Proceedings of the 2003 Winter Simulation Conference
S. Chick, P. J. Sánchez, D. Ferrin, and D. J. Morrice, eds.

AN INDIFFERENCE-ZONE SELECTION PROCEDURE WITH MINIMUM SWITCHING AND SEQUENTIAL SAMPLING

L. Jeff Hong
Barry L. Nelson

Department of Industrial Engineering and Management Sciences
Northwestern University
Evanston, IL 60208-3119, U.S.A.

ABSTRACT

Statistical ranking and selection (R&S) is a collection of experiment design and analysis techniques for selecting the "population" with the largest or smallest mean performance from among a finite set of alternatives. R&S procedures have received considerable research attention in the stochastic simulation community, and they have been incorporated in commercial simulation software. One of the ways that R&S procedures are evaluated and compared is via the expected number of samples (often replications) that must be generated to reach a decision. In this paper we argue that sampling cost alone does not adequately characterize the efficiency of ranking-and-selection procedures, and we introduce a new sequential procedure that provides the same statistical guarantees as existing procedures while reducing the expected total cost of application.

1 INTRODUCTION

Ranking-and-selection procedures (R&S) based on the indifference-zone formulation have been proposed to select the simulated system with the largest or smallest mean performance from among a finite number of alternative systems (see Bechhofer et al. 1995 for a summary). Among these, fully sequential procedures, which approximate the sum of differences between two systems as a Brownian motion process and use a triangular continuation region to determine the stopping time of the selection process, were first proposed by Paulson (1964). Figure 1 illustrates how triangular continuation region works. Hartmann (1988, 1991) improved Paulson's procedure by replacing Boole's inequality with a geometric inequality and replacing a large-deviation bound by a Brownian motion bound. These procedures were intended for the case of normally distributed data with known or unknown common variance across systems. Recently, Kim and Nelson (2001, 2002) further extended Hartmann's work to allow unknown and unequal variances, the use of common random numbers and single-replication experiment designs, yielding procedures that are more applicable in computer simulation experiments. In this paper we also refine sequential selection procedures with a triangular continuation region by considering the computation costs that are incurred in simulation experiments.

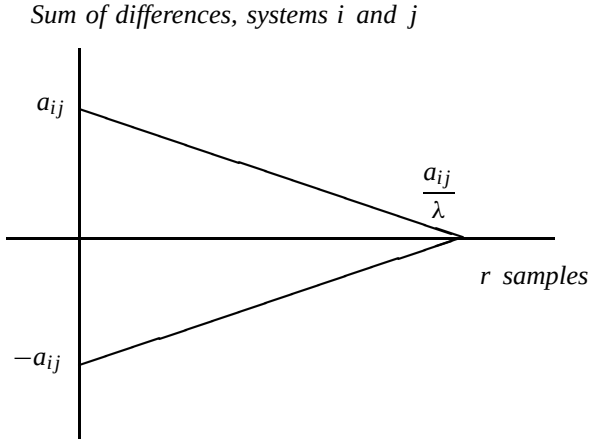

Figure 1: Triangular Continuation Region for the Fully Sequential, Indifference-zone Procedures; Selection of System i or j as Best Depends on Whether the Sum of Differences Exits the Region from the Top or Bottom

One important, and relatively recent, application of R&S is within optimization-via-simulation algorithms (see, for instance, Boesel, Nelson and Ishii 2003, and Pichitlamken and Nelson 2003). Many optimization algorithms attempt to move from a current good solution to an improved solution on each iteration by selecting the best from among a small set of candidates or neighbors. Ranking-and-selection procedures can be embedded within these algorithms to help recognize improved solutions efficiently, and with a statistical guarantee of correctness. Since some optimization algorithms revisit solutions to insure conver-

gence, Pichitlamken and Nelson (2002) developed fully sequential selection procedures that exploit the data already obtained on previous visits, even if the sample sizes are unequal. The procedure we introduce in this paper also extends easily to unequal initial samples, and it is particularly well suited to the optimization setting.

R&S procedures are typically evaluated in terms of the expected number of samples required to reach termination. Sequential procedures with elimination—such as Kim and Nelson's (2001, 2002) procedure KN—and Bayesian procedures—such as Chick and Inoue (2001)—reduce the expected total number of samples relative to well-known two-stage indifference-zone selection procedures, such as Rinott's (1978) procedure (RN). Two-stage procedures were designed to provide their guarantees in the face of a "least-favorable configuration," and not to try to adapt to more favorable situations (two-stage procedures with elimination at the first stage have been discussed in, for instance, Nelson et al. 2001, and these can be more efficient than procedures like RN).

The focus in R&S research has been on developing procedures that reduce the cost of *sampling*, which means generating data (typically replications) from a simulation model. To achieve this reduction, fully sequential procedures like KN repeatedly switch among the different simulation models, where a *switch* occurs when we change the model instance from which samples are generated. In fact, KN requires one switch for each new sample generated after an initial stage of sampling.

Unfortunately, the computational overhead of switching can be significant, sometimes orders of magnitude more than sampling. Further, the cost of switching may be incurred thousands of times if the R&S procedure is embedded within an optimzation algorithm that explores a large solution space. The work required to switch from one simulated system to another usually includes storing state information about the current system (including values of the random number seeds); saving all relevant output data; swapping the executible code for the current system out of, and the code for the next system into, active memory; and restoring the state information for the next system to the values it had on the last call. Therefore, focusing solely on sampling cost is misleading in many applications.

Two-stage indifference-zone selection procedures, such as RN, mininize the number of switches; however, they do not adapt to the observed differences among the systems as the sampling progresses. Thus, they are efficient from a switching perspective, but inefficient from a sampling perspective. KN, and similar procedures, are efficient from a sampling perspective, but may not be efficient when switching is considered. In this paper we propose a two-stage sequential procedure that is adaptive, like KN, but also has the minimum number of switches, like RN. Further, our procedure tends to allocate more samples to the better systems, which is desirable when R&S is also used to add statistical inference after an optimization algorithm has terminated (see Boesel, Nelson and Kim 2003).

In Section 2 we present the Minimum Switching Sequential Procedure (MSS), which consumes the minimal number of switches, just as the two-stage procedures do, while still maintaining sequential sampling. Then we provide some numerical examples in Section 3, followed by the conclusions and future research directions in Section 4.

2 A MINIMUM SWITCHING PROCEDURE

In this section we describe a sequential procedure that guarantees, with confidence level greater than or equal to $1-\alpha$, that the system ultimately selected has the largest true mean if the true mean of the best is at least δ better than the second best. When there are inferior systems whose means are within δ of the best, then the procedure guarantees to find one of these *good* systems with the same probability. The parameter δ, which defines the *indifference zone*, is set by the experimenter to the minimum difference in expected performance that it is important to detect. Differences of less than δ are considered practically insignificant.

The procedure, which is called the Minimum Switching Sequential Procedure (MSS), has two stages, like RN. A *stage* is a checkpoint at which the maximum number of samples that can be taken from each system until the next checkpoint is determined; it is the "maximum" number because fewer samples may be needed if some systems are eliminated from consideration. All of the procedures considered in this paper, including RN and KN, assume unknown output variances, and for that reason they require an initial stage of sampling that is not adaptive and whose size is somewhat arbitrary; we refer to this as the "zeroth stage." The first decision about how to proceed occurs after the zeroth stage.

MSS works as follows: After obtaining $n_0 \geq 2$ samples from each system in the zeroth stage, it then estimates the parameters of the triangular continuation region and checks if any system can be eliminated immediately. Let I be the set of systems still in play at the end of the zeroth stage, and let B and S *always* denote the systems in I with the best and the second-best zeroth-stage sample means. In stage 1, MSS takes the maximum number of samples implied by the continuation region from system B, so that no more samples are needed for system B under any circumstances (this contrasts with RN which takes the maximum number of samples from *all* systems, and KN which obtains one sample at a time from all systems still in play). Then MSS obtains one sample at a time from system S, comparing a weighted sample mean from system S to a weighted sample mean from system B, with elimination decisions after each sample. If B eliminates S, then the identity of S is updated (since the former system S is no longer in I, and S is always

the system in I with the second-largest zeroth-stage sample mean) and the process starts over again. If S eliminates B, then the identities of both B and S are updated, the new system B gets the maximum number of samples, and the process continues. The procedure always examines only two systems at a time, eliminating one, and stopping when there is only one system remaining in I. MSS is sequential in the way it obtains samples, but it requires at most k switches after the zeroth stage, where k is the number of systems.

Throughout this paper we use $\mathbf{x}_i$ to denote the ith system and use $X_{i\ell}$ to denote the ℓth independent sample from system i. Think of $\mathbf{x}_i$ as the vector of decision variables that define the ith system. We assume that $X_{i\ell} \sim \mathrm{N}\left(\mu_i, \sigma_i^2\right)$, with both μ_i and σ_i^2 unknown. The procedure has two stages, denoted $s = 0, 1$. Let $\bar{\bar{X}}_i(m) = m^{-1}\sum_{\ell=1}^{m} X_{i\ell}$ denote the sample mean of the first m samples from system i, and let $\bar{X}_i(n; s)$ denote the sample mean of the first n samples from system i *taken in stage* s, where $i = 1, 2, \ldots, k$.

Minimum Switching Sequential Procedure

Setup: Select confidence level $1/k < 1-\alpha < 1$, indifference-zone parameter $\delta > 0$ and zeroth-stage sample size $n_0 \geq 2$. Select λ such that $0 < \lambda < \delta$ ($\lambda = \delta/4$ is recommended by Paulson[1964]).

Initialization: Take n_0 samples $X_{i\ell}$, $\ell = 1, 2, \ldots, n_0$, from each system $i = 1, 2, \ldots, k$. For all $i \neq j$, calculate

$$S_{ij}^2 = \frac{1}{n_0 - 1}\sum_{\ell=1}^{n_0}\left(X_{i\ell} - X_{j\ell} - \left[\bar{X}_i\left(n_0; 0\right) - \bar{X}_j\left(n_0; 0\right)\right]\right)^2, \tag{1}$$

the sample variance of the difference between systems i and j, and let

$$a_{ij} = \frac{(n_0 - 1)\, S_{ij}^2}{4(\delta - \lambda)}\left\{\left[1 - (1-\alpha)^{1/(k-1)}\right]^{-2/(n_0-1)} - 1\right\}. \tag{2}$$

The parameters a_{ij} and λ define the triangular continuation region for system i and j. Let

$$N_{ij} = \left\lceil \frac{a_{ij}}{\lambda} \right\rceil - n_0,$$

where $\lceil a \rceil$ denotes the smallest integer which is greater than or equal to a.

Initial Screening: Calculate

$$Z_{ij}(n_0) = n_0\left[\bar{X}_i(n_0; 0) - \bar{X}_j(n_0; 0)\right].$$

Let

$$I = \left\{i : Z_{i,j}(n_0) \geq \min\left\{0, -a_{ij} + n_0\lambda\right\},\ i, j = 1, 2, \ldots, k,\ i \neq j\right\}$$

be the set of systems still in play. If $|I| = 1$, then stop and select the system whose index is in I as the best. Otherwise, sort the elements in I based on the zeroth-stage sample means $\bar{X}_i(n_0; 0)$ and let B and S *always* be the systems in I with the best and the second best zeroth-stage sample means. Let

$$N_B = \max_{j \in I, j \neq B} N_{Bj}.$$

Take N_B additional samples from system B. Let r be the number of samples taken from system S in stage 1, and set $r = 0$. Let ψ be an elimination indicator. If $\psi = 0$ then system B is still in contention; otherwise, system B has been eliminated. Set $\psi = 0$.

Screening: Take one sample from system S, and let $r = r + 1$. Calculate

$$Z_{BS}(n_0 + r) = Z_{BS}(n_0) + r\left[\bar{X}_B(N_B; 1) - \bar{X}_S(r; 1)\right], \tag{3}$$

$$W_{BS} = \max\left\{0, a_{BS} - \lambda(n_0 + r)\right\}.$$

If $Z_{BS}(n_0 + r) \geq W_{BS}$, then let $I = I - \{S\}$ and update S; if $Z_{BS}(n_0 + r) \leq -W_{BS}$, then let $I = I - \{B\}$, update B and S, and let $\psi = 1$; otherwise, go to **Screening**.

Stopping Rule: If $|I| = 1$, then stop and select the system whose index is in I as the best. Otherwise, if $\psi = 1$, then let

$$N_B = \max_{j \in I, j \neq B} N_{Bj},$$

and take $\max\{0, N_B - r\}$ samples from system B. Always let $r = 0$ and $\psi = 0$, then go to **Screening**.

Remark. $Z_{BS}(n_0 + r)$ is defined by Equation (3) instead of the more natural cumulative difference

$$Z_{BS}(n_0 + r) = \sum_{\ell=1}^{n_0+r}\left(\bar{\bar{X}}_B(n_0 + N_B) - X_{S\ell}\right). \tag{4}$$

Jennision et al. (1982) have shown that when $k > 2$ the sampling rule can only depend on the difference between sample means, not on the individual sample means, if Equation (4) is used. However, the individual sample means are required for sorting in MSS. Therefore, we use Equation (3) to insure independence between stages 0 and 1.

We show that our procedure satisfies the probability gurantee in Theorem 1, and a good system is selected when there are inferior systems within δ of the best system in Corollary 1. The proofs are given in Hong and Nelson (2003). Assume $\mu_k \geq \mu_{k-1} \geq \cdots \geq \mu_1$ so that system k is the best.

Theorem 1 *Suppose that $X_{i\ell}$, $\ell = 1, 2, \ldots$, are i.i.d. normally distributed, and that X_{ip} and X_{jq} are independent for $i \neq j$. Then MSS selects $\mathbf{x}_k$ with probability at least $1-\alpha$ whenever $\mu_k - \mu_i \geq \delta$ for $i = 1, 2, \ldots, k-1$.*

Corollary 1 *Suppose $\mu_k - \mu_{k-1} < \delta$. Then MSS selects a solution whose mean is within δ of μ_k with probability at least $1-\alpha$.*

Remark. We proved the use of a_{ij} defined in Equation (2) by using Paulson's probability bound. Pichitlamken and Nelson (2002) show that empirical results also suport the use of Fabian's probability bound used in KN, although it has not yet been proved. Fabian's probability bound is a Brownian motion bound. It is tighter than Paulson's probability bound which is a large-deviation bound. If Fabian's probability bound is used, we have

$$a_{ij} = \frac{(n_0 - 1)\, S_{ij}^2}{4(\delta - \lambda)} \left\{ \left[2 - 2(1-\alpha)^{1/(k-1)} \right]^{-2/(n_0-1)} - 1 \right\}, \tag{5}$$

for $\lambda = \delta/2$.

MSS has the same number of switches as two-stage procedures, such as RN, when both take additional samples for all systems after stage 0. However, MSS can eliminate clearly inferior systems at the end of stage 0, while RN only ceases sampling if n_0 happens to be larger than the total number of samples that RN required. Not surprisingly, our experiments show that the average sample size of MSS is less than or equal to RN in all situations. When the true difference between some system and the best system is larger than the indifference amount δ, which is the usual case in practice, MSS has a much smaller average sample size than RN. Since two-stage procedures have been implemented in commercial simulation software, such as Automod and Awesim, MSS is a good substitute for them.

In fully sequential procedures, such as KN, the elimination decisions between the better and the inferior systems occur with each having the same number of samples because the sampling is synchronized. In MSS, the inferior systems are typically eliminated by one of the (apparently) better systems that has received *all* of the samples that MSS will ever allocate to it. In some cases this fact will allow MSS to be more efficient than KN, even if switching cost is not significant. Consider, for instance, the case that among $k > 2$ systems the true mean difference between the best and second-best systems is small (less than or equal to δ), while the other systems are clearly inferior to these two. In this case, it is very likely that both MSS and KN will end up taking nearly the maximum samples from the two good systems. However, MSS will tend to eliminate the clearly inferior systems earlier (fewer samples), because the inferior systems are compared to a good system that has received all of its samples. This is because it is likely that one of the two good systems will have the largest zeroth-stage sample mean, and therefore will immediately be allocated all of its samples.

3 EMPIRICAL EVALUATION

In this section we summarize the results of an extensive empirical evaluation of MSS relative to the Rinott's procedure (RN) and Kim and Nelson's procedure (KN), which are representatives of two-stage procedures and fully sequential procedures, respectively. The systems are represented by various configurations of k normal distributions and, to assess the impact of nonnormality, lognormal distributions whose skewness and kurtosis (standardized third and fourth moments) differed from those of the normal distribution. In all cases system k is the true best (has the largest true mean). We evaluated each procedure on different variations of the systems, examing factors including the number of systems, the configuration of means, the configuration of variances and the relative cost of sampling and switching. The configurations, the experiment design and the results are described below.

3.1 Configurations and Experiment Design

Two configurations of the true means were used: the slippage configurations (SC) and the monotone increasing means (MIM). In SC μ_k was set to δ while $\mu_1 = \mu_2 = \cdots = \mu_{k-1} = 0$. This is a difficult configuration since all inferior systems are exactly δ from the best. In MIM $\mu_i = (i-1)\delta, i = 1, 2, \ldots, k$. MIM is used to investigate the effectiveness of the procedures in more favorable settings.

For each configuration of the means, we examined the effect of variances. There are three configurations of variances: equal variances (EV), increasing variances (IV) and decreasing variances (DV). In EV, $\sigma_1 = \sigma_2 = \cdots = \sigma_k = 1$; in IV, $\sigma_i = i$; and in DV $\sigma_i = k + 1 - i$.

Let c denote the switching cost in units of sampling cost (the cost of generating a sample). We used $c = 10$, $c = 100$ and $c = 1000$ to represent different cost structures. Considering the current status of computer simulation software, even $c = 1000$ is not excessive.

We also varied the number of systems in each experiment, with k taking values of 2, 5 and 10. In all experiments, the nominal probability of correct selection (PCS) was $1 - \alpha = 0.95$ and the first-stage sample size n_0 was 10. The indifference-zone parameter was set to $\delta = \sigma_k/\sqrt{n_0}$, where σ_k is the standard deviation of the best system. Thus, δ is one standard deviation of the first stage sample mean of the best system. For each configuration, 1000 macroreplications (complete repetitions) of the each procedure are performed to compare the performance measures, including the observed PCS, average number of

samples (ANSa), average number of switches (ANSw) and average total cost (ATC).

3.2 Summary of Results

MSS is very close to RN in the slippage configuration, but has much better performance than RN in the MIM configuration, which is more typical in practice. Therefore, MSS can be a good subsitute for the widely used RN procedure. We also found that the variance configuration does not significantly change the relative performances of all three procedures.

3.3 Some Specific Results

We do not attempt to present comprehensive results from such a large simulation study. Instead, we present selected results that emphasize the key conclusions.

3.3.1 Validity Check

In Sections 2 we showed the validity of MSS using Paulson's probability bound. However, we also pointed out that the empirical results conducted by Pichitlamken and Nelson (2002) support the use of the Fabian's Probability bound which is tighter than Paulson's bound. We checked the use of both bounds under the slippage configuration with equal variances. The results in Table 1 show that Paulson's probability bound is very conservative and the Fabian's probability bound is appropriate to use. In the rest of the paper, MSS refers to the MSS using Fabian's probability bound.

Table 1: Observed PCS of MSS using Paulson's and Fabian's Bounds

Procedure	$k = 2$	$k = 5$	$k = 10$
MSS Paulson	0.987	0.993	0.995
MSS Fabian	0.969	0.987	0.995

3.3.2 Effect of Mean Configuration

We compared the three procedures using the slippage configuration and monotone increasing means configuration. We used equal variances and $k = 10$ in both configurations. The results in Table 2 show that Rinott's procedure does not depend on the configuration of means while the other procedures require less work when the mean configuration deviates from the slippage configuration. In both configurations KN has the lowest ANSa, while MSS has the lowest ANSw. The reason why MSS has lower ANSw than RN is because some clearly inferior systems can be eliminated based on zeroth-stage samples in MSS. One can easily check that MSS has the lowest ATC in almost all cases under the three cost configurations ($c = 10, 100, 1000$).

Table 2: Effect of Mean Configuration

Procedure	SC		MIM	
	ANSa	ANSw	ANSa	ANSw
RN	1845.2	20.0	1845.2	20.0
KN	977.2	887.2	426.6	336.6
MSS	1950.2	19.9	981.7	18.5

3.3.3 Robustness Study

To assess the impact of nonnormal data on the procedures they were applied to lognormally distributed data with increasing levels of skewness and kurtosis, relative to the normal distribution (which has skewness 0 and kurtosis 3). The configurations of means and variances are same as Nelson et al. (2001). In all cases $n_0 = 10$, $c = 10$, $\sigma_i = 1$ for all i, $\delta = 1/\sqrt{n_0}$, $\mu_1 = \mu_2 = \cdots = \mu_{k-1} = 0$ and $\mu_k = \delta$. We also used different number of systems, $k = 2, 5, 10$. In the low deviation case the skewness and kurtosis are 1.8 and 9.1; in the medium deviation case they are 4.0 and 41.0; and in the high deviation case they are 6.2 and 113.2. Notice that the mean configuration is the slippage configuration which is a difficult configuration to achieve the nominal PCS.

Table 3 shows the estimated PCS for three lognormal cases for MSS, with the corresponding normal case included for comparison. When the level of deviation from the normal distribution increases, the observed PCS decreases. However, the observed PCS is close to the nominal value even when the deviation is significant.

Table 3: Observed PCS of MSS in the Robustness Study with Nominal Level 0.95

	Normal	Low	Medium	High
$k = 2$	0.969	0.970	0.960	0.941
$k = 5$	0.987	0.980	0.953	0.922
$k = 10$	0.995	0.979	0.950	0.926

Table 4: The Five Alternative Inventory Policies

Policy (i)	s	S	Expected Cost
1	20	40	114.176
2	20	80	112.742
3	40	60	130.550
4	40	100	130.699
5	60	100	147.382

3.3.4 An Illustrative Example

We provide a system simulation example to compare our procedures with RN and KN. Consider the (s, S) inventory system with the 5 inventory policies provided in Koenig and Law (1985). The objective of the study is to compare the 5 policies given in Table 4 on the basis of their corresponding expected average cost per month for the first 30 months of operation and select the policy with the minimum expected cost. The expected cost, which can be analytically computed, are also given in Table 4. We set $\delta = 1$, $n_0 = 10$ and $c = 100$. Table 5 includes the results of the simulation study based on 1000 complete macroreplications. The results are consistent to what we observed in the previous experiments. Notice that δ is smaller than the true difference between the best and the second best systems, which explains the high level of the observed PCS.

Table 5: Simulation Study of the Illustrative Example

Procedure	PCS	ANSa	ANSw	ATC
RN	1.000	1033.1	10.0	2033.1
KN	0.998	235.7	190.7	19305.7
MSS	0.999	635.0	7.56	1391.0

4 CONCLUSION AND FUTURE WORK

In this paper we have presented an indifference-zone selection procedure which is sequential and has minimum number of switches. As we discussed in Section 3, MSS has similar performance (in both ANSa and ANSw) in the slippage configuration, but much better performance in more typical configurations, than the widely used Rinott's procedure. It can be a substitute for the Rinott's procedure.

The disadvantage MSS has is that at least one system always gets the maximum number of samples, which may be more than needed when true differences between systems are much larger than the indifference-zone parameter δ. In that case the wasted sampling may overcome the savings from not switching. We are currently working on a more flexible indifference-zone selection procedure (Hong and Nelson 2003). It incorporates the cost structure into the design of the procedure, therefore achieves minimum total cost.

REFERENCES

Bechhofer, R. E., T. J. Santner and D. M. Goldsman. 1995. *Design and Analysis of Experiments for Statistical Selection, Screening, and Multiple Comparisons*. Wiley, New York.

Boesel, J., B. L. Nelson and N. Ishii. 2003. A framework for simulation-optimization software. *IIE Transactions*, forthcoming.

Boesel, J., B. L. Nelson and S.-H. Kim. 2003. Using ranking and selection to 'clean up' after simulation optimization. *Operations Research*, forthcoming.

Chick, S. E. and K. Inoue. 2001. New two-stage and sequential procedures for selecting the best simulated system. *Operations Research*, 49:732-743.

Hartmann, M. 1988. An improvement on Paulson's sequential ranking procedure. *Sequential Analysis*, 7:363-372.

Hartmann, M. 1991. An improvement on Paulson's procedure for selecting the population with the largest mean from k normal populations with a common unknown variance. *Sequential Analysis*, 10:1-16.

Hong, L. J. and B. L. Nelson. 2003. The Tradeoff Between Sampling and Switching: New Sequential Procedures for Indifference-Zone Selection. Working Paper, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois.

Jennison, C., I. M. Johnston and B. W. Turnbull. 1982. Asymptotically optimal procedures for sequential adaptive selection of the best of several normal means. In *Statistical Decision Theory and Related Topics III*, ed. S. S. Gupta and J. Berger, 2:55-86. New York: Academic Press.

Kim, S.-H. and B. L. Nelson. 2001. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation*, 11:251-273.

Kim, S.-H. and B. L. Nelson. 2002. On the asymptotic validity of fully sequential selection procedures for steady-state simulation. Working Paper, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois.

Koenig, L. W., and A. M. Law. 1985. A procedure for selecting a subset of size m containing the l best of k independent normal populations, with applications to simulation. *Communications in Statistics: Simulation and Computation*, 14:719–734.

Nelson, B. L., J. Swann, D. Goldsman and W. Song. 2001. Simple procedures for selecting the best simulated system when the number of alternatives is large. *Operations Research*, 49:950-963.

Paulson, E. 1964. A sequential procedure for selecting the population with the largest mean from k normal populations. *Annals of Mathematical Statistics*, 35:174-180.

Pichitlamken, J. and B. L. Nelson. 2002. A selection-of-the-best procedure for optimization via simulation. Working Paper, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, Illinois.

Pichitlamken, J. and B. L. Nelson. 2003. A combined procedure for optimization via simulation. *ACM Transactions on Modeling and Computer Simulation*, forthcoming.

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

AUTHOR BIOGRAPHIES

L. JEFF HONG is a Ph.D. student in the Department of Industrial Engineering and Management Sciences at Northwestern University. He received his B.E. in industrial engineering and automotive engineering from Tsinghua University, China, and M.S. in applied mathematics from University of Cincinnati. His research interests include ranking and selection and optimization via simulation. His e-mail and web addresses are <hongl@iems.northwestern.edu> and <http://users.iems.northwestern.edu/~hongl>.

BARRY L. NELSON is the Krebs Professor of Industrial Engineering and Management Sciences at Northwestern University, and is Director of the Master of Engineering Management Program there. His research centers on the design and analysis of computer simulation experiments on models of stochastic systems. He has published numerous papers and two books. Nelson has served the profession as the Simulation Area Editor of *Operations Research* and President of the INFORMS (then TIMS) College on Simulation. He has held many positions for the Winter Simulation Conference, including Program Chair in 1997 and current membership on the Board of Directors. His e-mail and web addresses are <nelsonb@northwestern.edu> and <www.iems.northwestern.edu/~nelsonb/>.