Proceedings of the 2013 Winter Simulation Conference
R. Pasupathy, S.-H. Kim, A. Tolk, R. Hill, and M. E. Kuhl, eds.

ROBUST SELECTION OF THE BEST

Weiwei Fan
L. Jeff Hong
Xiaowei Zhang

Department of Industrial Engineering and Logistics Management
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong, CHINA

ABSTRACT

Classical ranking-and-selection (R&S) procedures cannot be applied directly to select the best decision in the presence of distributional ambiguity. In this paper we propose a robust selection-of-the-best (RSB) formulation which compares decisions based on their worst-case performances over a finite set of possible distributions and selects the decision with the best worst-case performance. To solve the RSB problems, we design two-layer R&S procedures, either two-stage or fully sequential, under the indifference-zone formulation. The procedure identifies the worst-case distribution in the first stage and the best decision in the second. We prove the statistical validity of these procedures and test their performances numerically.

1 INTRODUCTION

Decision-making processes typically involve selecting the best decision among a set of competing alternatives and the best is often defined as the one with the largest or smallest mean performance. For instance, in risk management, investors call for the best portfolio of financial instruments (e.g., stocks, bonds, derivatives) to maximize the expected return. In inventory management, managers resort to the best decision rule for inventory control (e.g., the amounts to produce, pricing, etc.) to maximize the expected net profit of a firm. Where there is a finite (often small) number of alternatives, ranking and selection (R&S) serves as an important vehicle to select the best decision.

Most of the R&S procedures in the literature assume that a simulation model is provided and the input distributions of the simulation model can be specified accurately beforehand. However, when constructing a distributional model for a stochastic system in practice, the decision maker often faces an undesirable uncertainty in the specification of the distribution family and/or estimation of the pertinent parameters. This uncertainty arises due to either intrinsic randomness or incomplete information. For instance, in risk management, the historical data of financial instruments are often limited; and likewise, in inventory management, the demand distribution may be difficult to be characterized by often-used distribution families. In our paper, we use the term “ambiguity” to describe the above uncertainty issue, namely, the input distribution of a simulation model cannot be specified precisely. It is conceivable that ignoring the ambiguity may result in a misleading or false decision, especially when the ambiguity is deep and profound. This paper, therefore, is devoted to addressing the problem of how to select the best decision in the presence of ambiguity.

The ambiguity is introduced into our framework by assuming that the input distribution of the simulation model belongs to a so-called “ambiguity set”. Nevertheless, to specify an ambiguity set is by no means a trivial task. Morgan, Henrion, and Small (1990) construct the ambiguity set by beginning with a single “best-guess” distribution of the relevant system and then including all the distributions within a certain "distance" in the ambiguity set. The ambiguity sets in Ben-Tal and Nemirovski (2000) and Bertsimas and Sim (2004) have the form of a bounded and symmetric interval and ellipsoidal set, respectively. Another possible approach is based on the moments of the target distribution; see, e.g., Delage and Ye (2010). The ambiguity sets constructed from these approaches either are not flexible enough or often involve an infinite number of scenarios, which are not suitable for the classical R&S framework.

Hence, we use a different type of ambiguity set in this paper. In particular, we assume that the ambiguity set contains a finite number of scenarios for the underlying input distribution. Such a type of ambiguity set is very common in practice. In security pricing, risk managers may choose a finite set of generalized scenarios and use the real financial data to fit each of the scenarios, see for instance, Lesnevski, Nelson, and Staum (2007). In inventory management, the demand distribution is often modeled by a normal distribution; and then, one may consider an ambiguity set of different combinations of means and variances of a normal distribution. In a more general case, one could discretize the possible ranges (such as confidence intervals) of the pertinent parameters of the input distribution to form the ambiguity set.

It then becomes essential to determine how to measure the performance of an alternative decision with various possible scenarios (i.e., ambiguity set). We address this issue by a robust approach in the sense that the performance of an alternative decision is measured by its worst-case performance among all the possible scenarios in the ambiguity set. This robust approach is especially appealing for a conservative person in that it yields a solution which may be suboptimal and yet still performs well even in the worst-case scenario. For instance, an important goal of risk management is to quantify the chance of large losses and thus portfolio managers may prefer the worst-case (maximal) risk measures. Indeed, the worst-case risk measures stand in line with coherent risk measures (Artzner et al. 1999), which have been shown to have many advantages over a widely adopted risk measure called value-at-risk, such as realization of the benefits from diversification, see Lesnevski, Nelson, and Staum (2007) for detailed discussions on the problem. In this paper, we establish a robust framework for R&S and call it *robust selection of the best* (RSB).

We use the indifference-zone formulation proposed by Bechhofer (1954) to design the procedure for the RSB problem. In our robust framework, the RSB problem is reformulated as a minimax optimization problem and thus an natural solution is a two-layer R&S procedure. More specifically, one selects the worse-case scenario for each alternative decision in the first layer R&S, and selects the best decision by comparing the worst performances of all the systems in the second layer R&S. The merit of our two-layer procedure is two-fold. First, its implementation in practice does not require any additional knowledge beyond the classical R&S procedures. Second, the statistical validity of the classical R&S can be easily carried on to our two-layer R&S procedure.

A central problem in our two-layer R&S procedure is to determine the rule of error allocations in order to achieve a prescribed *probability of correct selection* (PCS), say $1-\alpha$. Suppose there are k alternative decisions in the RSB problem. Without loss of generality, we assume there are m scenarios for each decision. (The decisions may have different numbers of scenarios, but the analysis can be extended easily.) Note that one could view the RSB problem as to select the best from $k \times m$ "systems", where an "system" is a pair of a decision and a scenario. Consequently, one could intuitively allocate $1/(km-1)$ fraction of the total allowable error level α to each pairwise comparison between the best and the others. We call this way of error allocation the multiplicative rule.

However, the multiplicative rule is rather conservative in the sense that each comparison is allowed for a very small error level. Hence, it's computationally expensive even for moderate k and m. (Note that the smaller error level is allowed, the more simulation runs have to be computed.) To enhance the efficiency of our two-layer R&S procedure, we fine-tune our analysis and develop a so-called additive rule of error allocation, which allocates $1/(m+k-1)$ fraction of the total allowable error to each necessary pairwise comparison. Obviously, the advantage of the additive rule over the multiplicative rule becomes more significant as k or m increases.

Our research is related to two streams of literatures: selection of the best and robust optimization. In the select-of-the-best literature, R&S procedures are widely used to select the best decision among a fixed and finite set of alternatives, which seek a guaranteed lower bound on the PCS; see Rinott (1978), Branke, Chick, and Schmidt (2007), Kim and Nelson (2007). This literature rarely takes into account the distributional ambiguity of the simulation models. The robust optimization literature focuses on optimization problems in which the uncertainty appears in the objective function. See Ben-Tal, El Ghaoui, and Nemirovski (2009) for an extensive treatment on the subject. However, in the robust optimization literature, it typically assumes the objective function is available explicitly in closed-form, which does not hold in the context of simulation optimization.

The rest of paper is organized as follows. In Section 2, we introduce our framework for the robust selection of the best. Based on this framework, a two-layer R&S procedure is developed in Section 3. Section 4 discusses the associated error allocation rule and Section 5 presents numerical results. Section 6 concludes this paper.

2 A ROBUST FRAMEWORK

Define $\mathscr{S}$ as a group of decisions, where $\mathscr{S} = \{s_1, s_2, \dots, s_k\}$. Let $g(s,\xi)$ denote the performance value of the decision s, where ξ is the uncertainty parameter in the performance function and follows an unknown distribution P_0. Indeed, stochastic systems in reality are usually too complex to derive the closed-form expression of $g(s,\xi)$ and it can only be accessible by direct observations or simulation experiments given s and the distribution of ξ. In the simulation study, the performances are evaluated by $\mathbb{E}_{P_0}[g(s,\xi)]$ for every $s \in \mathscr{S}$.

Suppose that we are interested in selecting the best decision in $\mathscr{S}$, where the best decision is defined as the one with the smallest performance value. In particular, the selection of the best is formulated as

$$\min_{s\in\mathscr{S}} \mathbb{E}_{P_0}[g(s,\xi)].$$

In the presence of ambiguity in specifying P_0, it is difficult to accurately estimate the mean performance $\mathbb{E}_{P_0}[g(s,\xi)]$ of each decision s. To model the ambiguity, we assume that the possible probability distributions that ξ may follow are included in a pre-specified ambiguity set, defined as $\mathscr{P} = \{P_1, P_2, \dots, P_m\}$. Notice that ambiguity sets can be different for different systems. For simplicity, we assume that the ambiguity sets $\mathbb{P}$ are identical for all the systems. It is worth emphasizing that our ambiguity set incorporates the parameter ambiguity as well as distributional ambiguity. Particularly, it can contain a single family of probability distributions with various parameters, or various families of probability distributions.

Further, we employ a robust approach to evaluate the performance of a given decision s, which varies over the ambiguity set $\mathscr{P}$. More specifically, the performance of a given decision is measured by its worst-case performance, namely $\max_{P\in\mathscr{P}} \mathbb{E}_P[g(s,\xi)]$. It follows that the selection of the best in the presence of ambiguity can be formulated as

$$\min_{s\in\mathscr{S}} \max_{P\in\mathscr{P}} \mathbb{E}_P[g(s,\xi)], \tag{1}$$

which we call the robust selection of the best (RSB). This robust framework is appealing in that the decision it selects, though may not be optimal, performs well even in the worst-case scenario. The robust framework is able to prevent potential high risk.

In light of the minimax optimization formulation (1), we propose a two-layer R&S procedure. In the first stage, we solve the inner maximization problem. Particularly, for each fixed $s \in \mathscr{S}$, a R&S procedure is committed to select the worst scenario $P \in \mathscr{P}$ with an appropriate PCS, which poses the maximum $\mathbb{E}_P[g(s,\xi)]$ over $P \in \mathscr{P}$. In the second stage, we solve the outer minimization problem based on the selected worst scenario $P(s)$ for each s. Another R&S procedure is then conducted to select the best decision s with the minimal worst-case performance an appropriate PCS. After this two-layer R&S procedure, the best decision is selected with a PCS that is no less than the prescribed level.

We adopt the indifference-zone (IZ) formulation proposed by Bechhofer (1954), which guarantees to select the best decision with a pre-specified PCS whenever the mean performance of the best decision is smaller than that of the second-best one with a deviation at least δ, where δ is the smallest difference worth detecting; see, e.g., Rinott (1978) for more details. Recall that, in our two-layer procedure, two selections are made sequentially. Hence, we define δ_1, δ_2 as two corresponding IZ parameters, which refer to the smallest differences worth detecting while selecting the worst-case performance and the best alternative, respectively.

To facilitate our presentation, we refer "system (i,j)" as the pair of decision s_i and probability distribution scenario P_j. For each system (i,j), let $\mu_j^i = E_{P_j}[g(s_i,\xi)]$ denote its mean performance value and let $(\sigma_j^i)^2$ its unknown variance. For $i = 1,2,\ldots,k$, without loss of generality, we assume $\mu_1^i \geq \mu_2^i \geq \ldots \geq \mu_m^i$. Namely, $(i,1)$ is the worst system among all (i,j) for $j = 1,\ldots,m$. Furthermore, we assume $\mu_1^1 \leq \mu_1^2 \leq \ldots \leq \mu_1^k$. Hence, the correct selection (i.e. the best system based on the performance in the worst-case scenario) is to select system $(1,1)$.

3 A TWO-LAYER R&S PROCEDURE

In this section, we develop a two-layer R&S procedure for the RSB problem. In the first layer, for each fixed $i = 1,2,\ldots,k$ one selects the worst system among systems (i,j) for $j = 1,2,\ldots,m$; in the second layer, one selects the best system among the selected worst systems for $i = 1,2,\ldots,k$. Note that variances of the alternative systems are unknown, so when conducting the R&S in each layer one needs to first do an initial sampling in order to estimate the variances of the systems and then do additional sampling if necessary in order to eliminate any system.

Let β_1 and β_2 be the error level allocated to each necessary pairwise comparison in the first and the second layer of R&S. We will discuss how to appropriately determine β_1 and β_2 in order to achieve a prescribed PCS in Section 4. Supposing for now that β_1 and β_2 are given, our two-layer R&S procedure is described as follows.

Procedure 1 (Two-layer R&S Procedure for RSB)
(1) Setup:
Determine the target overall PCS $(1-\alpha)$, where $0 < \alpha < 1 - 1/(km-1)$ and β_1, β_2. For each fixed i, choose IZ parameter δ_1 across systems (i,j) for $j = 1,2,\ldots,m$, and IZ parameter δ_2 across systems $(i,1)$ for $i = 1,2,\ldots,k$.
(2) Select the worst scenario:
For each fixed $i = 1,2,\ldots,k$, conduct a R&S procedure to select a system which deviates at most δ_1 from the worst one among systems (i,j) for $j = 1,\ldots,m$ with probability at least $1-(m-1)\beta_1$.
(3) Select the best decision:
Conduct a R&S procedure to select a system which deviates at most δ_2 from the best one among the systems selected in Step (2) for each $i = 1,\ldots,k$ with probability at least $1-(k-1)\beta_2$.

Remark 1 In the step (2) and (3), any existing procedure (such as Rinott and KN procedures) can be plugged in to satisfy the IZ criteria.

Note that an incorrect selection (ICS) event happens when the best system is eliminated at some point during the procedure. Let $\varepsilon((i,j),(p,q))$ denote the event that system (i,j) eliminates system (p,q). Then, the target probabilities in Step (2) of the above procedure can be achieved if

$$\mathbb{P}(\varepsilon((i,j),(i,1))) \leq \beta_1, \text{ for } j = 2,3,\ldots,m, i = 1,2,\ldots,k. \tag{2}$$

Likewise, the target probability in Step (3) can be achieved if

$$\mathbb{P}(\varepsilon((i,1),(1,1))) \leq \beta_2, \text{ for } i = 2,3,\ldots,k. \tag{3}$$

Remark: In fact, we can streamline the above procedure by utilizing "batch elimination". In particular, for any i, one can eliminate all the systems (i,j) still in contention if there exists some l such that system (i,j_0) has a larger performance value than all system (l,j_1) in contention for some j_0. In this case, decision i is inferior than decision l in terms of worst-case performances. We refer the readers to the Appendix for details.

4 ERROR ALLOCATIONS

In our two-layer procedure, two separate ranking-and-selection's are conducted sequentially in order to select the best system among $k \times m$ systems. Intuitively, its efficiency is at least the same as the problem (M), where

(M): Selecting the best one from $k \times m$ systems using R&S.

In particular, an ICS event happens if system $(1,1)$ is eliminated by any one of the other $(km-1)$ systems. Therefore, it is natural to allocate $1/(km-1)$ fraction of the total allowable error α to each one of $(km-1)$ ICS events, namely $\{\varepsilon((i,j),(i,1)), i=1,\dots,k, j=2,\dots,m)\}$ and $\{\varepsilon((i,1)(1,1)), i=2,\dots,k\}$. We call this way of error allocations the multiplicative rule. Based on the multiplicative rule of error allocations, we will justify the statistical validity of our two-layer procedure in the following theorem.

Theorem 1 For the two-layer procedure in Section 3,

$$\mathbb{P}(CS) \geq 1-\alpha, \quad \text{if} \quad k(m-1)\beta_1 + (k-1)\beta_2 \leq \alpha$$

Specially, we can take $\beta_1 = \beta_2 = \alpha/(km-1)$.

Proof. Notice that,

$$\text{CS} \supset \left\{ \bigcap_{i=1,2,\dots,k} \bigcap_{j\neq 1} \{\text{system } (i,1) \text{ eliminates } (i,j)\} \right\} \bigcap \left\{ \bigcap_{i=2,3,\dots,k} \{\text{system } (1,1) \text{ eliminates } (i,1)\} \right\}.$$

Therefore,

$$\begin{aligned} \mathbb{P}(\text{CS}) &\geq 1 - \sum_{i=1}^{k}\sum_{j=2}^{m} \mathbb{P}(\{(i,j) \text{ eliminates } (i,1)\}) - \sum_{i=2}^{k} \mathbb{P}(\{(i,1) \text{ eliminates } (1,1)\}) \\ &\geq 1 - k(m-1)\beta_1 - (k-1)\beta_2 \\ &\geq 1-\alpha. \end{aligned}$$

The first inequality follows from the Bonferroni inequality and the second from (2) and (3). Moreover, it is straightforward to verify that $k(m-1)\beta_1 + (k-1)\beta_2 \leq \alpha$ if $\beta_1 = \beta_2 = \alpha/(km-1)$. □

However, the multiplicative rule is rather conservative, because it treats all the $(km-1)$ ICS events equally important. Note that our goal is to select the best system $(1,1)$, which does not necessarily require that each $(i,1)$ for $i=1,2,\dots,k$ is selected in the first layer R&S. Instead, we only need to guarantee that any system selected in the first layer R&S can be eventually eliminated by system $(1,1)$ except itself. In other words, some (in fact, a large fraction) of the $(km-1)$ ICS events are not critical. By virtue of this insight, we claim that our two-layer R&S procedure indeed performs as efficiently as problem (A), where

(A): Selecting the best one from $k+m-1$ systems using R&S.

Particularly, one can allocate $1/(k+m-2)$ fraction of the total allowable error α to each of the "critical" ICS events, $\varepsilon((i,j),(i,1))(j \neq 1)$ and $\varepsilon((i,1)(1,1))(i \neq 1)$. We call it the additive rule of error allocations. The following theorem will certify our claim and show that the additive rule is enough to ensure statistical validity, indicating multiplicative rule is too conservative.

Theorem 2 For the two-layer basic procedures stated in Section 3,

$$\mathbb{P}(\mathrm{CS}) \geq 1-\alpha, \quad \text{if} \quad (m-1)\beta_1+(k-1)\beta_2 \leq \alpha.$$

Specially, we can take $\beta_1=\beta_2=\alpha/(k+m-2)$.

Proof. Notice that the correct selection corresponds to the fact that s_1 is selected. Therefore, with a similar argument as in Theorem 1,

$$\begin{aligned}
\mathbb{P}(\mathrm{CS}) &= \mathbb{P}(\cap_{i=2}^{k}\{\text{ decision } s_1 \text{ eliminates decision } s_i\}) \\
&\geq \mathbb{P}(\cap_{i=2}^{k}\{\text{ decision } s_1 \text{ eliminates decision } s_i\} \cap \cap_{j=2}^{m}\{(1,1) \text{ eliminates } (1,j)\}) \\
&\geq \mathbb{P}(\cap_{i=2}^{k}\{(1,1) \text{ eliminates } (i,1)\} \cap \cap_{j=2}^{m}\{(1,1) \text{ eliminates } (1,j)\}) \\
&\geq 1-\sum_{i=2}^{k}\mathbb{P}(\{(i,1) \text{ eliminates } (1,1)\})-\sum_{j=2}^{m}\mathbb{P}(\{(1,j) \text{ eliminates } (1,1)\}) \\
&\geq 1-(k-1)\beta_2-(m-1)\beta_1 \\
&\geq 1-\alpha.
\end{aligned}$$

The first inequality holds because we add another constraint, i.e., that system (1,1) is the best among systems $(1,j)(j=1,2,\ldots,m)$. Recall the best system is defined as the system with the smallest average performance value while considering x_i and x_j. Thus it's easier to stand out for system (1,1) compared with system (i,1) than with all system $(i,j)(j=1,2,\ldots,m)$, meaning that the second inequality holds. The third inequality follows from Bonferroni inequality. Moreover, it is straightforward to verify that $(m-1)\beta_1+(k-1)\beta_2 \leq \alpha$ if $\beta_1=\beta_2=\alpha/(k+m-2)$. □

Clearly, the advantage of the additive rule becomes more significant as k or m increases. It is conceivable that the ambiguity set in practice could contain a large number of candidate probability distributions. In order to achieve a given PCS, the computational cost would be prohibitively high for the multiplicative rule of error allocation and yet may be acceptable if the additive rule is applied.

5 NUMERICAL EXPERIMENTS

Suppose that $X_{ij} \sim N(\mu_{ij},\sigma_{ij}^2)$ for $i=1,2,\ldots,k$ and $j=1,2,\ldots,m$. We consider the slippage configuration of means,

$$\mu_{ij} = \begin{cases} \delta, & i \neq 1, j=1, \\ -\delta, & i=1, j\neq 1, \\ 0, & \text{otherwise} . \end{cases} \tag{4}$$

The best system is defined with the system having $\min_{1\leq i\leq k}\max_{1\leq j\leq m}\mathbb{E}[X_{ij}]$. Hence, our target is to select the best system $(1,1)$, which corresponds to X_{11}. With the above configuration for the means, we will consider three variance configurations: (1) equal-variance configuration with $\sigma_{ij}^2=1$; (2) increasing-variance configuration with $\sigma_{ij}^2=1+(j-1)\delta$ for all i; (3) decreasing-variance configuration with $\sigma_{ij}^2=1/[1+(j-1)\delta]$ for all i.

Let the target PCS be $1-\alpha=0.95$. Suppose that β_1 and β_2 are chosen based on the additive rule in Theorem 2, i.e., $\beta_1=\beta_2=\alpha/(k+m-2)$. In addition, we set the IZ parameters as $\delta_1=\delta_2=0.2$.

The implementation of our two-layer R&S procedure follows that in **Procedure 1**. Under each variance configuration as well as combination of k and m, we replicate our procedure 1000 times in order to estimate the realized probability of correct selection (PCS), the sample size used to select the worst scenario (SS) for each decision s_i, and the average total sample size (TS).

Table 1: Summary of PCS, average sample sizes when m varies and $k = 5$ under three variance configurations

		Equal			Decreasing			Increasing		
m	k	PCS	SS	TS	PCS	SS	TS	PCS	SS	TS
	5	0.994	9.22E+02	5.03E+03	0.989	7.50E+02	4.13E+03	0.990	1.07E+03	5.83E+03
	10	0.997	1.05E+03	1.17E+04	0.990	8.55E+02	9.66E+03	0.996	1.20E+03	1.35E+04
5	15	0.993	1.13E+03	1.92E+04	0.983	9.35E+02	1.60E+04	0.993	1.29E+03	2.22E+04
	20	0.989	1.20E+03	2.74E+04	0.978	9.87E+02	2.26E+04	0.992	1.36E+03	3.14E+04
	25	0.994	1.25E+03	3.59E+04	0.992	1.03E+03	2.98E+04	0.994	1.42E+03	4.10E+04

Table 2: Summary of PCS, average sample sizes when k varies and $m = 5$ under three variance configurations

		Equal			Decreasing			Increasing		
k	m	PCS	SS	TS	PCS	SS	TS	PCS	SS	TS
	5	0.993	9.23E+02	5.03E+03	0.989	750.2218	4.13E+03	0.990	1.07E+03	5.83E+03
	10	0.997	2.03E+03	1.06E+04	0.993	1.39E+03	7.32E+03	0.997	2.77E+03	1.45E+04
5	15	0.998	3.26E+03	1.68E+04	0.994	1.90E+03	9.93E+03	0.999	5.09E+03	4.20E+03
	20	0.997	4.56E+03	2.33E+04	0.990	2.32E+03	1.20E+04	1.000	8.00E+03	4.12E+04
	25	0.998	5.97E+03	3.03E+04	0.998	2.67E+03	1.38E+04	0.998	1.16E+04	5.93E+04

The results listed in Table 1 and Table 2 show that the realized PCS values are all above the desired PCS value, which confirms the statistical validity of the procedure. Moreover, we find that it takes fewer samples to select the best for decreasing-variance configuration than for the equal-variance configuration, while the contrary happens to increasing-variance configuration. It is reasonable because it is often more difficult to eliminate inferior systems while their variances are large.

6 CONCLUSIONS AND FUTURE WORK

In this paper, we consider the selection of the best in the presence ambiguity of the input distributions. To address this problem, we have proposed a robust framework for the selection of the best and developed a two-layer R&S procedure. In particular, this procedure selects the worst scenario for each alternative decision first, and the best system based on their performances in the worse-case scenario. Though intuitive, the multiplicative rule of error allocation is rather conservative. We have proposed that the additive rule of error allocation to enhance the efficiency of the two-layer R&S procedure.

Note that the numerical results show that the two-layer R&S procedure works well for moderate k and m. However, it may cost a large amount of computational budget when k or m is large. Hence, it would be interesting to make this procedure adapted to large-scale problems.

ACKNOWLEDGEMENTS

The authors are supported in part by the Hong Kong Research Grants Council under General Research Funds 613011 and 613012.

A Fully Sequential Procedure with Batch Eliminations for RSB problem

Setup: Select the total desired PCS $1-\alpha$. For each fixed i, choose IZ parameter δ_1 across systems $(i,j)(j=1,2,\ldots,m)$ and the error tolerance β_1 to each pair comparison between them, IZ parameter δ_2 across systems $(i,1)(i=1,2,\ldots,k)$, error tolerance β_2 to each pair comparison between them. Determine common first-stage sample size $n_0 \geq 2$. Calculate η_1, η_2, c_1 and c_2 as described below.

Initialization: Let $I = \{(i,j) : i = 1,2,\ldots,k, j = 1,2,\ldots,m\}$ be the set of system still in contention. Obtain n_0 observations $X_{ij}(r), r = 1,2,\ldots,m$, from each system (i,j). Let $h_1^2 = 2c_1\eta_1$ and $h_2^2 = 2c_2\eta_2$. Set $n = n_0$.

Update: Let $X_{ij}(r) = n^{-1}\sum_{r=1}^{n} X_{ij}(r)$ denote the sample mean, for all $(i,j) \in I$. Calculate

$$S^2_{(i,j)(s,l)}(n) = \frac{1}{n-1}\sum_{r=1}^{n}(X_{ij}(r) - X_{sl}(r) - [\bar{X}_{ij}(n) - \bar{X}_{sl}(n)])^2,$$

the sample variance between system (i,j) and (s,l).

Screening: Set $I^{\text{old}} = I$.

1. Screening among systems $(i,j)(j = 1,2,\ldots,m)$ for each fixed i. Let

$$I = I^{\text{old}} \setminus \{(i,j) : (i,j) \in I^{\text{old}} \text{ and } \bar{X}_{ij}(n) \leq \bar{X}_{il}(n) - W_{(i,j)(i,l)}(n), \text{ for some } (i,l) \in I^{\text{old}}, l \neq j\},$$

where

$$W_{(i,j)(i,l)}(n) = \max\left\{0, \frac{\delta_1}{2c_1 n}\left(\frac{h_1^2 S^2_{(i,j)(i,l)}(n)}{\delta_1^2} - n\right)\right\}.$$

2. Screening among systems (i,j) for different i. Let

$$I = I \setminus \{(i,j) \in I : \exists (s,l) \in I, \forall (i,t) \in I, s.t., s \neq i \text{ and } \bar{X}_{it}(n) \geq \bar{X}_{sl}(n) + W_{(i,t)(s,l)}(n)\},$$

where

$$W_{(i,t)(s,l)}(n) = \max\left\{0, \frac{\delta_2}{2c_2 n}\left(\frac{h_2^2 S^2_{(i,j)(s,l)}(n)}{\delta_2^2} - n\right)\right\}.$$

Stopping Rule: If $|I| = 1$, then stop and select the system whose index is in I as the best. Otherwise take one additional output from each system in I and set $n = n+1$ and go to ***Update***.

Constants: The constants c_1, c_2 may be any non-negative integers. Respectively, the constants η_1, η_2 are the solutions to the following two equations,

$$\begin{aligned} g(\eta_1) &= \sum_{l=1}^{c_1}(-1)^{l+1}(1 - I(l = c_1)/2)\exp\left(-\frac{\eta_1}{c_1}(2c_1 - l)l\right) = \beta_1, \\ g(\eta_2) &= \sum_{l=1}^{c_2}(-1)^{l+1}(1 - I(l = c_2)/2)\exp\left(-\frac{\eta_2}{c_2}(2c_2 - l)l\right) = \beta_2. \end{aligned}$$

Remarks:

- In our fully sequential procedure, $W_{(ij)(sl)}(n)$ defines a continuous triangular region for partial sum, i.e., $\sum_{r=1}^{n}(X_{ij}(r) - X_{sl}(r))$. As long as the partial sum stays in this triangular region, sampling continuous and no elimination is made. Otherwise, sampling stops and a system is eliminated.
- In the ***Screening*** step, we conduct a two-step elimination. In the first step, pair comparisons only involves systems $(i,j)(j = 1,2,\ldots,m)$ for each i. Unlike classical fully sequential procedure, we add a second stage into our procedure, which enables our procedure to be more efficient. In this step, all systems (i,j) in I involving a fixed i will be deleted if they are found to be inferior than some system (s,l) in I for $s \neq i$. In other words, the second step may bring batch eliminations, which boosts searching for the best system.
- An initial sample of size n_0 is taken to estimate weights and variances. n_0 is usually selected to be guarantees that enough data are available to estimate the variance, but that much to distinguish one system from another.
- After applying this procedure stated above, the best system is selected out with a guaranteed at least $1 - \alpha$ probability, when errors terms β_1, β_2 in the *Initialization* Step are appropriately chosen beforehand.

REFERENCES

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. 1999. "Coherent measures of risk". *Mathematical Finance* 9 (3): 203–228.

Bechhofer, R. E. 1954. "A single-sample multiple decision procedure for ranking means of normal populations with known variances". *The Annals of Mathematical Statistics*:16–39.

Ben-Tal, A., L. El Ghaoui, and A. Nemirovski. 2009. *Robust optimization*. Princeton University Press.

Ben-Tal, A., and A. Nemirovski. 2000. "Robust solutions of linear programming problems contaminated with uncertain data". *Mathematical Programming* 88 (3): 411–424.

Bertsimas, D., and M. Sim. 2004. "The price of robustness". *Operations Research* 52 (1): 35–53.

Branke, J., S. E. Chick, and C. Schmidt. 2007. "Selecting a selection procedure". *Management Science* 53 (12): 1916–1932.

Delage, E., and Y. Ye. 2010. "Distributionally robust optimization under moment uncertainty with application to data-driven problems". *Operations Research* 58 (3): 595–612.

Kim, S.-H., and B. L. Nelson. 2007. "Recent advances in ranking and selection". In *Proceedings of the 39th Conference on Winter Simulation*, 162–172. IEEE Press.

Lesnevski, V., B. L. Nelson, and J. Staum. 2007. "Simulation of coherent risk measures based on generalized scenarios". *Management Science* 53 (11): 1756–1769.

Morgan, M. G., M. Henrion, and M. Small. 1990. *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge University Press.

Rinott, Y. 1978. "On two-stage selection procedures and related probability-inequalities". *Communications in Statistics-Theory and Methods* 7 (8): 799–811.

AUTHOR BIOGRAPHIES

WEIWEI FAN is a PhD candidate in the Department of Industrial Engineering and Logistics Management at the Hong Kong University of Science and Technology. Her research interests include robust simulation and data analysis. Her email address is wfan@ust.hk.

L. JEFF HONG is a professor in the Department of Industrial Engineering and Logistics Management at the Hong Kong University of Science and Technology. His research interests include Monte Carlo method, financial engineering, and stochastic optimization. He is currently an associate editor for *Operations Research*, *Naval Research Logistics* and *ACM Transactions on Modeling and Computer Simulation*. His email address is hongl@ust.hk.

XIAOWEI ZHANG is an Assistant Professor in the Department of Industrial Engineering and Logistics Management at Hong Kong University of Science and Technology. He received his Ph.D. in Operations Research from Stanford University in 2011. He is a member of INFORMS and his research interests include rare event simulation and stochastic modeling in service engineering. His email address is xiaoweiz@ust.hk.