

Large Deviations Perspective on Ordinal Optimization of Heavy-tailed Systems

Jose Blanchet
Department of Industrial Engineering
and Operations Research
Columbia University
New York, NY 10027

Jingchen Liu
Department of Statistics
Columbia University
New York, NY 10027

Bert Zwart
H. Milton Stewart School of Industrial and Systems Engineering
Georgia Institute of Technology
765 Ferst Drive
Atlanta, GA 30332, USA

ABSTRACT

We consider the problem of selecting the best among several heavy-tailed systems from a large deviations perspective. In contrast to the light-tailed setting studied by Glynn and Juneja (2004), in the heavy-tailed setting, the probability of false selection is characterized by a rate function that does not require as detailed information about the probability distributions of the system's performance. This motivates the question of studying static policies that could potentially provide convenient implementations in heavy-tailed settings. We concentrate on studying sharp large deviations estimates for the probability of false detection which suggest precise optimal allocation policies when the systems have comparable heavy-tails. Additional optimality insights are given for systems with non-comparable tails.

1 INTRODUCTION

We assume that there are d systems, each of which can be independently simulated and for each of the simulations we can evaluate the performance of the corresponding system and obtain a score. The value of the score for the j -th replication of the i -th system is denoted as $X_{i,j}$ and the mean score is denoted by $\mu_i = EX_{i,j}$, which is assumed to be unknown.

Given a computer budget n (in terms of total number of replications), we are interested in studying the problem of optimizing the allocation of the computer budget across the different systems in order to select, via simulation, the system with the smallest mean score in such a way that the probability of false selection (PFS) is minimized. Our focus is on the analysis of heavy-tailed systems – in particular, we will not assume that the $X_{i,j}$'s possess a finite moment

generating function. Our approach involves the use of large deviations theory for heavy-tailed random walks; therefore, our allocation policies are shown to be asymptotically optimal as the computational budget, n , increases.

The problem of maximizing the probability of correct selection in ordinal optimization has been studied substantially in the literature. A popular approach is based on the so-called indifferent zone formulation (see, for instance, Goldsman and Nelson (2001) and Kim and Nelson (2003) for an overview). For instance, in the case of two systems, one assumes a known lower bound in the difference between the mean scores of the systems. One then tries to allocate the computational budget in order to control the probability of false selection in terms of such a bound. In turn, the analysis of such allocation policies is based on Gaussian assumptions, which are justified by means of the Central Limit Theorem (CLT) (see, for instance, Chen et al (2000) for allocation analysis under Gaussian assumptions). These types of assumptions may be reasonable if the indifference zone is relatively small. More precisely, the use of the CLT is appropriate if the size of the indifference zone (potentially after appropriate scaling relative to the largest mean score) is of order $O(n^{-1/2})$ where n is the total number of simulation runs. When the indifference zone is relatively large (say of order $O(1)$ relative to the number of simulation runs) a large deviations approach seems more appropriate. Our approach here follows the same spirit as that proposed by Glynn and Juneja (2004), which take a large deviations approach to ordinal optimization in the context of light-tailed systems. Broadie, Han and Zeevi (2007) have also taken a similar perspective to that Glynn and Juneja in the setting of heavy-tailed systems but our results here strengthen theirs and also, as we shall discuss, provide a somewhat different interpretation of the optimal allocation policy.

In order to discuss our main results and to put them in perspective, let us start by summarizing some key ideas behind the analysis of Glynn and Juneja (2004). In the light-tailed setting, the most likely way in which false detection occurs is dictated by the so-called rate functions associated to large deviations probabilities for empirical means (there is one such rate for each of the empirical means corresponding to each of the d systems). In particular, the i -th rate function (for $1 \leq i \leq d$) is computed as the Legendre transform of the associated cumulant generating function of $X_{i,j}$. The rate of decay for the probability of false detection is computed combining the d rate functions in a concave optimization problem over a region that represents false detection. It turns out that, in principle, one can recover the cumulant generating functions from the rate functions (by applying once again Legendre transforms). Therefore, since the rate of decay for the probability of false detection requires knowledge of the rate functions, it follows that a static allocation policy would require knowledge of the whole distribution (via the cumulant generating function) and therefore such a policy is typically difficult to implement.

Another feature implicit in the analysis of Glynn and Juneja (2004) is the fact that, in its most likely scenario, false detection occurs due to atypical behavior of basically all of the $X_{i,j}$'s. In other words, the most likely way in which false detection occurs is given by "cooperation" among all the different sampled scores in a way that is described (asymptotically) by the solution to the concave optimization problem mentioned in the previous paragraph.

Our analysis in this paper, in the context of heavy-tailed $X_{i,j}$'s (which we shall assume to have essentially power-law type tails), provides different qualitative insights relative to the light-tailed situation. First, the large deviations rate function for the empirical means of each of the d systems is basically polynomial and does not characterize the distribution of the increments. In fact, not even the mean of the μ_i 's can be recovered if one has perfect knowledge of such rate functions. Second, false selection occurs as a consequence of at most d atypical outcomes (in contrast to cooperation among basically n samples in the light-tailed case). In particular, the large deviations analysis of the probability of false detection has a very intuitive interpretation. Namely, false detection can be decomposed in d scenarios, corresponding to false detection caused by large deviations behavior of the empirical estimate of μ_i for $1 \leq i \leq d$. The chance that false detection is caused by large deviations behavior involving two or more empirical estimates is negligible. When the associated large deviations rates for each of the d empirical means have comparable behavior as the number of samples grows, then the allocations are proportional as in the light-tailed setting and take a non-trivial form based on the solution to a concave program.

Heavy-tailed systems in the context of ordinal optimization have also been studied by Broadie, Han and Zeevi

(2007). Their analysis is different from ours in the following aspects. First, they concentrate on the analysis of the log-rate function (i.e. $\log PFS$), so the exact decay rate of the PFS is not discussed. Second, they focus only on proportional allocation rules, much in the spirit of the light-tailed setting. Based on these two items, the analysis of Broadie, Han and Zeevi (2007) suggests that the PFS is, in a rough sense, asymptotically independent of the selected policy.

In contrast to the analysis of Broadie, Han and Zeevi, we study the exact rate of decay of the PFS and we also allow arbitrary allocation policies. Our analysis suggests somewhat different interpretations. In particular, we conclude that allocation policies might have a significant impact, specially when several systems possess comparable tail behaviors. Moreover, proportional policies do not necessarily might have substantially different performance when it comes to the exact rate of decay for the PFS . In particular, we conclude that when the systems do not have comparable tail behavior, then one should allocate substantially more resources to the system that has the highest "risk" (i.e. heaviest tail) but one should still allocate a substantial amount of computational budget (described explicitly in terms of the large deviations rate functions) to each of the less risky systems.

The rest of the paper is organized as follows. Section 2 reviews the basic ideas in the analysis of Glynn and Juneja (2004) that gives rise to proportional allocation rules. In Section 3 we discuss some large deviations results that are used in our analysis. The rate of convergence analysis of the PFS is given in Section 4. Finally, Section 5 discusses a simulation experiment in which the various parameters required to implement the policy are also estimated.

2 ORDINAL OPTIMIZATION FOR LIGHT-TAILED SYSTEMS

Without loss of generality we assume that $\mu_1 < \mu_2 \leq \dots \leq \mu_d$. We introduce some additional notation. Let

$$\bar{X}_i(n_i) = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{i,j}$$

be the empirical mean of the scores of the i system given n_i replications. We assume that $\sum_{i=1}^d n_i = n$ and the goal is to select the n_i 's (as a function of n) in order to minimize the probability of false selection (PFS) which is given by

$$PFS = P \left(\bar{X}_1(n_1) > \min_{j \geq 2} \bar{X}_j(n_j) \right).$$

In the context of light-tailed systems studied by Glynn and Juneja (2004) one has

$$\log P(\bar{X}_i(n_i) \geq \varepsilon + \mu_j) = -n_i I_i(\mu_j + \varepsilon) + o(n_i I_i(\mu_j + \varepsilon))$$

as $n_i \nearrow \infty$ uniformly over $\varepsilon \geq 1/n_i^{1/2-\delta}$ for each $\delta > 0$, where $I_j(\cdot)$ is a positive and convex rate function with the property that $I_j(\mu_j + \varepsilon) \sim \varepsilon^2/(2\sigma_j^2)$ as $\varepsilon \searrow 0$ with $\sigma_j^2 = \text{Var}(X_i)$ (see, for instance, Dembo and Zeitouni (1998)). In the light-tailed setting, the rate functions $I_j(\cdot)$'s governs the exponential rate of decay of small probabilities and they can be used to provide asymptotics for PFS.

In particular, following the analysis of Glynn and Juneja (2004) we see that if one has $n_j \geq n^{1/2+\delta}$ for any fix $\delta > 0$, then

$$\begin{aligned} \log P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j)\right) \\ = -H(n_1, \dots, n_d) + o(H(n_1, \dots, n_d)) \end{aligned}$$

as $n \nearrow \infty$, where

$$H(n_1, \dots, n_d) = \min_{2 \leq j \leq d} \inf_x (n_1 I_1(\mu_1 + x) + n_j I_j(\mu_j + x)).$$

Now proportional allocation follows from the analysis of the previous rate function. In particular, it follows from the fact that the $I_j(\cdot)$'s are positive and $I_j(\mu_j + \varepsilon) \sim \varepsilon^2/(2\sigma_j^2)$ that in order to maximize the growth rate of $H(n_1, \dots, n_d)$ as $n \nearrow \infty$ as a function of (n_1, \dots, n_d) as long as the n_i 's and subject to the constraint $\sum_{j=1}^d n_j = 1$, one is forced to select $n_j = p_j n$ (where the p_j 's are strictly positive numbers, independent of n , that add up to 1). Glynn and Juneja provide the optimal selection of the p_j 's, which clearly depends on the fine structure of the I_j 's, which in turn depends on the whole distribution of the $X_{i,j}$'s through the moment generating function. In fact, the I_j 's are so intimately connected to the distribution of the $X_{i,j}$'s that, in principle, if one knows $I_j(\cdot)$, then one can basically recover the whole distribution of $X_{i,j}$'s via the associated moment generating function. This makes the asymptotic optimal policy of Glynn and Juneja difficult to implement in practice.

3 BASIC RESULTS FOR AVERAGES OF REGULARLY VARYING RANDOM WALKS

In contrast to the light-tailed setting, large deviations approximations for heavy-tailed systems often require less detailed information about the structure of the distribution. In particular, assuming that the $X_{i,j}$'s have a right tail that is regularly varying with index $\alpha_i > 0$ which means that $P(X_{i,j} > x) = L(x)x^{-\alpha_i}$, $L(tx)/L(x) \rightarrow 1$, as $x \rightarrow \infty$ for some

$\alpha_i > 0$ and every $t > 0$ we have that

$$P(\bar{X}_i(n_i) > (\mu_i + \varepsilon)) = n_i G_i(n_i \varepsilon) + o(n_i G_i(n_i \varepsilon)),$$

as $n_i \nearrow \infty$ (for fixed ε) where $G_i(\cdot)$ is any function such that $G_i(t) \sim P(X_{i,j} > t)$ as $t \nearrow \infty$. A completely analogous estimate holds for left tails. For extensions to other types of heavy-tailed random variables, see for instance, Nagaev (1979) and Rozovskii (1993). A review is given by Mikosch & Nagaev (1998), and a recent study is Denisov *et al.* (2008).

Note that if one knows an associated "rate function" (a function with the same asymptotic behavior as $G_i(\cdot)$) then the distribution of the increments cannot be recovered. In fact, not even the mean can be identified given precise knowledge of such a rate function. This observation suggests that a large deviations perspective to ordinal optimization in the context of heavy-tailed systems could provide some guideline that may be helpful in practical settings. This is one of the reasons that motivates our analysis in this paper.

An additional feature of typical heavy-tailed random variables is that the conditional overshoot over a large level increases, in some way, as the level in question increases. For instance, let X_1, \dots, X_n be iid mean zero random variables with regularly varying right tail. Now, consider $S_n = \sum_{i=1}^n X_i$. Consider the conditional overshoot of S_n given that $S_n > cn$ for some $c > 0$. We then have that for any $\delta > 0$,

$$\begin{aligned} P(S_n - cn > \delta n | S_n > cn) &\sim \frac{P(X_1 > (\delta + c)n)}{P(X_1 > cn)} \\ &\rightarrow P(Z_\alpha > \delta/c) = \frac{1}{(1 + \delta/c)^\alpha}. \end{aligned}$$

In other words, $Z_\alpha + 1$ has a Pareto distribution with tail index α . As a consequence, if the $X_{i,j}$'s are regularly varying with tail index, α_i , then for all ε_1 and $\varepsilon_2 > 0$,

$$P(\bar{X}_i(n_i) - \mu_i > \varepsilon_1 + \varepsilon_2 | \bar{X}_i(n_i) - \mu_i > \varepsilon_1) \rightarrow P(Z_\alpha > \varepsilon_2/\varepsilon_1)$$

as $n_i \rightarrow \infty$. The previous basic feature is basically all what we shall need for our development of sharp asymptotics for the probability of false detection in the next section.

4 ANALYSIS OF PROBABILITY OF FALSE DETECTION

The main result in this section is the following:

Theorem 1 *Let $EX_{i,j} = \mu_i$ and suppose that $-\infty < \mu_1 < \mu_2 \leq \dots \leq \mu_d$. In addition, suppose that*

$$P(X_{1,j} > t) = L_1(t)t^{-\alpha_1} + o(L_1(t)t^{-\alpha_1})$$

as $t \nearrow \infty$ and that

$$P(X_{1,j} \leq t) = L_j(|t|)|t|^{-\alpha_j} + o(L_j(|t|)|t|^{-\alpha_j})$$

as $t \searrow -\infty$. Then, assuming that $n_j \nearrow \infty$ for all $1 \leq j \leq d$, we obtain

$$\begin{aligned} PFS &\sim P(\bar{X}_1(n_1) > \mu_2) + \sum_{j=2}^d P(\bar{X}_j(n_j) < \mu_1) \\ &\sim (\mu_2 - \mu_1)^{-\alpha_1} n_1^{-\alpha_1+1} L_1(n_1) + \sum_{j=2}^d (\mu_j - \mu_1) n_j^{-\alpha_j+1} L_j(n_j). \end{aligned}$$

The proof is inspired by the methods discussed in Samorodnitsky (2002).

Proof. Suppose that

$$P(X_{i,j} > t) = L_i(t) t^{-\alpha_i} + o(L_i(t) t^{-\alpha_i})$$

as $t \nearrow \infty$. Define

$$\begin{aligned} A_1 &= \{\bar{X}_1(n_1) \geq \mu_1 + \varepsilon\}, \\ A_j &= \{\bar{X}_j(n_j) \leq \mu_j - \varepsilon\}, \end{aligned}$$

for $2 \leq j \leq d$. Then, select $\varepsilon \in (0, (\mu_2 - \mu_1)/2)$ so that

$$\begin{aligned} &P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j)\right) \\ &= P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j), A_1 \cup \dots \cup A_d\right) \\ &\leq \sum_{j=1}^d P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j), A_j\right). \end{aligned}$$

We have that $y \geq \varepsilon$

$$\begin{aligned} &P(\bar{X}_1(n_1) - \mu_1 > y | \bar{X}_1(n_1) - \mu_1 \geq \varepsilon) \\ &\rightarrow \left(\frac{y}{\varepsilon}\right)^{-\alpha_1} \end{aligned}$$

Therefore, using the law of large numbers and the bounded convergence theorem we obtain that

$$\begin{aligned} &P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j), A_1\right) \\ &= P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j) \middle| A_1\right) P(A_1) \\ &\sim \left(\frac{\mu_2 - \mu_1}{\varepsilon}\right)^{-\alpha_1} n_1^{-\alpha_1+1} \varepsilon^{-\alpha_1} L_1(n_1) \\ &= (\mu_2 - \mu_1)^{-\alpha_1} n_1^{-\alpha_1+1} L_1(n_1) \end{aligned}$$

as long as $\min_j n_j \nearrow \infty$. Similarly, we have that

$$\begin{aligned} &P(\bar{X}_1(n_1) - \mu_j > \bar{X}_j(n_j) - \mu_j | \bar{X}_j(n_j) - \mu_j \leq -\varepsilon) \\ &\rightarrow \left(\frac{\mu_j - \mu_1}{\varepsilon}\right)^{-\alpha_j}. \end{aligned}$$

and therefore, once again using dominated convergence, we obtain

$$\begin{aligned} &P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j), A_1\right) \\ &= P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j) \middle| A_j\right) P(A_j) \\ &\sim \left(\frac{\mu_j - \mu_1}{\varepsilon}\right)^{-\alpha_j} n_j^{-\alpha_j+1} \varepsilon^{-\alpha_j} L_j(n_j) = (\mu_j - \mu_1)^{-\alpha_j} n_j^{-\alpha_j+1} L_j(n_j). \end{aligned}$$

Combining the previous estimates, we obtain that an asymptotic upper bound for the *PFS* is given, as long as $n_j \nearrow \infty$ for all j simultaneously, by

$$(\mu_2 - \mu_1)^{-\alpha_1} n_1^{-\alpha_1+1} L_1(n_1) + \sum_{j=2}^d (\mu_j - \mu_1) n_j^{-\alpha_j+1} L_j(n_j). \quad (1)$$

For the lower bound we isolate the most important events. In particular, we define the events

$$\begin{aligned} B_1 &= \{\bar{X}_1(n_1) > \mu_2 + \varepsilon \geq \bar{X}_2(n_2)\}, \\ B_j &= \{\bar{X}_1(n_1) > \mu_1 - \varepsilon \geq \bar{X}_j(n_j)\}, \end{aligned}$$

for $2 \leq j \leq d$. Note that

$$P\left(\bar{X}_1(n_1) > \min_{2 \leq j \leq d} \bar{X}_j(n_j)\right) \geq P\left(\cup_{j=1}^d B_j\right).$$

The analysis of the probability $P\left(\cup_{j=1}^d B_j\right)$ is easily done using the inclusion-exclusion principle, in particular, we obtain

$$P\left(\cup_{j=1}^d B_j\right) \geq \sum_{j=1}^d P(B_j) - \sum_{1 \leq j_1 < j_2 \leq d} P(B_{j_1} \cap B_{j_2}).$$

Assuming $n_j \nearrow \infty$ as $n \nearrow \infty$ for all $1 \leq j \leq d$ we have that

$$\begin{aligned} P(B_1) &\sim P(\bar{X}_1(n_1) - \mu_1 > \mu_2 - \mu_1 + \varepsilon) \\ &\sim (\mu_2 - \mu_1 + \varepsilon)^{-\alpha_1} n_1^{-\alpha_1+1} L(n_1), \\ P(B_j) &\sim P(\bar{X}_j(n_1) - \mu_j \leq \mu_1 - \mu_j - \varepsilon), \\ &\sim (\mu_j - \mu_1 + \varepsilon)^{-\alpha_j} n_j^{-\alpha_j+1} L(n_j). \end{aligned}$$

On the other hand,

$$\begin{aligned} P(B_1 \cap B_{j_2}) &\leq P(\bar{X}_1(n_1) > \mu_2 + \varepsilon) P(\mu_1 - \varepsilon > \bar{X}_j(n_j)) \\ &= O(P(B_1)P(B_{j_2})). \end{aligned}$$

Similarly, we obtain that

$$P(B_{j_1} \cap B_{j_2}) \leq O(P(B_{j_1})P(B_{j_2})).$$

As a consequence, since $\varepsilon > 0$ is arbitrary, we conclude that an asymptotic lower bound for PFS is also given by (1). ■

In the case in which the right tail of $X_{1,j}$ is comparable to all of the left tails of the $X_{i,j}$'s for $1 \leq i \leq d$, then it follows directly from Theorem 1 that it is optimal to have proportional allocations. These allocations are the heavy-tailed analogue of those given by Glynn and Juneja (2004). The next result provides the precise form of such allocations.

Corollary 2 *Suppose that $P(X_{1,j} \leq t) \sim c_1 L(t) t^{-\alpha}$ as $t \nearrow \infty$ and that $P(X_{i,j} \leq t) \sim c_i L(|t|) |t|^{-\alpha}$ as $t \searrow -\infty$ then, the optimal allocation policy takes the form $n_j = np_j$ where the p_j 's are obtained by solving the problem*

$$\min c_1 (\mu_2 - \mu_1) p_1^{-\alpha+1} + \sum_{j=2}^d c_j (\mu_j - \mu_1) p_j^{-\alpha+1}$$

subject to $\sum_{j=1}^d p_j = 1$ and $p_j \geq 0$. In particular, we have that

$$p_j = \frac{\beta_j^{1/\alpha}}{\sum_{j=1}^d \beta_j^{1/\alpha}},$$

where $\beta_1 = c_1 (\mu_2 - \mu_1)$ and $\beta_j = c_j (\mu_j - \mu_1)$.

In the case in which the tails do not have comparable growth, then it is efficient to allocate most of the resources into systems that have the heaviest tails. However, still, large enough samples should be allocated to all systems. This is the content of the next result. For simplicity, we shall assume that the slowly varying component is common to all systems.

Corollary 3 *Suppose that $P(X_{1,j} \leq t) \sim c_1 L(t) t^{-\alpha_i}$ as $t \nearrow \infty$ and that $P(X_{i,j} \leq t) \sim c_i L(|t|) |t|^{-\alpha_i}$ as $t \searrow -\infty$. Let (i_1, \dots, i_d) be a set of indexes such that*

$$\alpha_{i_1} \leq \alpha_{i_2} \leq \dots \leq \alpha_{i_d}.$$

Suppose that $\alpha_{i_1} = \alpha_{i_2} = \dots = \alpha_{i_{k^*}}$ for some $k^* \leq d$. Then, the decay rate of PFS is maximized by selecting $n_{i_j} = np_{i_j}^*$ for

$$p_{i_j}^* = \frac{\beta_{i_j}^{1/\alpha_{i_1}}}{\sum_{j=1}^{k^*} \beta_{i_j}^{1/\alpha_{i_1}}},$$

where $\beta_{i_j} = c_1 (\mu_2 - \mu_1)$ if $i_j = 1$ and $\beta_{i_j} = c_{i_j} (\mu_{i_j} - \mu_1)$ assuming $i_j \neq 1$. Moreover, for $k^* < j \leq d$ we must have that $n_{i_j} \nearrow \infty$ as $n \nearrow \infty$ in such a way that $n^{\alpha_{i_1}/\alpha_{i_j}} = o(n_{i_j})$.

5 SIMULATION EXPERIMENT

In this section, we carry out the implementation and assume different systems have the same but unknown tail index, α , same slowly varying function, $L(\cdot)$, and different c_i 's. Our strategy is to spend a small portion of the samples to estimate α and c_i . Then, estimate the optimal allocation using the estimated parameters according to the solution in Corollary 2. As an illustration of our analysis, we performed a simulation study of two systems in the following example.

Example 4 *Suppose we have two systems. One is t distribution with three degrees of freedom and expectation zero. The other one is t distribution with three degrees of freedom, scale factor two, and expectation one. That is,*

$$X_{1,j} \stackrel{d}{=} T_3, \text{ and } X_{2,j} \stackrel{d}{=} 1 + 2T_3,$$

where T_3 follows standard t distribution with three degrees of freedom. We use 20% samples to estimate the tail index, α , (with the knowledge that both distributions have the same tail index), and the constants, c_i . We used Hill's estimator to estimate tail index; see, for instance, Embrechts et al (1997). In particular, let $n_0 = \lfloor n/10 \rfloor$, and $X_{i,(1)}, \dots, X_{i,(n_0)}$ be the ordered statistics of $X_{i,1}, \dots, X_{i,n_0}$ ($X_{i,(1)}$ is the largest value). We let

$$\hat{\alpha} = \left[\frac{1}{2n_0} \sum_{i=1}^2 \sum_{j=1}^{\lfloor \sqrt{n_0} \rfloor} \log X_{i,(j)} - \log X_{i,(\lfloor \sqrt{n_0} \rfloor + 1)} \right]^{-1}$$

be the estimator of α and

$$\hat{c}_i = \exp \left[\frac{1}{1 + \lfloor \sqrt{n_0} \rfloor - \lfloor \sqrt{n_0} \rfloor / 2} \sum_{i=n - \lfloor \sqrt{n_0} \rfloor}^{n - \lfloor \sqrt{n_0} \rfloor / 2} \log \left(\frac{\lfloor \sqrt{n_0} \rfloor}{n_0} X_{i,(\lfloor \sqrt{n_0} \rfloor)} \right) \right]$$

be the estimator of the constant, c_i . We allocate the rest 80% samples according to Corollary 2. The false selection rates and the asymptotic optimal false selection rates according to Theorem 1 are presented in Table 1.

REFERENCES

- M. Brodie, M. Han, A. Zeevi. Implications of heavy tails on Simulation-based ordinal optimization. In: *Proceedings of the 2007 WSC*.
 N. Bingham, C. Goldie, J. Teugels (1987). *Regular Variation*.
 Chen, C. H., Lin, J., Yucesan, E., and Chick, S. (2000). Simulation Budget Allocation for Further Enhancing the

n	Optimal	FS	Std. Error	Simulation
500	1.19E-04	2.06E-04	1.44E-05	1.00E+06
1000	2.98E-05	3.45E-05	4.15E-06	2.00E+06
2000	7.44E-06	5.00E-05	1.58E-06	2.00E+06
5000	1.19E-06	1.80E-06	4.24E-07	1.00E+07

Table 1: Simulation. “Optimal” - asymptotic false selection rate according to Theorem 1; “FS” - estimated false selection probability; “Std. Error” - estimated standard error of “FS”; “Simulation” - number of replications to produce “FS” and “Std. Error”.

Efficiency of Ordinal Optimization. *Journal of Discrete Event Dynamic Systems: Theory and Applications*, 10, 251-270.

Dai, L. (1996) Convergence Properties of Ordinal Comparison in the Simulation of Discrete Event Dynamic systems. *Journal of Optimization Theory and Applications*, 91, 263-388.

Dembo, A., and Zeitouni, O. (1998) *Large deviations techniques and applications*. Springer-Verlag. New York.

Denisov, D., Dieker, T., Shneer, V. (2008). Large deviations for random walks with subexponential increments. To appear in *Annals of Probability*.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag. New York.

Glynn, P. and Juneja, S. (2004) A Large Deviations Perspective for Ordinal Optimization. *Proc. of the 2004 Winter Simulation Conference*, Eds. B.A. Peters, J.S. Smith, D.J. Madeiros, and M.W. Rohrer. 139-146. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Goldsman, D., and Nelson, B. (2001) Statistical Selection of the Best System. *Proceedings of the 2001 Winter Simulation Conference*, Eds. B. A. Peters, J. S. Smith, D. J. Madeiros, and M. W. Rohrer. 139-146. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers.

Kim, S. H., and Nelson, B. (2003) Selecting the Best system: Theory and Methods. *Proceedings of the 2003 Winter Simulation Conference*, Ed. S. Chick, P. J. Sanchez, D. Ferrin, and D. J. Morrice, 101-112.

Mikosch, T., Nagaev, A.V. (1998). Large Deviations of Heavy-Tailed Sums with Applications in Insurance. *Extremes*, 1, 81-110.

Nagaev, S. (1979) Large Deviations of Sums of Independent Random Variables. *Ann. of Probab.*, 7, 745-789.

Rozovskii, L. V. (1993) Probabilities of large deviations on the whole axis. *Theory Probab. Appl.*, 38, 53-79.

Probabilities of Large Deviations of Sums of Independent Random Variables with Common Distribution Function

in the Domain of Attraction of the Normal Law, *Theory Probab. Appl.*, 34, 625-644.

Samorodnitsky, G. (2002) Lectures on Heavy-tailed Large Deviations. <http://www.math.ku.dk/mikosch/maphysto/gena1/gena4.pdf>

AUTHOR BIOGRAPHIES

Jose Blanchet is an Assistant Professor in the Department of Industrial Engineering and Operations Research at Columbia University. Jose holds a M.Sc. in Engineering-Economic Systems and Operations Research and a Ph.D. in Management Science and Engineering, both from Stanford University. He also holds two B.Sc. degrees: one in Actuarial Science and another one in Applied Mathematics from ITAM (Mexico). Jose worked for two years as an analyst in Protego Financial Advisors, a leading investment bank in Mexico. He has research interests in applied probability, computational finance, performance engineering, queueing theory, risk management, rare-event analysis, statistical inference, stochastic modelling, and simulation.

Jingchen Liu is an Assistant Professor in the Department of Statistics at Columbia University. Jingchen received his Ph.D. in Statistics from Harvard University in June, 2008.

Bert Zwart is Coca Cola Associate Professor of Industrial and Systems Engineering at Georgia Tech. Bert holds an M.A. in Econometrics from the Free University, Amsterdam, and a Ph.D in Applied Mathematics from Eindhoven University of Technology. He serves on the editorial boards of *Operations Research*, *Mathematics of Operations Research*, *Mathematical Methods of Operations Research*, and *Queueing Systems*. His research is concerned with modeling, analysis and simulation of stochastic systems arising in actuarial and financial mathematics, computer and communication systems, manufacturing systems, and service systems.