Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

A BAYESIAN APPROACH TO FEASIBILITY DETERMINATION

Roberto Szechtman

Department of Operations Research
Naval Postgraduate School
1411 Cunningham Road
Monterey, CA 93943, USA

Enver Yücesan

Technology and Operations Management Area
INSEAD
Boulevard de Constance
77305 Fontainebleau, FRANCE

ABSTRACT

We propose a computing budget allocation scheme for feasibility determination in a stochastic setting. More formally, we propose a Bayesian approach to determine whether a system belongs to a given set based on performance measures estimated through Monte Carlo simulation. We introduce two adaptive approaches in the sense that the computational budget is allocated dynamically based on the samples obtained thus far. The first approach determines the number of additional samples required so that the posterior probability that a system's mean performance is correctly classified is at least $1-\delta$ in expectation, while the second approach determines the number of additional samples so that the posterior probability that the system mean lies inside or outside of the feasible region is at least $1-\delta$ with a desired probability. Preliminary numerical experiments are reported.

1 INTRODUCTION

We consider r systems, each with unknown performance measure $\mu_1,\ldots,\mu_r \in R^c$. We assume a normal prior structure, whereby the prior of each μ_i, $i=1,\ldots,r$, is normal, and the analyst can draw samples $Y_{i,1},Y_{i,2},\ldots$, that are normally distributed around μ_i. Given a set $\Gamma \subseteq R^c$, system i is said to be *in* Γ if the posterior probability that μ_i (given the samples $Y_{i,1},\ldots,Y_{i,n_i}$) is in Γ is larger than some threshold $1-\delta$, and *not in* Γ if the posterior probability that μ_i is not in Γ is at least $1-\delta$. The goal is to allocate a budget of N total simulation replications in order to classify as many systems as possible.

An interesting problem in its own right, feasibility determination has recently attracted much attention within the context of ranking and selection (R&S) in the presence of stochastic constraints. Traditionally, the overwhelming majority of the R&S research focuses on a single unconstrained performance measure. To the best of our knowledge, Santner and Tamhane (1984) were the first authors to propose a two-stage procedure with a constraint on variance. In a setting where R&S is based on a primary performance measure subject to the feasibility of a (possibly correlated) secondary performance measure, Andradottir, Goldsman, and Kim (2005) and Andradottir and Kim (2010) propose a two-phase approach whereby phase I identifies feasible systems while phase II selects the best one among them. With the objective of accelerating the first phase, Batur and Kim (2010) have introduced procedures for finding a set of feasible or near-feasible systems in the presence of orthant constraints. Under the assumption of normality, Lee et al. (2012) have extended the OCBA approach to the selection of the best design under stochastic constraints. Hunter and Pasupathy (2012) have addressed the problem of selecting an optimal system in the presence of stochastic constraints using large deviations theory under the assumption that the underlying distributional family of the simulation estimator is known or assumed. Szechtman and Yücesan (2008) relax this assumption by estimating the large deviations rate function on the fly. Pasupathy et al. (2014) also offer a tractable instance of a bi-level optimization problem to solve the constrained R&S problems efficiently.

Another approach is to maximize the number of correctly classified systems given a budget constraint. However, this problem is intractable even in the simple case when Γ is defined by linear inequalities (Powell and Ryzhov (2012)). An alternative approach proposed here is a class of heuristics for allocating a budget of N replications to sample from the r systems in such a way that, any system that is classified as in Γ or not in Γ, is classified with a user-specified level of confidence, $1-\delta$, using an *adaptive budget allocation algorithm*, which can be outlined as follows:

Step 0. Perform some initial number of simulation replications, n_i, for each system i and compute the posterior distribution for μ_i.

Step 1. For each system i that has not yet been classified, use the estimates from the simulations to decide whether this system can be declared as in Γ or not in Γ with confidence $1-\delta$. If so, eliminate this system from further budget considerations.

Step 2. If the budget is depleted, stop the algorithm after classifying any remaining unclassified system, noting that these systems are not classified at the desired level of confidence.

Step 3. For each unclassified system i, decide how many additional simulation replications, m_i, would be required to classify it. Decide which systems to sample from, and return to Step 1.

The specific algorithms proposed here are based on a Bayesian approach, which requires a prior $\mu_\mathbf{i} \sim N(\mu_{i,0},\Sigma_{i,0})$, where $\mu_{i,0} \in R^c$ and $\Sigma_{i,0}$ is a $(c \times c)$ diagonal covariance matrix with its j'th diagonal element equal to $\sigma^2_{i,j,0}$; the independence assumption will be subsequently relaxed. Furthermore, using simulation, assume that we can draw i.i.d. samples $Y_{i,1}, Y_{i,2}, \ldots$ from $N(\mu_i,\Sigma_i)$, where Σ_i is a diagonal matrix with diagonal elements $\sigma^2_{i,1},\ldots,\sigma^2_{i,c}$. In settings where variances are known, we have that, after n_i samples are drawn from system i, $\mu_{i,j}|Y_{i,j,1},Y_{i,j,2},\ldots,Y_{i,j,n_i}$ also is normally distributed.

Using the posterior distribution, we introduce two adaptive methods through which, in Step 3, the computational budget is allocated dynamically based on the samples obtained thus far. The first approach (introduced in Section 3) determines the number of additional samples required so that the posterior probability that a system's mean performance is correctly classified to be inside or outside of Γ is at least 1-δ, in expectation. The second approach (introduced in Section 5) determines the number of additional samples so that the posterior probability that the system mean lies inside or outside of Γ is at least $1-\delta$ with probability $1-\eta$, for $0<\eta<1$. In either procedure, the system that requires the fewest additional samples for classification is selected for sampling. Preliminary numerical illustrations are presented for both approaches.

2 THE FRAMEWORK

Let $\Gamma = \{y \in R^c : y \le b\}$, where $b \in R^c$. Using the constraint structure, we get for system i that

$$P(\mu_i \in \Gamma|Y_{i,1},\ldots,Y_{i,n_i}) = P(\cap_{j=1}^c \mu_{i,j} \le b_j|Y_{i,1},\ldots,Y_{i,n_i}) \le \min_{j=1,\ldots,c} P(\mu_{i,j} \le b_j|Y_{i,j,1},\ldots,Y_{i,j,n_i}),$$

and

$$\begin{aligned} P(\mu_i \notin \Gamma|Y_{i,1},\ldots,Y_{i,n_i}) &= P(\cup_{j=1}^c \mu_j > b_j|Y_{i,1},\ldots,Y_{i,n_i}) \\ &\le c \max_{j=1,\ldots,c} P(\mu_{i,j} > b_j|Y_{i,j,1},\ldots,Y_{i,j,n_i}) \\ &= c(1-\min_{j=1,\ldots,c} P(\mu_{i,j} \le b_j|Y_{i,j,1},\ldots,Y_{i,j,n_i})). \end{aligned}$$

For some $\delta \in (0,1/2)$, we declare system i not in Γ if $P(\mu_i \in \Gamma|Y_{i,1},\ldots,Y_{i,n_i}) \le \delta$ and declare the system in Γ if $P(\mu_i \notin \Gamma|Y_{i,1},\ldots,Y_{i,n_i}) \le \delta$; otherwise, we continue sampling the system. From the equations

above, we declare system i not in Γ if $\min_j P(\mu_{i,j} \le b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i}) \le \delta$, declare the system in Γ if $\min_j P(\mu_{i,j} \le b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i}) \ge 1-\delta/c$, or continue sampling if $\delta < \min_j P(\mu_{i,j} \le b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i}) < 1-\delta/c$. Remark that $\delta \in (0,1/2)$ ensures that $\delta < 1-\delta/c$ for $c \ge 1$. The upper bound can be made less conservative by using the bound $P(\cup_j \mu_{i,j} > b_j | Y_{i,1}, \ldots, Y_{i,n_i}) \le \sum_{j=1}^{c} P(\mu_{i,j} > b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i})$, but this makes the developments of Sections 3 and 5 less tractable. If follows from these developments that each system will be correctly classified with confidence $1-\delta$.

A benefit of this procedure is that it does not follow a "worst case" configuration approach; instead the constraint that makes it most likely to make a misclassification dictates when the analyst can stop sampling. Other performance measures, such as maximizing the probability of classifying k systems, for $k \in \{1, \ldots, r\}$, also are of interest and will be treated elsewhere.

In Step 3 of our adaptive budget allocation algorithm, a greedy type of policy selects to sample from the system that increases the most the probability of classification, among those that remain unclassified; i.e., that pushes $\min_j P(\mu_{i,j} \le b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i+m})$ the most to either exit threshold after drawing m extra samples, among the systems with $\delta < \min_j P(\mu_{i,j} \le b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i}) < 1-\delta/c$. Since $P(\mu_{i,j} \le b_j | Y_{i,j,1}, \ldots, Y_{i,j,n_i+m})$ is random given the samples $Y_{i,j,1}, \ldots, Y_{i,j,n_i}$, we present two approaches, one dealing with its mean (in Sections 3 and 4) and another with a percentile (in Sections 5 and 6).

3 STEP 3: NUMBER OF SAMPLES TO CLASSIFY A SYSTEM ON AVERAGE

In order to understand how $P(\mu_j \le b_j | Y_{j,1}, \ldots, Y_{j,n})$ changes for each constraint j after drawing a new sample for a given system, we develop an expression for $E_{Y_{j,n+1}|Y_{j,1},\ldots,Y_{j,n}}[P(\mu_j \le b_j | Y_{j,1}, \ldots, Y_{j,n}, Y_{j,n+1})]$. In this and the next sections, we consider an arbitrary system, and drop the i index to simplify the notation.

Recall that there is a prior $\mu_j \sim N(\mu_{j,0}, \sigma^2_{j,0})$ and that the samples $Y_{j,1}, Y_{j,2}, \ldots$, conditional on μ_j, are distributed as $N(\mu_j, \sigma^2_j)$. It is well known (see Gelman et al. (2013)), that, given $n+1$ observations, the posterior mean $\mu_{j,n+1}$ is distributed as

$$\mu_{j,n+1} | Y_{j,1}, \ldots, Y_{j,n}, Y_{j,n+1} \sim N\left(\frac{\frac{\mu_{j,0}}{\sigma^2_{j,0}} + \frac{n\bar{Y}_j(n) + Y_{j,n+1}}{\sigma^2_j}}{1/\sigma^2_{j,0} + (n+1)/\sigma^2_j}, \frac{1}{1/\sigma^2_{j,0} + (n+1)/\sigma^2_j} \right),$$

and $Y_{j,n+1} | Y_{j,1}, \ldots, Y_{j,n} \sim N(\alpha_{j,n}, \beta^2_{j,n})$, with

$$\alpha_{j,n} = \frac{\mu_{j,0}/\sigma^2_{j,0} + n\bar{Y}_j(n)/\sigma^2_j}{1/\sigma^2_{j,0} + n/\sigma^2_j} \text{ and } \beta^2_{j,n} = \sigma^2_j + \frac{1}{1/\sigma^2_{j,0} + n/\sigma^2_j}. \tag{1}$$

Then,

$$\begin{aligned}
&E_{Y_{j,n+1}|Y_{j,1},\ldots,Y_{j,n}}[P(\mu_j \le b_j | Y_{j,1}, \ldots, Y_{j,n}, Y_{j,n+1})] \\
&= \int P(\mu_j \le b_j | Y_{j,1}, \ldots, Y_{j,n}, Y_{j,n+1}) \times p(Y_{j,n+1} | Y_{j,1}, \ldots, Y_{j,n}) \\
&= E_Z\left[\Phi\left(\left(b_j - \frac{\mu_{j,0}/\sigma^2_{j,0} + \frac{n\bar{Y}_j(n) + \alpha_{j,n} + \beta_{j,n} Z}{\sigma^2_j}}{1/\sigma^2_{j,0} + (n+1)/\sigma^2_j} \right) (1/\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2} \right) \right] \\
&= E_Z\left[\Phi\left(\left(\frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n} - \beta_{j,n} Z}{\sigma^2_j} \right) \frac{1}{(1/\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2}} \right) \right] \qquad (2) \\
&= \Phi\left(\frac{\frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n}}{\sigma^2_j}}{\left(1/\sigma^2_{j,0} + (n+1)/\sigma^2_j + \beta^2_{j,n}/\sigma^4_j\right)^{1/2}} \right), \qquad (3)
\end{aligned}$$

where $\Phi(\cdot)$ is the cdf of $Z \sim N(0,1)$.

The same approach applies if the analyst adds m new samples $Y_{j,n+1},\ldots,Y_{j,n+m}$, resulting in

$$E_{Y_{j,n+1},\ldots,Y_{j,n+m}|Y_{j,1},\ldots,Y_{j,n}}[P(\mu_j \le b_j|Y_{j,1},\ldots,Y_{j,n},Y_{j,n+1},\ldots,Y_{j,n+m})] = \Phi\left(\frac{\frac{b_j-\mu_{j,0}}{\sigma^2_{j,0}}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma^2_j}}{\left(1/\sigma^2_{j,0}+(n+m)\sigma^2_j+m\beta^2_{j,n}/\sigma^4_j\right)^{1/2}}\right). \quad (4)$$

If the system is unclassified after n samples, the analyst finds the smallest number of samples m such that

$$\min_j \Phi\left(\frac{\frac{b_j-\mu_{j,0}}{\sigma^2_{j,0}}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma^2_j}}{\left(1/\sigma^2_{j,0}+(n+m)\sigma^2_j+m\beta^2_{j,n}/\sigma^4_j\right)^{1/2}}\right) \le \delta$$

or

$$\min_j \Phi\left(\frac{\frac{b_j-\mu_{j,0}}{\sigma^2_{j,0}}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma^2_j}}{\left(1/\sigma^2_{j,0}+(n+m)\sigma^2_j+m\beta^2_{j,n}/\sigma^4_j\right)^{1/2}}\right) \ge 1-\delta/c.$$

The interpretation is that m is the number of samples required until the system is classified, on average. A greedy algorithm selects the system with the smallest m to sample from. Starting from somewhere in $(\delta, 1-\delta/c)$, the expression inside the parenthesis of (4) is of order $\sqrt{m}$, and thus goes to $\pm\infty$ as m gets large. Hence, it must hit $\Phi^{-1}(1-\delta/c)$ or $\Phi^{-1}(\delta)$ for some $m > 0$. Solving for m leads to a quadratic equation of the form

$$\underbrace{\left(\frac{b_j-\alpha_{j,n}}{\sigma^2_j}\right)^2}_{A_j} m^2 + \underbrace{\left[2\left(\frac{b_j-\nu_{j,n}}{\sigma^2_j}\right)\left(\frac{n(b_j-\bar{Y}_j(n))}{\sigma^2_j}+\frac{b_j-\mu_{j,0}}{\sigma^2_{j,0}}\right)-\Phi^{-1}(\cdot)^2\times(\sigma^2_j+\beta^2_{j,n}/\sigma^4_j)\right]}_{B_j(\cdot)} m + \underbrace{\left(\frac{n(b_j-\bar{Y}_j(n))}{\sigma^2_j}+\frac{b_j-\mu_{j,0}}{\sigma^2_{j,0}}\right)^2-\Phi^{-1}(\cdot)^2\left(\frac{1}{\sigma^2_{j,0}}+n\sigma^2_j\right)}_{C_j(\cdot)} = 0, \quad (5)$$

where $\cdot$ equals δ or $1-\delta/c$. Let $m_j(\cdot)$ be the smallest positive root of (5), with $m_j(\cdot)=\infty$ if the root is not real (i.e., $B^2_j(\cdot) < 4A_jC_j(\cdot)$). Hence, $m_j(\delta)$ is the number of additional samples required for constraint j to exit through δ, and $m_j(1-\delta/c)$ is the number of extra samples required for constraint j to exit through $1-\delta/c$. Let

$$m = \begin{cases} \min_j\{m_j(\delta)\} & \text{if } m_j(\delta) < m_j(1-\delta/c) \le \infty \text{ for some } j, \\ \max_j\{m_j(1-\delta/c)\} & \text{otherwise.} \end{cases}$$

Thus, in the top case m is the number of additional samples required to have $E_{Y_{j,n+1},\ldots,Y_{j,n+m}|Y_{j,1},\ldots,Y_{j,n}}[P(\mu_j \le b_j|Y_{j,1},\ldots,Y_{j,n},Y_{j,n+1},\ldots,Y_{j,n+m})] = \delta$ for at least one constraint j, while in the bottom case m is the number of extra samples to have $E_{Y_{j,n+1},\ldots,Y_{j,n+m}|Y_{j,1},\ldots,Y_{j,n}}[P(\mu_j \le b_j|Y_{j,1},\ldots,Y_{j,n},Y_{j,n+1},\ldots,Y_{j,n+m})] = 1-\delta/c$ for all constraints j.

4 NUMERICAL ILLUSTRATION I

We illustrate the first algorithm by determining whether the ten systems whose parameter values are given in Table 1 are feasible, i.e., we wish to determine whether $\mu_j \leq \gamma$ for $j = 1,2,\ldots,10$. μ_j is the (unknown) mean system performance estimated through simulation and $\gamma = 0$ is the threshold of the feasible region. We assume that $\mu_j \sim N(\mu_{j,0}, \sigma^2_{j,0})$, where $\mu_{j,0}$ and $\sigma^2_{j,0}$ are the prior values of the system mean and variance, respectively. We further assume that the true values of σ^2_j are known.

Figure 1 shows the results of 30 independent experiments where the x-axis shows the values of the prior as a percentage of the true mean, μ_j. In particular, a value of 1 implies that the prior is exactly equal to the true mean while a value of -1 implies that, while the value of the prior is equal to the true mean, its sign is wrong. The y-axis reflects the proportion of the time the systems are classified correctly in our experiments. As reflected by the figure, the algorithm's ability to correctly classify the systems is largely driven by the value of the prior for μ_j. When the prior is close to the true value, the performance of the algorithm is very good. On the other hand, when the prior is far from the true mean, its performance is quite poor. This is indeed a well-known challenge (Shen and Wasserman (2001)) with various proposed solutions (Kass and Wasserman (1996)).

Table 1: Parameter values of the systems.

System, j	μ_j	σ^2_j	$\mu_{j,0}$	$\sigma^2_{j,0}$
1	1	1	0.5	10
2	-1	1	-0.5	10
3	2	1	1.5	10
4	-2	1	-1.5	10
5	3	1	2.5	10
6	-3	1	-2.5	10
7	5	1	5.5	10
8	-5	1	-5.5	10
9	10	1	9.0	10
10	-10	1	-9.0	10

5 STEP 3: NUMBER OF SAMPLES TO CLASSIFY A SYSTEM WITH PROBABILITY $1-\eta$

The approach of Section 3 finds the number of samples required so that the posterior probability that the system mean lies inside or outside of Γ is at least $1-\delta$, on average. An alternative approach is to determine the number of samples so that the posterior probability that the system mean lies inside or outside of Γ is at least $1-\delta$ with (say) probability $1-\eta$, for $0 < \eta < 1$. That is, starting from

$$\delta < P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n}) < 1 - \delta/c,$$

the problem is to find the number of additional samples $m \geq 1$ such that

$$P_{Y_{j,n+1},\ldots,Y_{j,n+m}|Y_{j,1},\ldots,Y_{j,n}}((\min_j P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+m}) \leq \delta) \cup (\min_j P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+m}) \geq 1 - \delta/c)) \geq 1 - \eta, \quad (6)$$

so that a system is classified with probability $1-\eta$ after drawing m new samples. As before, if after drawing the m new samples $\min_j P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+m}) \leq \delta$ then the system is declared not in Γ, and if $\min_j P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+m}) \geq 1 - \delta/c$ then the system is declared in Γ.

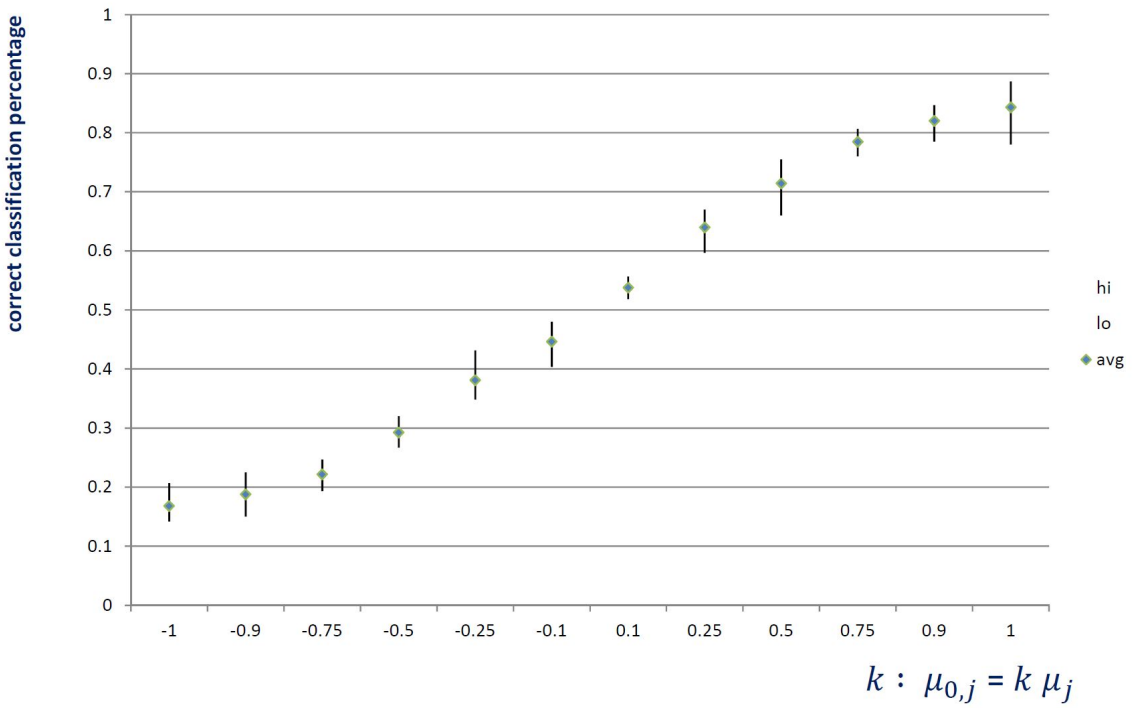

Figure 1: Correct Classification of Alternative Systems.

To wit, from (2) we have $P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+1}) \leq \delta$ if and only if

$$\Phi\left(\left(\frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n} - \beta_{j,n}Z}{\sigma^2_j}\right)\frac{1}{(1/\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2}}\right) \leq \delta$$

$$\iff \frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n} - \beta_{j,n}Z}{\sigma^2_j} \leq \Phi^{-1}(\delta)(\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2},$$

where $Z \sim N(0,1)$, and $\alpha_{j,n}$ and $\beta_{j,n}$ are as in (1). Likewise, $P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+1}) \geq 1 - \delta/c$ if and only if

$$\frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n} - \beta_{j,n}Z}{\sigma^2_j} \geq \Phi^{-1}(1-\delta/c)(\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2}.$$

Hence

$$P_{Y_{j,n+1}|Y_{j,1},\ldots,Y_{j,n}}(\min_j(P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+1}) \leq \delta) \cup (\min_j P(\mu_j < b_j | Y_{j,1}, \ldots, Y_{j,n+1}) \geq 1 - \delta/c))$$

$$= 1 - \Phi\left(\min_j \frac{\sigma^2_j}{\beta_{j,n}}\left(-\Phi^{-1}(\delta)(\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2} + \frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n}}{\sigma^2_j}\right)\right)$$

$$+ \Phi\left(\min_j \frac{\sigma^2_j}{\beta_{j,n}}\left(-\Phi^{-1}(1-\delta/c)(\sigma^2_{j,0} + (n+1)/\sigma^2_j)^{1/2} + \frac{b_j - \mu_{j,0}}{\sigma^2_{j,0}} + \frac{(n+1)b_j - n\bar{Y}_j(n) - \alpha_{j,n}}{\sigma^2_j}\right)\right). \quad (7)$$

Using (7), Eq. (6) is the same as

$$\Phi\left(\min_j \frac{\sigma_j^2}{m\beta_{j,n}}\left(-\Phi^{-1}(\delta)(\sigma_{j,0}^2+(n+m)/\sigma_j^2)^{1/2}+\frac{b_j-\mu_{j,0}}{\sigma_{j,0}^2}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma_j^2}\right)\right)$$
$$-\Phi\left(\min_j \frac{\sigma_j^2}{m\beta_{j,n}}\left(-\Phi^{-1}(1-\delta/c)(\sigma_{j,0}^2+(n+m)/\sigma_j^2)^{1/2}+\frac{b_j-\mu_{j,0}}{\sigma_{j,0}^2}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma_j^2}\right)\right)$$
$$\leq \eta. \quad (8)$$

The left-hand side in (8) approaches zero as m increases because the dominant term inside each $\Phi(\cdot)$ is $\min_j(b_j-\alpha_{j,n})/\beta_{j,n}$. Hence, there is a positive solution of

$$\Phi\left(\min_j \frac{\sigma_j^2}{m\beta_{j,n}}\left(\Phi^{-1}(1-\delta)(\sigma_{j,0}^2+(n+m)/\sigma_j^2)^{1/2}+\frac{b_j-\mu_{j,0}}{\sigma_{j,0}^2}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma_j^2}\right)\right)$$
$$-\Phi\left(\min_j \frac{\sigma_j^2}{m\beta_{j,n}}\left(\Phi^{-1}(\delta/c)(\sigma_{j,0}^2+(n+m)/\sigma_j^2)^{1/2}+\frac{b_j-\mu_{j,0}}{\sigma_{j,0}^2}+\frac{(n+m)b_j-n\bar{Y}_j(n)-m\alpha_{j,n}}{\sigma_j^2}\right)\right)=\eta,$$

which can be found numerically for m.

6 NUMERICAL ILLUSTRATION II

We illustrate the second algorithm by determining whether the ten systems whose parameter values are given in Table 2 are feasible, i.e., we wish to determine whether $\mu_j \leq \gamma$ for $j = 1,2,\ldots,10$. $\mu_j \in R^2$ is the (unknown) mean system performance estimated through simulation and $\gamma = (0,0)$ is the threshold of the feasible region. We assume that $\mu_j \sim N(\mu_{j,0}, \sigma_{j,0}^2)$, where $\mu_{j,0}$ and $\sigma_{j,0}^2$ are the prior values of the system mean and variance, respectively.

In this case, the algorithm's ability to correctly classify the systems is largely driven by the value of the prior for σ_j. Figure 2 shows the results of 100 independent experiments. Panels (a) and (b) show that the required sample size to achieve a desired probability of correct classification ($\eta = 0.90$) increases as the system variance increases. Panel (c) further shows that as the prior values for the system variance increase, it becomes increasingly costly, in terms of the required number of samples, to achieve the desired probability of correct classification.

Table 2: Parameter values of the systems.

System, j	$\mu_{j,0}$	σ_j^2	$\sigma_{j,0}^2$
1	-0.11,-0.12	10,20	3,4
2	-0.23,-0.24	10,20	3,4
3	-0.35,-0.35	10,20	3,4
4	0.41,0.42	10,20	3,4
5	0.50,0.50	10,20	3,4
6	-1.00,-1.00	10,20	3,4
7	1.00,1.00	10,20	3,4
8	-2.00,-2.00	10,20	3,4
9	-0.75,-0.75	10,20	3,4
10	0.75,0.85	10,20	3,4

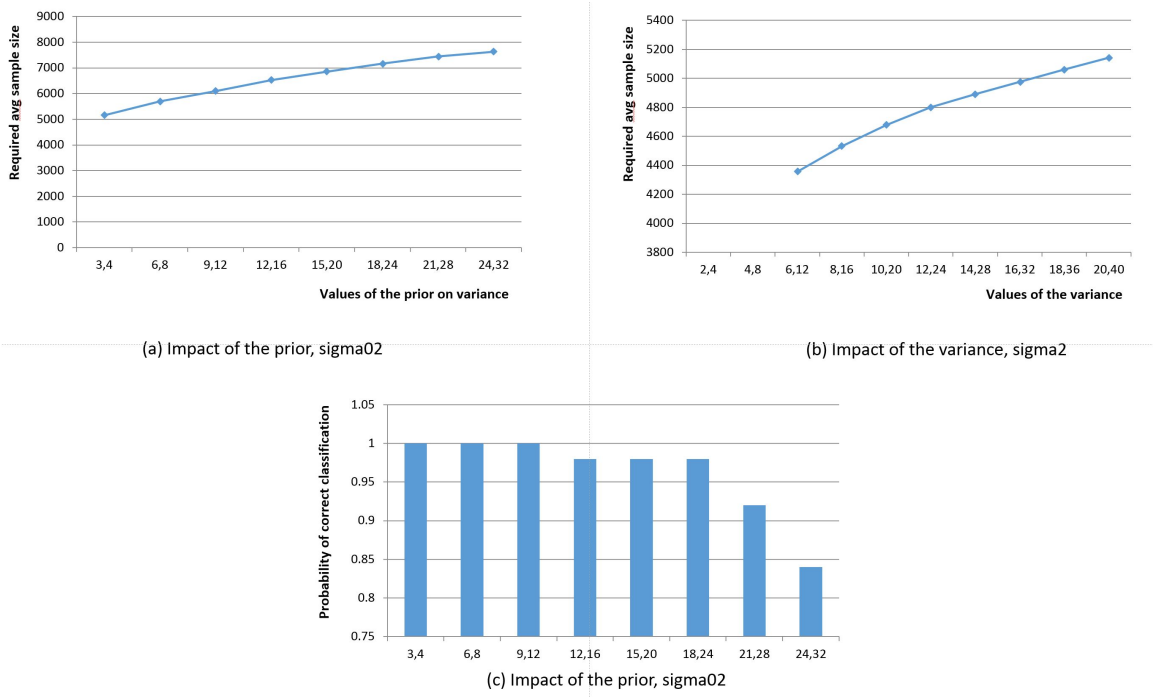

Figure 2: Probability of Correct Classification of Alternative Systems.

7 CONCLUDING COMMENTS

We consider a Bayesian approach to determine whether a system belongs to a given set $\Gamma \subseteq R^c$ based on performance measures estimated through Monte Carlo simulation. In particular, Γ is assumed to be defined by linear constraints that form an orthant. Within this setting, we introduce two adaptive procedures. The first approach determines the number of additional samples required so that the posterior probability that a system's mean performance is correctly classified to be inside or outside of Γ is at least 1-δ, in expectation. The second method determines the number of additional samples so that the posterior probability that the system mean lies inside or outside of Γ is at least $1-\delta$ with probability $1-\eta$, for $0 < \eta < 1$. In either approach, the system that requires the fewest additional samples for classification is selected for sampling. Preliminary numerical experiments show that the algorithms do not converge rapidly if the prior mean is far away from the true (unknown) mean. In other words, the number of iterations needed to get a certain probability of correct classification grows with the distance of the prior sample to the prior mean. Current work is focused on formally characterizing this growth rate, and at extending the proposed approaches to settings with unknown variances.

REFERENCES

Andradottir, S., D. Goldsman, and S.-H. Kim. 2005. "Finding the best in the presence of stochastic constraints". In *Proceedings of the 2005 Winter Simulation Conference*, edited by M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 732–738. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Andradottir, S., and S. H. Kim. 2010. "Fully sequential procedures for comparing constrained systems via simulation". *Naval Research Logistics* 57:403–421.

Batur, D., and S. H. Kim. 2010. "Finding feasible systems in the presence of constraints on multiple performance measures". *ACM Transactions on Modeling and Computer Simulation* 20:118–130.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2013. *Bayesian Data Analysis*. 3rd ed. Boca Raton, FL: CRC Press.

Hunter, S. R., and R. Pasupathy. 2012. "Optimal sampling laws for stochastically constrained simulation optimization on finite sets". *INFORMS Journal on Computing*:527 – 542.

Kass, R. E., and L. Wasserman. 1996. "The selection of prior distributions by formal rules". *Journal of the American Statistical Association* 91:1343 – 1370.

Lee, L. H., N. A. Pujowidianto, C. H. Chen, and C. M. Yap. 2012. "Asymptotic simulation budget allocation for selecting the best design in the presence of stochastic constraints". *IEEE Transactions on Automatic Control* 57:2940–2945.

Pasupathy, R., S. R. Hunter, N. A. Pujowidianto, L. H. Lee, and C. H. Chen. 2014. "Stochastically constrained ranking and selection via SCORE". *ACM Transactions on Modeling and Computer Simulation* 25:1 – 26.

Powell, W. B., and I. O. Ryzhov. 2012. *Optimal Learning*. Hoboken, NJ: Wiley.

Santner, T. J., and A. C. Tamhane. 1984. "Designing experiments for selecting a normal population with a large mean and a small variance". In *Design of Experiments - Ranking and Selection: Essays in Honor of Robert E. Bechhoffer*, edited by T. J. Santner and A. C. Tamhane, 179–198. New York: Marcel-Dekker.

Shen, X., and L. Wasserman. 2001. "Rates of convergence of posterior distributions". *The Annals of Statistics* 29:687 – 714.

Szechtman, R., and E. Yücesan. 2008. "A new perspective on feasibility determination". In *Proceedings of the Winter Simulation Conference*, edited by R. G. Ingalls, M. D. Rossetti, J. S. Smith, and B. A. Peters, 192–199. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

AUTHOR BIOGRAPHIES

ROBERTO SZECHTMAN received his Ph.D. from Stanford University, and currently is Associate Professor in the Operations Research Department at the Naval Postgraduate School. His research interests include applied probability and military operations research. His email address is rszechtm@nps.edu.

ENVER YÜCESAN is a Professor of Operations Management at INSEAD. He holds a Ph.D. in Operations Research from Cornell University. His research interests include stochastic simulation optimization, with applications in supply chain and risk management. His email address is enver.yucesan@insead.edu.