–295.

Dudewicz, E. J. and J. S. Ramberg. 1972. Multiple comparison with a control: Unknown variances. *Ann. Tech. Conf. Amer. Soc. Quality Control* 26:483–488.

Dudewicz, E. J., J. S. Ramberg, and H. J. Chen. 1972. New tables for multiple comparisons with a control (unknown variances). *Biometrische Zeitschrift* 17:437–445.

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

Goldsman, D. and B. L. Nelson. 1994. Ranking, selection and multiple comparisons in computer simulation. In *Proceedings of the 1994 Winter Simulation Conference*, ed. J. D. Tew, S. Manivannan, D. A. Sadowski, and A. F. Seila, 192–199. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

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

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

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

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

Tamhane, Y. 1977. Multiple comparisons in model I: One-way ANOVA with unequal variances. *Communications in Statistics—Theory and Methods* A6:15–32.

Whitt, W. 1989a. Planning queueing simulations. *Management Science* 35:1341–1366.

Whitt, W. 1989b. Simulation run length planning. In *Proceedings of the 1989 Winter Simulation Conference*, ed. E. A. MacNair, K. J. Musselman, and P. Heidelberger, 106–112. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Wilcox, R. R. 1984. A table for Rinott's selection procedure. *Journal of Quality Technology* 16:97–100.

AUTHOR BIOGRAPHIES

HALIM DAMERDJI is an assistant professor in the Department of Industrial Engineering at North Carolina State University. His research interests are in simulation and applied stochastic processes. He is a member of INFORMS.

MARVIN K. NAKAYAMA is an assistant professor in the Department of Computer and Information Science at the New Jersey Institute of Technology. Previously, he has held positions in the Management Science/Computer Information Systems Department of the Graduate School of Management at Rutgers University, and at the IBM Thomas J. Watson Research Center in Yorktown Heights, New York. He received Ph.D. in Operations Research from Stanford University. His research interests include applied probability, simulation output analysis, gradient estimation, rare-event simulation, and random-variate generation.