Proceedings of the 1999 Winter Simulation Conference
P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, eds.

SELECTION PROCEDURES WITH STANDARDIZED TIME SERIES VARIANCE ESTIMATORS

David Goldsman
William S. Marshall

School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.

ABSTRACT

This article studies a modification of Rinott's two-stage procedure for selecting the normal population with the largest (or smallest) mean. The modification, which is appropriate for use in the simulation environment, uses in the procedure's first stage different variance estimators than the usual batch means (BM) variance estimator. In particular, we will use variance estimators arising from the method of standardized time series (STS). On the plus side, certain STS estimators have more degrees of freedom than that of the BM estimator. On the other hand, STS variance estimators tend to require larger sample sizes than the BM estimator in order to converge to their assumed distributions. These considerations result in trade-offs involving the procedure's achieved probability of correct selection as well as the procedure's expected sample size.

1 INTRODUCTION

Statistical selection procedures are often used in computer simulations to compare alternative designs. These methods are applicable when we are interested in making comparisons among a finite, but not necessarily small, number of systems (say 2 to 1000). For example, such procedures could be appropriate in any of the following practical situations:

- A manufacturer would like to know which of ten plant layouts under consideration will maximize expected revenues.
- A network news division wants to determine the most popular candidate before an election.
- A medical research team conducts a clinical study comparing the success rates of three different drug regimens to combat a certain disease.

One class of selection procedures is designed to find the alternative with the largest (or smallest) mean; these procedures typically assume that the experimenter will have at his disposal independent and identically distributed (i.i.d.) normal observations from each competitor. Usually, these normal observations are simply the batch means from one long run of each competitor or the sample means from independent replications of the competing simulation processes — so these observations are probably not quite normal, nor, in the case of batch means, quite independent. This can be troublesome, since violations of the i.i.d. normal assumption can lead to improper conclusions on the part of the experimenter, e.g., selecting the wrong alternative as best, or asserting that the selected alternative is best with a confidence level that is too high.

For elementary tutorials on selection procedures, the reader should see Goldsman and Nelson (1998ab); some implementation issues are discussed in Goldsman, et al. (1999); and more-advanced treatments are given in Gibbons, Olkin, and Sobel (1977) and Bechhofer, Santner, and Goldsman (BSG) (1995). Law and Kelton (1991) describe a number of selection procedures that have proven useful in simulation applications.

This paper concerns the use of Rinott's (1978) two-stage selection procedure in the simulation environment. Perhaps the key to Rinott's procedure is that it uses its first stage to estimate the variance of the allegedly i.i.d. normal observations, almost always using the sample variance of the batch or replicate sample means; these variance estimates then determine how many additional observations to take in the second stage, after which a decision on the best alternative is finally made.

Our aim in the current article is to incorporate into the Rinott procedure's first stage different variance estimators than the usual sample variance (or so-called "batch means" (BM) variance estimator). Namely, we will use variance estimators arising from the method of standardized time series (STS). Certain STS estimators have more degrees of freedom than that of the BM estimator, making it easier to pin down the appropriate number of observations to take from

each alternative during the second stage. Unfortunately, as opposed to the BM variance estimator, these STS estimators sometimes require larger sample sizes within each batch or replication before the necessary asymptotic distribution theory works itself out. We will examine the trade-offs between these good and bad characteristics.

The remainder of this paper consists of the following. §2 presents background material that describes pertinent selection procedure notation and terminology. §3 gives a description of Rinott's procedure, using "generic" variance estimators in the first stage. In order to develop an arsenal of variance estimators for use in the first stage, we give in §4 a quick tutorial on BM and STS variance estimation in the simulation environment. §5 gives a comparison of the new methods, with advantages and disadvantages, while §6 offers some conclusions and recommendations for future research.

2 BACKGROUND

Selection procedures are used to select outright the best of a number of competing scenarios, where "best" refers here to the alternative having the largest (or smallest) expected value.

To facilitate what follows, we define some notation: Let Y_{ij} represent the jth simulation output from scenario i, for $i = 1, 2, \ldots, k$ scenarios and $j = 1, 2, \ldots$. For fixed i, the usual assumption is that the outputs from scenario i, $Y_{i1}, Y_{i2}, \ldots$, are i.i.d. normal. These assumptions are *roughly* plausible if $Y_{i1}, Y_{i2}, \ldots$ are sample means across independent replications, or if they are appropriately defined batch means from a single replication after accounting for initialization effects. We also assume that the (normal) observations among scenarios are independent, i.e., Y_{ij} is independent of $Y_{i',j}$ for all $i \neq i'$ and all j. Finally, let $\mu_i = \mathsf{E}[Y_{ij}]$ denote the expected value of a sample-mean output from the ith alternative simulation scenario, and let $\sigma_i^2 = \mathsf{Var}(Y_{ij})$ denote its variance. The Rinott (1978) method we describe herein makes comparisons based on estimates of μ_i.

Selection procedures allow the experimenter to specify a "practical-significant" difference, often denoted by δ and often referred to as the "indifference-zone" parameter. Any scenario whose performance is within δ of the best can be considered as a candidate for the best, perhaps "as good as" the best, for all practical purposes; alternatives that are not within δ of the best are to be considered as clearly inferior, and we would like to avoid selecting such poor candidates.

Our goal is to correctly select the true best scenario (or at least a "good" one within δ of the best). Of course, in a stochastic simulation such a correct selection (CS) can never be guaranteed with certainty. But the Rinott selection procedure guarantees to select the best alternative with user-specified high probability $1-\alpha$ whenever the true best is at least a user-specified amount δ better than the others, i.e., we have the following:

Probability Requirement: For constants (δ, α) with $0 < \delta < \infty$ and $1/k < 1 - \alpha < 1$, specified prior to the start of experimentation, we require

$$\Pr\{\text{CS}\} \;\geq\; 1-\alpha \;\text{ whenever }\; \mu_{[k]} - \mu_{[k-1]} \;\geq\; \delta,$$

where $\mu_{[1]} \leq \mu_{[2]} \leq \cdots \leq \mu_{[k]}$ are the ordered, but unknown μ_i's.

3 RINOTT'S PROCEDURE

We now describe a generalization of Rinott's (1978) procedure. In what follows, we let V_i denote a generic variance estimator for σ_i^2. In terms of the original procedure, V_i is simply the sample variance of the first-stage observations from alternative i, $i = 1, 2, \ldots, k$. Recall that an "observation" Y_{ij} is taken here to denote the jth batch mean or replicate sample mean from scenario i.

1. Specify the indifference-zone parameter δ, the desired probability of correct selection $1-\alpha$, and the common first-stage sample size $n_1 \geq 2$ to be taken from each alternative. Let h_α solve Rinott's integral, that is, the constant $h_\alpha = h(k, 1-\alpha, \nu)$ is the solution to

$$\int_0^\infty \int_0^\infty \Phi^{k-1}\left(\frac{h}{\sqrt{\nu(1/x + 1/y)}}\right) f_\nu(x) f_\nu(y)\, dy\, dx = 1-\alpha,$$

   where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $f_\nu(\cdot)$ is the probability density function of the χ^2-distribution with ν degrees of freedom, and ν depends on the variance estimator we use (see the appropriate tables in Wilcox 1984 or BSG 1995).
2. Take an i.i.d. sample $Y_{i1}, Y_{i2}, \ldots, Y_{in_1}$ from each of the k scenarios simulated independently.
3. Calculate the first-stage sample means $\bar{Y}_i^{(1)} = \sum_{j=1}^{n_1} Y_{ij}/n_1$, and marginal variance estimates V_i, for $i = 1, 2, \ldots, k$.
4. Compute the final sample sizes

$$N_i \;=\; \max\left\{n_1, \left\lceil h_\alpha^2 V_i/\delta^2 \right\rceil\right\}$$

   for $i = 1, 2, \ldots, k$, where $\lceil\cdot\rceil$ is the integer "round-up" function.
5. Take $N_i - n_1$ additional i.i.d. observations from scenario i, independently of the first-stage sample and the other scenarios, for $i = 1, 2, \ldots, k$.
6. Compute the overall sample means $\bar{Y}_i = \sum_{j=1}^{N_i} Y_{ij}/N_i$ for $i = 1, 2, \ldots, k$.
7. Select the scenario with the largest $\bar{Y}_i$ as best.

We make below a series of remarks concerning other aspects of the Rinott procedure.

Remark: If we had been interested in selecting the scenario with the *smallest* expected value, then the final step above obviously would instead be to select the scenario with the smallest $\bar{Y}_i$ as best.

Remark: In the "usual" implementation of the procedure (with batch means or replicate sample means), we would have taken as our variance estimators the sample variances of the batch or replicate means, each with $\nu = n_1 - 1$ degrees of freedom,

$$S_i^2 = \sum_{j=1}^{n_1} (Y_{ij} - \bar{Y}_i^{(1)})^2/(n_1 - 1),$$

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

Remark: Even though the current paper will concentrate only on the pure select-the-best problem, we mention a few alternative methodologies one could consider. As pointed out in Goldsman and Nelson (1998a), we could invoke a *screening* procedure to pare down a large number of alternatives into a palatable number; at that point, we might use a selection procedure to make the more fine-tuned choice of the best. Provided that certain assumptions are met, a screening procedure will choose a subset containing the best (or a good) scenario, and a selection procedure will then pick the best, with a user-specified confidence level. In fact, Nelson, et al. (1998) show how to combine a certain subset procedure with the Rinott procedure. This *two-phase* procedure is of great utility when the experimenter is initially faced with a large number of alternatives — the idea is for the subset procedure to pare out poor scenarios, after which Rinott selects the best from the survivors.

Remark: Multiple comparison procedures (MCPs) approach the problem of determining the best scenario by forming simultaneous confidence intervals on the means $\mu_i - \max_{j \neq i} \mu_j$ for $i = 1, 2, \ldots, k$. These confidence intervals are known specifically as multiple comparisons with the best (Hsu 1984), and they bound the difference between the expected performance of each alternative and the best of the others. See Hochberg and Tamhane (1987) for a thorough review. MCPs are often used in conjunction with selection procedures (Matejcik and Nelson 1995 and Nelson and Matejcik 1995) at no additional cost in Pr{CS} or sampling or calculation.

4 SOME VARIANCE ESTIMATORS

The question now arises as to what other variance estimators V_i are eligible to be used in Rinott's procedure? And which estimators do well when implemented?

A number of other estimators besides BM-based sample variances have been suggested for use in selection procedures. For example, Iglehart (1977) worked with a regenerative estimator on a selection procedure due to Dudewicz and Dalal (1975); Sullivan and Wilson (1984, 1989) tackled a similar problem using spectral variance estimation techniques. As far back as Goldsman (1985), the STS class of estimators was informally implemented in selection procedures.

We will discuss here only BM- and STS-based variance estimators, where the goal is to estimate the variance of the Y_{ij}'s, i.e., the batch or replicate sample means, from a particular simulation alternative.

4.1 Batch Means

Instead of working with the batch means directly, we can consider the lower-level observations that comprise the batch means. For a generic alternative, let $X_{i1}, X_{i2}, \ldots, X_{in}$ denote the stochastic output process of the simulation observations from the ith alternative. For example, $X_{i\ell}$ could be the ℓth individual waiting time in the ith queueing system under consideration. These observations, arising from some steady-state simulation, are rarely i.i.d. or normal, thus necessitating the need for a non-trivial variance estimator.

In order to explicitly relate the lower-level $X_{i\ell}$'s with the higher-level Y_{ij}'s (the batch or replicate means) we can divide $X_{i1}, X_{i2}, \ldots, X_{in}$ into b contiguous batches, each of length m (where we assume for convenience that $n = bm$); the observations $X_{i,(j-1)m+1}, X_{i,(j-1)m+2}, \ldots, X_{i,jm}$ comprise the jth batch, $j = 1, 2, \ldots, b$. The quantity

$$\bar{X}_{i,j,m} \equiv \frac{1}{m} \sum_{p=1}^{m} X_{i,(j-1)m+p}$$

is called the jth *batch mean* from scenario i — what we have been referring to as Y_{ij}, the jth high-level "observation" when we speak of a generic scenario i.

The point estimator that we shall always use for the mean μ_i of the ith system is the sample mean

$$\bar{X}_{i,n} \equiv \frac{1}{n} \sum_{p=1}^{n} X_{ip} = \frac{1}{b} \sum_{j=1}^{b} Y_{ij}.$$

In order to estimate the variance of the Y_{ij}'s, we could use their sample variance; this is what is commonly known as the batch means (BM) estimator for $\mathsf{Var}(Y_{ij})$, which we henceforth denote by V_B. The idea behind BM is that, for a fixed number of batches, a central limit theorem kicks in, so that the batch means, $Y_{i1}, Y_{i2}, \ldots, Y_{ib}$, are approximately i.i.d. normal

for large enough batch size m. In fact, under mild conditions, it is well known that

$$\begin{aligned} mV_B &\equiv \frac{m}{b-1}\sum_{j=1}^{b}(\bar{X}_{i,j,m} - \bar{X}_{i,n})^2 \\ &\xrightarrow{\mathcal{D}} \frac{v_i^2\chi^2(b-1)}{b-1}, \quad b > 1, \end{aligned}$$

where $v_i^2 \equiv \lim_{n\to\infty} n\mathsf{Var}(\bar{X}_{i,n})$ is the *variance parameter* of the stochastic process, and the symbol "$\xrightarrow{\mathcal{D}}$" denotes convergence in distribution as $m \to \infty$.

4.2 STS Estimators

We now look at a completely different methodology for estimating the variance of the Y_{ij}'s — standardized time series.

For $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, b$, and $\ell = 1, 2, \ldots, m$, the ℓth *cumulative mean* from batch j of scenario i is

$$\bar{X}_{i,j,\ell} \equiv \frac{1}{\ell}\sum_{p=1}^{\ell} X_{i,(j-1)m+p}.$$

For $i = 1, 2, \ldots, k$, $j = 1, 2, \ldots, b$, and $0 \le t \le 1$, the *standardized time series* from batch j of scenario i is given by

$$T_{i,j,m}(t) \equiv \frac{\lfloor mt \rfloor(\bar{X}_{i,j,m} - \bar{X}_{i,j,\lfloor mt \rfloor})}{v_i\sqrt{m}},$$

where $\lfloor\cdot\rfloor$ is the greatest integer function,. Schruben (1983) shows that if $X_{i1}, X_{i2}, \ldots, X_{in}$ is a stationary sequence satisfying certain mild moment and mixing conditions, then as $m \to \infty$, we have $T_{i,j,m}(t) \xrightarrow{\mathcal{D}} \mathcal{B}(t)$, $0 \le t \le 1$, a standard Brownian bridge process. All finite-dimensional joint distributions of $\mathcal{B}$ are normal and $\mathsf{Cov}(\mathcal{B}(s), \mathcal{B}(t)) = \min(s,t)(1-\max(s,t))$, $0 < s, t < 1$. Schruben also shows that $T_{i,j,m}(t)$ and $\bar{X}_{i,j,m}$ are asymptotically independent as the batch size m becomes large.

We denote the weighted area under the standardized time series formed by the jth batch of observations from scenario i, $i = 1, 2, \ldots, k$, and $j = 1, 2, \ldots, b$, by

$$A_{i,j} \equiv \frac{v_i}{m}\sum_{\ell=1}^{m} w(\ell/m)T_{i,j,m}(\ell/m),$$

where $w(\cdot)$ is a pre-specified weighting function that is continuous on [0,1], not dependent on m, and normalized so that

$$\begin{aligned} &\mathsf{Var}\left(\int_0^1 w(t)\mathcal{B}(t)\,dt\right) \\ &\quad = 2\int_0^1\int_0^u w(u)w(t)t(1-u)\,dt\,du = 1. \end{aligned}$$

(This expression can be simplified considerably; see Goldsman, Meketon, and Schruben 1990 for details.) Finally, define the location on $[0, 1]$ of the maximum of the standardized time series from the jth batch of observations from scenario i, $i = 1, 2, \ldots, k$, and $j = 1, 2, \ldots, b$, by

$$t_{i,j} \equiv \frac{\mathrm{argmax}_{1\le\ell\le m}\{T_{i,j,m}(\ell/m)\}}{m}.$$

In addition to the BM estimator V_B, we then have a collection of estimators for v_i^2 (cf. Glynn and Iglehart 1990, and Goldsman and Schruben 1990):

(Weighted) Area estimator:

$$\begin{aligned} mV_A &\equiv \frac{1}{b}\sum_{j=1}^{b} A_{i,j}^2 \\ &\xrightarrow{\mathcal{D}} \frac{v_i^2\chi^2(b)}{b}, \quad b \ge 1. \end{aligned}$$

Combined BM + Area estimator:

$$\begin{aligned} mV_{B+A} &\equiv \frac{m((b-1)V_B + bV_A)}{2b-1} \\ &\xrightarrow{\mathcal{D}} \frac{v_i^2\chi^2(2b-1)}{2b-1}, \quad b > 1. \end{aligned}$$

Maximum estimator:

$$\begin{aligned} mV_M &\equiv \frac{v_i^2}{3b}\sum_{j=1}^{b}\frac{T_{i,j,m}^2(t_{i,j})}{t_{i,j}(1-t_{i,j})} \\ &\xrightarrow{\mathcal{D}} \frac{v_i^2\chi^2(3b)}{3b}, \quad b \ge 1. \end{aligned}$$

Combined BM + Maximum estimator:

$$\begin{aligned} mV_{B+M} &\equiv \frac{m((b-1)V_B + 3bV_M)}{4b-1} \\ &\xrightarrow{\mathcal{D}} \frac{v_i^2\chi^2(4b-1)}{4b-1}, \quad b > 1. \end{aligned}$$

5 COMPARISONS

Notice that the variance estimators from the previous section are all asymptotically χ^2, and that we have conveniently arranged them in order of increasing degrees of freedom. It turns out that more degrees of freedom is inherently a good thing, but this boon is sometimes purchased at a price. In particular, the increased degrees of freedom ν yields

- A smaller variance of the variance estimator V_i, $i = 1, 2, \ldots, k$.
- A smaller Rinott h-value for a given number n_1 of first-stage batch or replicate means.
- Possible slower convergence of the variance estimator to its limiting χ^2 distribution.

The first two items are favorable; they (usually) result in a smaller expected total number of observations taken by the procedure. The third item is problematic, since the invalid distributional assumption may cause trouble involving lower-than-desired Pr{CS}.

5.1 Expected Number of Required Samples

It is possible to calculate the expected number of observations, say $\mathsf{E}[N_i]$, that the Rinott procedure will need from each scenario. This quantity is a function of the experimenter's specified parameter choices as well as the underlying variances of the competing scenarios. Without going into the details, we can consider a simple example. Suppose that the experimenter specifies a required Pr{CS} of $1 - \alpha = 0.95$, an indifference parameter of $\delta = 0.5$, and various choices of n_1 for the first-stage sample size (in terms of batches or replications). Further suppose that we are conducting the comparison among $k = 2$ scenarios, both of which have $\sigma_i^2 = 1$. In this example, we will use the BM estimator. Figure 1 displays $\mathsf{E}[N_i]$ (which will be the same for both of the $k = 2$ scenarios) as a function of n_1.

For small n_1, increasing the degrees of freedom, $\nu = n_1 - 1$, initially decreases the variance of the variance estimator, resulting in lower $\mathsf{E}[N]$ for the procedure. Eventually, however, we collect so many first-stage samples that we do not need to take any in the second stage — a wasteful situation resulting in the linear slope of $\mathsf{E}[N]$ for $n_1 > 30$ or so.

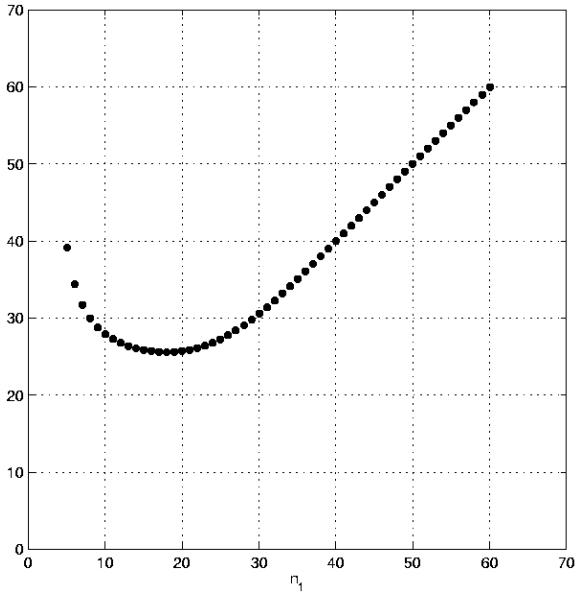

Figure 1: Relationship Between $\mathsf{E}[N]$ and n_1 for BM Variance Estimator

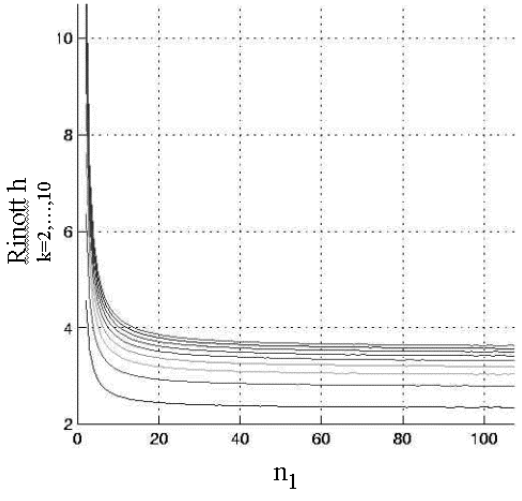

Figure 2: Relationship Between h and n_1 for BM Variance Estimator

5.2 Rinott h-value

Generally speaking, the value of h decreases with increased degrees of freedom — up to a point, where it begins to level off. This is illustrated in Figure 2, where we plot h as a function of the BM method's first-stage sample size n_1 for $1 - \alpha = 0.95$ and a selection of k-values; the number of scenarios is $k = 2, \ldots, 10$, with the largest value of k corresponding to largest values of h, and thus the top-most curve.

5.3 Convergence of Variance Estimators

A high degrees of freedom does not necessarily guarantee that one estimator is superior to another, especially if the higher degrees of freedom is purchased at the cost of a slower convergent rate to the limiting χ^2 distribution. In fact, it was shown in Sargent, Kang, and Goldsman (1992),

and it is well known, that some of the STS estimators converge to their limiting distributions more slowly than does the BM estimator; the maximum and combined BM + maximum estimators converge especially slowly. A consequence of this slow convergence is that the offending variance estimators often underestimate the true variances σ_i^2 of the batch means. This underestimation results in smaller-than-needed second-stage sample sizes, and then smaller-than-anticipated Pr{CS}.

6 CONCLUSIONS

On the plus side of things, we saw that certain STS estimators have more degrees of freedom than that of the BM estimator. On the other hand, STS variance estimators tend to require larger sample sizes than the BM estimator in order to converge to their assumed distributions. These considerations result in trade-offs involving the Rinott procedure's achieved probability of correct selection as well as the procedure's expected sample size, and are subjects of ongoing research.

The problem of dealing with the underlying "low-level" process observations (as opposed to assuming that the batch means are obligingly i.i.d. normal) has not been studied a great deal. Besides a couple of robustness papers in the literature, it seems that only Dudewicz and Zaino (1977) present a procedure to handle an explicit non-i.i.d. process (namely, a first-order autoregressive model). This too, is the subject of ongoing research on our part.

The reader may have noticed that all of the variance estimators studied in the current article were (asymptotically, at least) χ^2. Proving that the use of these estimators satisfies the Rinott probability requirement is not a problem since, as with BM, the STS estimators are not only χ^2, they are also (asymptotically) independent of the batch means. An interesting question to investigate is that of using other, non-χ^2 variance estimators in the Rinott procedure. For example, the low variance and reasonable convergence properties of the overlapping batch means estimator (Meketon and Schmeiser 1984) or the STS Cramér–von Mises estimator (Goldsman, Kang, and Seila 1999) might make them attractive candidates for inclusion in Rinott.

ACKNOWLEDGMENTS

This research was partially supported by National Science Foundation grant DMI-9622269. A small portion of this article is paraphrased from Goldsman and Nelson (1998a).

REFERENCES

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

Dudewicz, E. J., and S. R. Dalal. 1975. Allocation of observations in ranking and selection with unequal variances. *Sankhyā* B37:28–78.

Dudewicz, E. J., and N. A. Zaino, Jr. 1977. Allowance for correlation in setting simulation run-length via ranking-and-selection procedures. *TIMS Studies in the Management Sciences* 7:51–61.

Gibbons, J. D., I. Olkin and M. Sobel. 1977. *Selecting and Ordering Populations: A New Statistical Methodology*. New York: John Wiley.

Glynn, P. W., and D. L. Iglehart. 1990. Simulation output analysis using standardized time series. *Mathematics of Operations Research* 15:1–16.

Goldsman, D. 1985. Ranking and selection procedures using standardized time series. In *Proceedings of the 1985 Winter Simulation Conference*, ed. D. T. Gantz, G. C. Blais, and S. L. Solomon, 120–124, Piscataway, New Jersey: IEEE.

Goldsman, D., K. Kang, and A. F. Seila. 1999. Cramér–von Mises variance estimators for simulations. *Operations Research* 47:299–309.

Goldsman, D., M. S. Meketon, and L. W. Schruben. 1990. Properties of standardized time series weighted area variance estimators. *Management Science* 36:602–612.

Goldsman, D., and B. L. Nelson. 1998a. Statistical screening, selection, and multiple comparison procedures in computer simulation. In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. J. Medeiros, E. F. Watson, M. Manivannan, and J. S. Carson, 159–166. Piscataway, New Jersey: IEEE.

Goldsman, D., and B. L. Nelson. 1998b. Comparing systems via simulation. In *Handbook of Simulation*, ed. J. Banks, Chapter 8. New York: John Wiley.

Goldsman, D., B. L. Nelson, T. Opicka, and A. A. B. Pritsker. 1999. A ranking and selection project: Experiences from a university-industry collaboration. In *Proceedings of the 1999 Winter Simulation Conference*, ed. P. A. Farrington, H. B. Nembhard, D. Sturrock, and J. Evans. Piscataway, New Jersey: IEEE.

Goldsman, D., and L. W. Schruben. 1990. New confidence interval estimators using standardized time series. *Management Science* 36:393–397.

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

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

Iglehart, D. L. 1977. Simulating stable stochastic systems, VII: Selecting the best system. *TIMS Studies in the Management Sciences* 7:37–50.

Law, A. M., and W. D. Kelton. 1991. *Simulation Modeling & Analysis*. 2nd ed. New York: McGraw-Hill, Inc.

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

Meketon, M. S., and B. W. Schmeiser. 1984. Overlapping batch means: Something for nothing? *Proceedings of the 1984 Winter Simulation Conference*, ed. S. Sheppard, U. W. Pooch, and C. D. Pegden, 227–230. Piscataway, New Jersey: IEEE.

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.

Nelson, B. L., J. Swann, D. Goldsman and W.-M. Song. 1998. Simple procedures for selecting the best system when the number of alternatives is large. Technical Report. Dept. of IEMS, Northwestern Univ., Evanston, Illinois.

Rinott, Y. 1978. On two-stage selection procedures and related probability-inequalities. *Comm. Stat.—Thy. and Meth.* A7:799–811.

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

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

Sullivan, D. W., and J. R. Wilson. 1984. Restricted subset selection for normal populations with unknown and unequal variances. In *Proceedings of the 1984 Winter Simulation Conference*, ed. S. Sheppard, U. Pooch and C. D. Pegden, 123–128. Piscataway, New Jersey: IEEE.

Sullivan, D. W., and J. R. Wilson. 1989. Restricted subset selection procedures for simulation. *Operations Research* 37:52–71.

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

AUTHOR BIOGRAPHIES

DAVID GOLDSMAN is an Associate Professor in the School of Industrial and Systems Engineering at Georgia Tech. His research interests include simulation output analysis and ranking and selection. He is the Simulation Department Editor for *IIE Transactions*, and an Associate Editor for *Operations Research Letters*. He was the Program Chair for the 1995 Winter Simulation Conference.

WILLIAM S. MARSHALL is a Research Engineer at the Georgia Tech Research Institute and Ph.D. candidate in the School of Industrial and Systems Engineering at Georgia Tech. His research interests include radar signal analysis and tracking algorithms, radar system simulation, and ranking and selection procedures.