Working Paper

Asymptotically optimal allocation of simulation experiments in discrete stochastic optimization

Andreas Futschik
Georg Ch. Pflug

WP-96-020
February 1996

IIASA International Institute for Applied Systems Analysis ◻ A-2361 Laxenburg ◻ Austria
Telephone: +43 2236 807 ◻ Fax: +43 2236 71313 ◻ E-Mail: info@iiasa.ac.at

Asymptotically optimal allocation of simulation experiments in discrete stochastic optimization

Andreas Futschik
Georg Ch. Pflug

WP-96-020
February 1996

International Institute for Applied Systems Analysis ◻ A-2361 Laxenburg ◻ Austria
Telephone: +43 2236 807 ◻ Fax: +43 2236 71313 ◻ E-Mail: info@iiasa.ac.at

Abstract

Approximate solutions for discrete stochastic optimization problems are often obtained via simulation. It is reasonable to complement these solutions by confidence regions for the argmin-set. We address the question, how a certain total number of random draws should be distributed among the set of alternatives. We propose a one-step allocation rule which turns out to be asymptotically optimal in the case of normal errors for two goals: To minimize the costs caused by using only an approximate solution and to minimize the expected size of the confidence sets.

Key words: Discrete Stochastic Optimization, Simulation, Sampling Strategy, Large Deviations.

Asymptotically optimal allocation of simulation experiments in discrete stochastic optimization

*Andreas Futschik**
Georg Ch. Pflug

1 Introduction

Suppose we have to find the optimal decision i^* out of a finite set $S := \{1, \ldots, k\}$ of possible alternatives for a decision problem under uncertainty. Let ξ_i be the random variable modeling the uncertainty. The probability law of ξ_i may depend on the decision i. If $f(i, \xi_i)$ measures the result of decision i and the random outcome ξ_i, the discrete stochastic optimization problem is given as

$$\left\| \begin{array}{l} \text{Minimize } F_i = \mathbf{E}(f(i, \xi_i)) \\ \text{for } \ i \in S. \end{array} \right. \tag{1}$$

Subsequently we will write $\boldsymbol{F}$ for the vector $(F_i)_{i=1}^k$, F^* instead of $\min_{i \in S} F_i$ and i^* for the argmin of F (assuming that it is unique).

If $\boldsymbol{F}$ can be evaluated easily, problem (1) is a discrete optimization problem and techniques like Branch–and–Bound or Simulated Annealing can be applied. If on the other hand the exact evaluation of F is impossible, one has to use Monte Carlo sampling. We observe m_i i.i.d. replicates $\xi_{i,k}$ of the random variables ξ_i and approximate the problem (1) by the empirical problem

$$\left\| \begin{array}{l} \text{Minimize } \hat{F}_i^{(n)} = \frac{1}{m_i} \sum_{j=1}^{m_i} f(i, \xi_{i,j}) \\ \text{for } \ i \in S \end{array} \right. \tag{2}$$

where n is the total sample size, $n = \sum_{i \in S} m_i$.

When all m_i are sufficiently large the solutions of (2) will provide reasonable (approximate) solutions for (1). Indeed, by the law of large numbers, the solutions $\hat{X}^{(n)}$ of (2) satisfy

$$P[\hat{X}^{(n)} = i^*] \to 1,$$

provided that $m_i \to \infty$.

However, the above fact does not tell anything about the quality of the solutions $\hat{X}^{(n)}$ for small sample sizes and a restricted time budget. As pointed out by Ho, Sreenivas and

*Department of Statistics, University of Vienna

Vakili (1992) there are also many situations where the set S is very large and we cannot expect the solutions $\hat{X}^{(n)}$ to be optimal for (1).

Therefore it seems important to carry out the simulations cleverly: Observations should be allocated in a way that provides as much information as possible for the identification of the minimal point. This goal (which will be made more precise later) is quite different from the objective pursued in multi–armed bandit problems, namely to minimize the expected number of observations taken from non-optimal points. The objective of multi–armed bandit problems makes sense in biostatistical applications, where each alternative corresponds to a drug and an observation to an application of this drug to a patient. The same goal occurs when choosing among gambling machines the one with the highest expected outcome. Lai and Robbins (1985) construct asymptotically efficient strategies for this type of problem. Further information concerning optimal allocation rules based on index policies may be found e.g. in Gittins (1989). In our situation, however, observations correspond to computer simulations and will typically cause the same costs for all alternatives. So there is no model-inherent reason why sampling from non-optimal populations should be avoided.

Ho's work on ordinal optimization (Ho, Sreenivas and Vakili (1992)) is more in the spirit of our approach. There mainly heuristic rules for huge discrete problems are proposed in the case when one has almost no chance of truly identifying the optimal solution.

To formulate our objective in a precise way, we introduce nonnegative real functions $c(x)$ with $c(0) = 0$ to measure the costs incurred by selecting a non-optimal $\hat{X}^{(n)}$ as solution for (1). The costs associated with $\hat{X}^{(n)}$ are given as $c(F(\hat{X}^{(n)}) - F^*)$. Typical cost functions $c(x)$ are monotonous. In particular $c(x) = x$ is appropriate, if the values of F have already an interpretation as costs.

A clever sample allocation rule should choose the decision variables m_i as to minimize the above costs in some sense. Let again $\hat{X}^{(n)}$ be a solution of (2). Then a possible goal would be

$$\left\| \begin{array}{l} \text{Minimize (in } (m_i)_{i\in S}) \quad u_1^{(n)} := \mathbf{E}[c(F(\hat{X}^{(n)}) - F^*)], \\ \text{such that } \sum_{i\in S} m_i = n, \quad m_i \geq 0. \end{array} \right. \tag{3}$$

It is often desirable to obtain not only an approximate solution for (1), but also to construct a confidence set that contains exact solutions of (1) with a certain prescribed probability. To be informative, the constructed confidence set should also contain as few points of S as possible. More formally, we will call a (random) subset $\hat{S}^{(n)}$ of S a level $1-\alpha$ confidence set, if

$$\mathbf{P}\{i^* \in \hat{S}^{(n)}\} \geq 1 - \alpha. \tag{4}$$

In statistical literature (see e.g. Gupta (1965)) such confidence sets are called subset selection rules.

An (obviously bad) level $1-\alpha$ confidence set is always given by choosing $\hat{S}^{(n)} = S$. Assume that including non-optimal i in $\hat{S}^{(n)}$ causes costs $c(F_i - F^*)$. Then the quality of a confidence set $\hat{S}^{(n)}$ can be measured by the expected costs and an optimal sampling plan should minimize these costs, i.e. solve

$$\left\| \begin{array}{l} \text{Minimize (in } (m_i)_{i\in S}) \quad u_2^{(n)} := \mathbf{E}[\sum_{i\in S} c(F_i - F^*)\mathbb{1}_{[i\in\hat{S}^{(n)}]}], \\ \text{such that } \sum_{i\in S} m_i = n, \quad m_i \geq 0. \end{array} \right. \tag{5}$$

If $c(x) = 1_{(0,\infty)}(x)$ then (5) is just the expected number of points in the confidence set not counting i^*.

Remark 1 *Rules that define confidence sets for normally distributed estimates $\hat{\boldsymbol{F}}^{(n)}$ are usually of the form*

$$i \in \hat{S}^{(n)} \Leftrightarrow \hat{F}_i^{(n)} \leq \min_j (\hat{F}_j^{(n)} + d_i \sqrt{\sigma_i^2 + \sigma_j^2}).$$

In literature one can find different proposals how to choose d_i. All of them satisfy (4) for arbitrary functions $\boldsymbol{F}$. With $\Phi(\cdot)$ denoting the normal distribution function, possible choices of d_i are

1. *Bonferroni rule: $d_i = d = \Phi^{-1}(1 - \frac{\alpha}{k-1})$*
2. *Gupta–Huang rule (see Gupta and Huang (1976)): Independently of i, d_i is the solution of*

$$\int \prod_{j \neq i^*} \Phi \left(\frac{d\sqrt{\sigma_{i^*}^2 + \sigma_j^2} - y}{\sigma_j} \right) d\Phi \left(\frac{y}{\sigma_{i^*}} \right) = 1 - \alpha,$$

   where $\sigma_{i^*}^2 = \min_j \sigma_j^2$.

3. *Gupta rule: Choose d_i as solution (in d) of*

$$\int \prod_{j \neq i} \Phi \left(\frac{d\sqrt{\sigma_i^2 + \sigma_j^2} - y}{\sigma_j} \right) d\Phi \left(\frac{y}{\sigma_i} \right) = 1 - \alpha.$$

The rules obtained by the above three choices for d_i will be denoted by $S_1^{(n)}$, $S_2^{(n)}$, and $S_3^{(n)}$ respectively. The order $S_3^{(n)} \subseteq S_2^{(n)} \subseteq S_1^{(n)}$ with respect to size is easy to verify. Notice however, that the smaller confidence sets are more tedious to implement.

Given one of the above rules we might ask how to carry out the sampling as to minimize the expected costs (5).

2 Asymptotically optimal sampling

Since the exact objective function occurring in our sample size allocation problems (3) and (5) is complicated, an asymptotic approximation is of interest. To obtain such an approximation we consider the following model:

Assume that for each $i \in S$ a normally $N(F_i, \sigma_i^2/m_i)$ distributed estimate $\hat{F}_i^{(n)}$ is available. The estimate might be thought as the arithmetic mean of independent samples.

We call a sample allocation asymptotically optimal for problem (3) or (5), if the rate of convergence of the expected costs i.e. $\lim_{n\to\infty} \frac{-1}{n} \log u_t^{(n)}$ (for $t = 1$ or $t = 2$) is maximal.

We will now address the question how to maximize the above rate. To this end define $\mathcal{S}$ to be the standard simplex, i.e. the set of points $\alpha = (\alpha_i)_{1 \leq i \leq k}$ with $\alpha_i \geq 0$ and

$\sum_{1\le i\le k}\alpha_i = 1$. Let π_{A_i} denote the projection w.r.t. the norm $\|\cdot\|_{\Lambda(\alpha)}$ onto the closed convex cone A_i, where $\|x\|_{\Lambda(\alpha)} = [x^t\Lambda(\alpha)x]^{1/2}$ with

$$\Lambda(\alpha) = \begin{pmatrix} \alpha_1/\sigma_1^2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_k/\sigma_k^2 \end{pmatrix},$$

and

$$A_i = \{x \in \mathbf{R}^k : x_i \le x_j,\ 1 \le j \le k\}.$$

We need a weak regularity condition: There is a sequence $a_n \to 0$ such that

$$i \in S^{(n)} \Longrightarrow \hat{F}_i^{(n)} \le \min_j \hat{F}_j^{(n)} + a_n. \tag{6}$$

Notice that all rules given in Remark 1 satisfy condition (6) provided that for all $i \in S$ we have $\liminf_i m_i/n > 0$.

Lemma 1 *Suppose that* $\frac{m_i}{n} \to \alpha_i$ *as* $n \to \infty$. *Let*

$$\gamma(\alpha) = \min_{i\ne i^*} \gamma_i(\alpha)$$

with

$$\gamma_i(\alpha) = \|\boldsymbol{F} - \pi_{A_i}\boldsymbol{F}\|^2_{\Lambda(\alpha)}.$$

Then for any cost function c *satisfying* $c(x) > 0$ *on* $(0,\infty)$ *and* $c(0) = 0$,

$$\lim_{n\to\infty} \frac{-2}{n} \log u_1^{(n)} = \gamma(\alpha).$$

If additionally either condition (6) holds or at least one $\alpha_i = 0$, *then also*

$$\lim_{n\to\infty} \frac{-2}{n} \log u_2^{(n)} = \gamma(\alpha).$$

Proof. Assume first $\alpha_i > 0$ for $1 \le i \le k$. We may write

$$u_1^{(n)} = \sum_{i\ne i^*} c(F_i - F^*)\mathbf{P}\{\hat{X}^{(n)} = i\} \tag{7}$$

and

$$u_2^{(n)} = \sum_{i\ne i^*} c(F_i - F^*)\mathbf{P}\{i \in \hat{S}^{(n)}\}. \tag{8}$$

To derive approximations for the probabilities in the above expressions, notice that

$$\mathbf{P}\{\hat{X}^{(n)} = i\} = \mathbf{P}\{\hat{F}_i^{(n)} = \min_{j\ne i} \hat{F}_j^{(n)}\} \le \mathbf{P}\{i \in \hat{S}^{(n)}\} \le \mathbf{P}\{\hat{F}_i^{(n)} \le \min_{j\ne i} \hat{F}_j^{(n)} + a_n\}, \tag{9}$$

where $a_n \to 0$. Defining $\hat{\boldsymbol{Y}}^{(n)}$ as vector with components

$$\begin{aligned} \hat{Y}_j^{(n)} &= \hat{F}_j^{(n)} + a_n \quad \text{for } j \neq i, \\ \hat{Y}_i^{(n)} &= \hat{F}_i^{(n)}, \end{aligned}$$

we will derive a large deviation approximation for

$$\mathbf{P}\{\hat{F}_i^{(n)} \leq \min_{j \neq i} \hat{F}_j^{(n)} + a_n\} = \mathbf{P}\{\hat{\boldsymbol{Y}}^{(n)} \in A_i\}.$$

For this we introduce

$$\begin{aligned} \varphi_n(\boldsymbol{t}) &= \frac{1}{n} \log \mathbf{E}(\exp[\boldsymbol{t}(n\hat{\boldsymbol{Y}}^{(n)})]) \\ &= \frac{1}{n}\left[nF_i t_i + \frac{n\sigma_i^2 t_i^2}{2\alpha_i} + \sum_{j \neq i}[n(F_j + a_n)t_j + \frac{n\sigma_j^2 t_j^2}{2\alpha_j}]\right]. \end{aligned}$$

Now

$$\varphi(\boldsymbol{t}) := \lim_{n \to \infty} \varphi_n(\boldsymbol{t}) = \sum_{j=1}^{k}\left[F_j t_j + \frac{n\sigma_j^2 t_j^2}{2\alpha_j}\right]$$

leads to the rate function (i.e. the dual function)

$$\begin{aligned} J(\boldsymbol{x}) &= \sup_t [\langle \boldsymbol{t}, \boldsymbol{x} \rangle - \varphi(\boldsymbol{t})] \\ &= \sum_{j=1}^{k} \frac{\alpha_j (x_j - F_j)^2}{2\sigma_j^2} \end{aligned}$$

Now, according to Ellises Theorem (see Bucklew (1990), p. 21)

$$\limsup \frac{1}{n} \log \mathbf{P}\{\hat{\boldsymbol{Y}}^{(n)} \in A_i\} \leq -\inf_{x \in A_i} J(\boldsymbol{x}),$$

and

$$\liminf \frac{1}{n} \log \mathbf{P}\{\hat{\boldsymbol{Y}}^{(n)} \in A_i^\circ\} \geq -\inf_{x \in A_i^\circ} J(\boldsymbol{x}).$$

Noting that $\inf_{x \in A_i} J(\boldsymbol{x}) = \inf_{x \in A_i^\circ} J(\boldsymbol{x}) = \gamma_i(\alpha)/2$ we obtain

$$\mathbf{P}\{\hat{\boldsymbol{Y}}^{(n)} \in A_i\} = \exp(-n[\gamma_i(\alpha)/2 + o(1)]) \tag{10}$$

which is an upper bound for $\mathbf{P}\{i \in \hat{S}^{(n)}\}$. The special choice $a_n = 0$ in the above arguments leads to the same bound (10) also for $\mathbf{P}\{\hat{X}^{(n)} = i\}$.

To shorten the notation we write c_i instead of $c(F_i - F^*)$. Then we obtain using (7) and (9)

$$u_1^{(n)} = \exp(-n[\gamma(\alpha)/2 + o(1)])\left(\sum_{i \neq i^*} c_i \exp[-n/2(\gamma_i(\alpha) - \gamma(\alpha) + o(1))]\right). \tag{11}$$

Notice that the above sum is equal to $\sum_{i:\gamma_i(\alpha)=\gamma(\alpha)} c_i + o(1)$. Therefore, for $t = 1$

$$-\frac{2}{n}\log(u_t^{(n)}) = -\gamma(\alpha) + o(1). \tag{12}$$

Starting from (8) the identical arguments as above yield (12) also for $t = 2$.

We now consider the case $\alpha_i = 0$ for at least one i. In this situation Lemma 2 (ii) states that $\gamma(\alpha) = 0$. Therefore, and since $u_t^{(n)}$ is bounded from above, it suffices to show that for $t \in \{1, 2\}$

$$\limsup_{n\to\infty}[-\log u_t^{(n)}/n] \leq 0. \tag{13}$$

Introduce the random quantities $\hat{u}_1^{(n)} = \sum_{i=1}^k c_i 1_{[\hat{X}^{(n)}=i]}$ and
$\hat{u}_2^{(n)} = \sum_{i=1}^k c_i 1_{[i\in\hat{S}^{(n)}]}$. Let furthermore $c^* = \min_{i\neq i^*} c_i > 0$. Since for $t \in \{1, 2\}$

$$u_t^{(n)} \geq c^* \mathbf{P}\{\hat{u}_t^{(n)} \geq c^*\}, \tag{14}$$

we may obtain (13) by establishing an adequate lower bound for (14). Assume first that $\alpha_{i^*} > 0$, and choose a j such that $\alpha_j = 0$. Then

$$\mathbf{P}\{\hat{u}_t^{(n)} \geq c^*) \geq \mathbf{P}(\hat{F}_{i^*}^{(n)} > \hat{F}_j^{(n)}\}.$$

By using the normal tail approximation $1 - \Phi(x) = \varphi(x)/x(1 + o(1))$, (see e.g. Barndorff–Nielsen and Cox (1989), p.56) the r.h.s. is equal to

$$1 - \Phi\left(\frac{F_j - F^*}{\sqrt{\frac{\sigma_{i^*}^2}{m_{i^*}} + \frac{\sigma_j^2}{m_j}}}\right) = \frac{e^{-m_j\delta}}{\sqrt{m_j}\delta'}(1 + o(1)),$$

for suitable constants δ and δ'. Thus

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbf{P}\{\hat{F}_{i^*}^{(n)} > \hat{F}_j^{(n)}\} = 0,$$

which establishes (13) for the case $\alpha_{i^*} > 0$. In the case $\alpha_{i^*} = 0$, (13) may be proved in the same way as above by choosing a j such that $\alpha_j > 0$. (Such an index j always exists.) □

Since asymptotically the optimal sample allocation may be formulated in terms of allocated proportions α_i, we consider in the sequel the approximate problem:

$$\left\| \begin{array}{l} \text{Maximize } \gamma(\alpha) \\ \text{such that } \sum_{i\in S}\alpha_i = 1, \quad \alpha_i \geq 0 \end{array} \right. \tag{15}$$

3 The approximate problem

To solve (15) we have to find the maximum of a concave (but nondifferentiable) function over the $(k-1)$–dimensional standard simplex. This could be done in principle by a convex

optimization procedure like the bundle method. However in our applications $\boldsymbol{F}$ is unknown and is replaced by estimates $\hat{\boldsymbol{F}}_n$. Furthermore $\hat{\boldsymbol{F}}_n$ (and thus our optimization problem) has to be updated after each new observation. Since solving an optimization problem in each step of the simulation would be quite tedious in most applications, it seems crucial to have a simple one-step rule which decides where to invest the next observation. Ideally each new observation should bring us closer to the optimal sampling plan.

The rule we will propose is of the Frank-Wolfe type in the sense that the direction of move is the solution of the linearized problem. In our case, the move will always be towards a corner of the simplex.

Unfortunately rules of the Frank-Wolfe type are not generally converging for nondifferentiable objective functions like that occurring in (15). A possible solution is to smooth the objective function by introducing a regularization. (Details as well as other approaches may be found in the paper by Ruszczynski (1987) and the references therein.) As regularization we propose to replace $\gamma(\alpha) = \min_{i \neq i^*}[\gamma_i(\alpha)]$ by $\gamma^\epsilon(\alpha) = \epsilon\text{-}\min_{i \neq i^*}[\gamma_i(\alpha)]$, where

$$\epsilon\text{-}\min(x_1, \ldots, x_\ell) := -\epsilon \log \left(\sum_{i=1}^{\ell} e^{-x_i/\epsilon} \right) \tag{16}$$

and solve the regularized problem

$$\left\| \begin{array}{l} \text{Maximize } \gamma^\epsilon(\alpha) \\ \text{such that } \sum_{i \in S} \alpha_i = 1, \quad \alpha_i \geq 0. \end{array} \right. \tag{17}$$

This is a smooth concave optimization problem, since by Lemma 5 (iii) the function ϵ-min is concave, monotone and smooth.

Remark 2 *Another reasonable approximation could be obtained by choosing*

$$\gamma^\epsilon(\alpha) = -\epsilon \log \left(\sum_{i \neq i^*} c(F_i - F^*) e^{-\gamma_i(\alpha)/\epsilon} \right).$$

According to (11) the above expression can be viewed as an approximation of $u_1^{(n)}$ *and* $u_2^{(n)}$.

If all parameters F_i and σ_i are known, the function γ^ϵ can be optimized by a Frank-Wolfe type algorithm. We state here a general convergence result:

Theorem 1 *Let* $\gamma(\alpha)$ *be a convex differentiable function defined on the simplex* $\mathcal{S}$. *Let the gradient* $\nabla\gamma(\alpha)$ *be Lipschitz. For every* $\alpha \in \mathcal{S}$, *let* $e(\alpha)$ *be the* $i-th$ *unit vector, where* i *is the smallest index such that*

$$\frac{\partial \gamma(\alpha)}{\partial \alpha_i} = \max_j \frac{\partial \gamma(\alpha)}{\partial \alpha_j}.$$

Let $\alpha^{(s)}$ *be the sequence generated by the following Frank-Wolfe type algorithm*

$$\alpha^{(s+1)} = (1 - \frac{1}{s})\alpha^{(s)} + \frac{1}{s} e(\alpha^{(s)}). \tag{18}$$

Then

$$\lim_{s \to \infty} \gamma(\alpha^{(s)}) = \max_{\alpha \in \mathcal{S}} \gamma(\alpha).$$

Proof. Let

$$\varphi(\alpha) = \max_i \frac{\partial \gamma(\alpha)}{\partial \alpha_i} - \langle \nabla \gamma(\alpha), \alpha \rangle. \tag{19}$$

Notice that $\varphi(\alpha) \geq 0$ and that φ is Lipschitz. Moreover the necessary and sufficient optimality condition

$$\nabla \gamma(\alpha) - \frac{1}{k} \langle \nabla \gamma(\alpha), \mathbb{1} \rangle \mathbb{1} = 0$$

is equivalent to $\varphi(\alpha) = 0$. Here $\mathbb{1}$ denotes the vector with all components equal 1. Let

$$\gamma^* = \max_{\alpha \in \mathcal{S}} \gamma(\alpha).$$

By the mean value theorem

$$\begin{aligned} \gamma(\alpha^{(s+1)}) - \gamma(\alpha^{(s)}) &= \langle \nabla \gamma[(1-t)\alpha^{(s)}) + t\alpha^{(s+1)}], \frac{1}{s}[e(\alpha^{(s)}) - \alpha^{(s)}] \rangle \\ &= \langle \nabla \gamma(\alpha^{(s)}), \frac{1}{s}[e(\alpha^{(s)}) - \alpha^{(s)}] \rangle + R_s \\ &= \frac{1}{s}\varphi(\alpha^{(s)}) + R_s. \end{aligned}$$

The remainder terms R_s are of order $O(s^{-2})$. Fix an $\epsilon > 0$. By continuity, there is an $\eta > 0$ such that $\varphi(\alpha) \leq \eta$ implies $\gamma(\alpha) \geq \gamma^* - \epsilon$. Hence the following inequality holds true

$$\gamma(\alpha^{(s+1)}) \geq \min[\gamma(\alpha^{(s)}) + \frac{\eta}{s} + R_s, \gamma^* - \epsilon]. \tag{20}$$

Since $\sum_s R_s < \infty$, the relation (20) implies that

$$\liminf_s \gamma(\alpha^{(s)}) \geq \gamma^* - \epsilon$$

and because ϵ was arbitrary,

$$\liminf_s \gamma(\alpha^{(s)}) = \gamma^*.$$

□

4 Practical implementation and simulation

Since $\boldsymbol{F}$ and $(\sigma_i^2)_{i \in S}$ will be unknown in practice, it is natural to replace them by estimates obtained during sampling. We propose the following algorithm for practical application.

Algorithm:

1. *Choose N as the total number of observations to be taken.*
2. *Choose some $\epsilon > 0$ and some $n_0 \geq 1$.*
3. *Take an initial sample of size n_0 from each $i \in S$. and set the total number of already taken observations $s = k \cdot n_0$.*

4. *If $s > N$ stop.*

5. *Suppose that $\hat{F}_i^{(s)}$, $\hat{\sigma}_i^{(s)}$ are the actual estimates of F_i, σ_i after taking $m_i^{(s)}$ observations at alternative i. Let $\alpha^{(s)}$ be the actual vector of relative frequencies, i.e. $\alpha_i(s) = \frac{m_i^{(s)}}{s}$.*

6. *Calculate (for all $i \neq i^*$) $\hat{\gamma}_i(\alpha^{(s)})$ according to Remark 3 (iii) and $\nabla\hat{\gamma}_i(\alpha)$ according to Lemma 3. (Replace in all formulas F_i and σ_i by $\hat{F}_i^{(s)}$ and $\hat{\sigma}_i^{(s)}$).*

7. *Calculate*

$$\nabla\hat{\gamma}^{\epsilon}(\alpha) = \sum_{i\neq i^*} \frac{e^{-\hat{\gamma}_i(\alpha^{(s)})/\epsilon}}{\sum_{j\neq i^*} e^{-\hat{\gamma}_j(\alpha^{(s)})/\epsilon}} \nabla\hat{\gamma}_i(\alpha^{(s)})$$

8. *Find the component of $\nabla\hat{\gamma}^{\epsilon}(\alpha)$ with maximal value and denote it by l , i.e.*

$$[\nabla\hat{\gamma}^{\epsilon}(\alpha)]_l = \max_j [\nabla\hat{\gamma}^{\epsilon}(\alpha)]_j \,.$$

Make one additional observation ξ_l for alternative l and adjust the estimates

$$\begin{aligned}
\hat{F}_l^{(s+1)} &= \frac{m_l^{(s)}}{m_l^{(s)}+1}\hat{F}_l^{(s)} + \frac{1}{m_l^{(s)}+1}\xi_l, \\
\left[\hat{\sigma}_l^{(s+1)}\right]^2 &= \frac{1}{m_l^{(s)}}\left[(m_l^{(s)}-1)\left[\sigma_l^{(s)}\right]^2 + \xi_l^2 - (m_l^{(s)}+1)[\hat{F}_l^{(s+1)}]^2 + (m_l^{(s)})[\hat{F}_l^{(s)}]^2\right], \\
m_l^{(s+1)} &= m_l^{(s)} + 1
\end{aligned}$$

and keep all other estimates unchanged.

9. *Increase s by 1 and go to 4.*

The convergence of this algorithm is given by the following theorem

Theorem 2 *The just described algorithm satisfies*

$$\lim_s \gamma^{\epsilon}(\alpha^{(s)}) = \max_{\alpha\in f} \gamma^{\epsilon}(\alpha) a.s.$$

Proof. We omit ϵ for simplicity in the proof. Let $\hat{\gamma}_s$ be the actual estimate of the function γ^{ϵ} in step s. and let $\hat{\varphi}_s$ the expression analogous to (19). Let $\hat{e}_s$ be the direction of move in step s. We have by the mean value theorem

$$\begin{aligned}
\gamma(\alpha^{(s+1)}) - \gamma(\alpha^{(s)}) &= \langle\nabla\gamma((1-t)\alpha^{(s)}) + t\alpha^{(s+1)}, \frac{1}{s}[\hat{e}_s(\alpha^{(s)}) - \alpha^{(s)}]\rangle \\
&= \langle\nabla\gamma(\alpha^{(s)}), \frac{1}{s}[\hat{e}_s(\alpha^{(s)}) - \alpha^{(s)}]\rangle + R_s \\
&= \langle\nabla\hat{\gamma}_s(\alpha^{(s)}), \frac{1}{s}[\hat{e}_s(\alpha^{(s)}) - \alpha^{(s)}]\rangle \\
&\quad + \langle\nabla\gamma(\alpha^{(s)}) - \nabla\hat{\gamma}_s(\alpha^{(s)}), \frac{1}{s}[\hat{e}_s(\alpha^{(s)}) - \alpha^{(s)}]\rangle + R_s \\
&= \frac{1}{s}\hat{\varphi}_s(\alpha^{(s)}) + \frac{1}{s}V_s + R_s.
\end{aligned}$$

Here $|V_s| \leq \|\nabla\gamma(\alpha^{(s)}) - \nabla\hat{\gamma}_s(\alpha^{(s)})\|$. By the Law of Large Numbers, $V_s \to 0$ a.s. Choose ϵ and η as in the proof of theorem 1. Choosing s_0 large enough to make $\sup_{s \geq s_0} |V_s| \leq \eta/2$ and $\sup_\alpha |\hat{\varphi}_s(\alpha) - \varphi(\alpha)| \leq \eta/2$ (on a set of arbitrary large probability) we get the recursion

$$\gamma(\alpha^{(s+1)}) \geq \min[\gamma(\alpha^{(s)}) + \frac{\eta}{4s} + R_s, \gamma^* - \epsilon], \qquad s \geq s_0. \tag{21}$$

By the same argument as in (20), this implies that

$$\lim_s \gamma(\alpha^{(s)}) = \gamma^*.$$

□

s	$m_1^{(s)}$	$m_2^{(s)}$	$m_3^{(s)}$	$m_4^{(s)}$	$m_5^{(s)}$	$m_6^{(s)}$	$m_7^{(s)}$	$m_8^{(s)}$	$m_9^{(s)}$	$m_{10}^{(s)}$
200	20	20	20	20	20	20	20	20	20	20
600	20	20	20	20	20	20	21	161	162	136
1000	20	20	20	20	20	55	66	307	239	233
1400	20	20	20	20	20	55	66	307	442	430
1800	20	20	20	20	20	55	66	307	643	629
2200	20	20	20	20	20	55	66	307	833	839
5000	20	20	24	37	27	104	104	351	2133	2180
10000	20	20	41	60	62	164	332	686	4204	4411
15000	29	26	56	99	140	227	450	1133	6257	6583

Table 1.

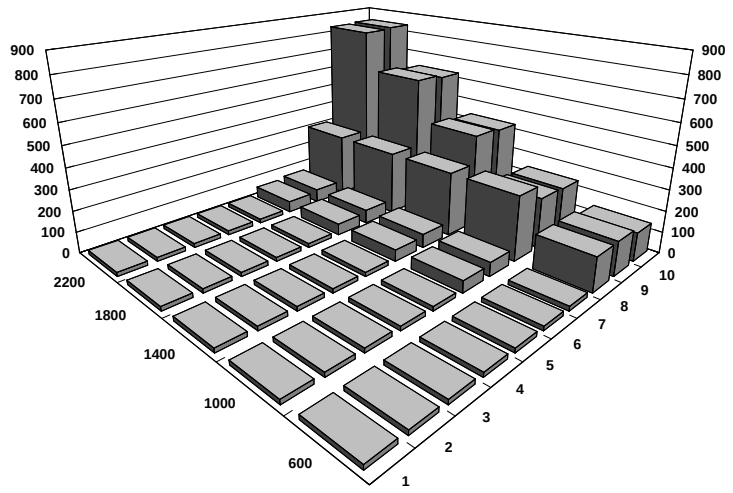

Figure 1: The development of the sample sizes for the first 2200 allocations.

To investigate the performance of the algorithm, it has been applied to the following example: Assume that $S = \{1, \ldots, 10\}$ and that normally distributed observations from

$\boldsymbol{F}$ with $F_i = 0.008i^2 - 0.3762i$ are available. The variance of an observation at point i is $0.88 + 0.11i$. This choice seems reasonable, since in practice the variance is often largest for observations at the optimum, which is here $i^* = 10$. An initial sample of size $n_0 = 20$ has been taken for each observation. Then our algorithm has been applied to obtain further observations. The regularized gradients in step 7 have been based on $\epsilon = 0.001$. Table 1 respectively Figure 1 show the cumulative allocations for the first 15000 resp. 2200 observations.

Figure 2 displays the rates $\hat{\gamma}_i(\alpha^{(s)})$ as they occurred during the first 5000 steps. One sees that not only the minimal rate increases but also that the maximal rate decreases. We conjecture but could not prove that for the optimal α all γ_i's are equal.

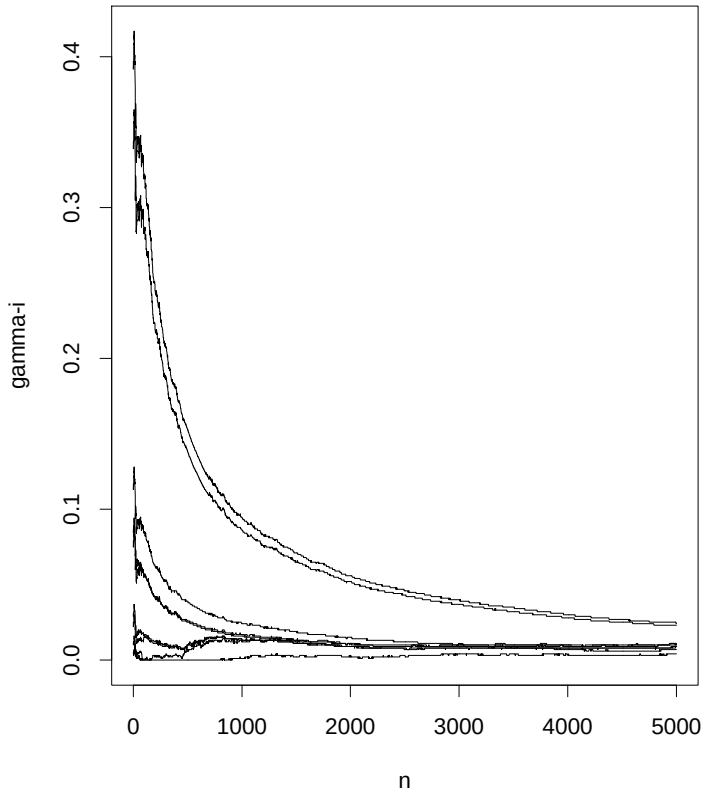

Figure 2: The rate functions $\hat{\gamma}_i(\alpha^{(s)})$.

5 Appendix: Properties of γ_i and γ.

We will first summarize what is known from literature.

Remark 3 *The following results may be verified immediately e.g. from Robertson et al. (1988), example 1.3.2.*

(i) $\gamma_i(\alpha) = \min_m \left[\frac{\alpha_i}{\sigma_i^2}(F_i - m)^2 + \sum_{j \neq i} \frac{\alpha_j}{\sigma_j^2}(F_j - m)^2 1_{[F_j \leq m]} \right]$.

(ii) The quantity $m_i^*(\alpha) = argmin_m \gamma_i(\alpha)$ *is unique and can be calculated as follows: Let* $d(l)$ *denote the antiranks of* $(F_j)_{j \neq i}$, *i.e* $F_{d(l)} = F_{[l]}$ *where* $F_{[l]}$ *is the l-th smallest element of* $(F_j)_{j \neq i}$. *Define sums with indexes from* 1 *to* 0 *to be* 0. *Then, with* $w_j = \alpha_j / \sigma_j^2$

$$m_i^*(\alpha) = \min_{0 \leq l \leq k-1} \left(\frac{w_i F_i + \sum_{j=1}^{l} w_{d(j)} F_{d(j)}}{w_i + \sum_{j=1}^{l} w_{d(j)}} \right).$$

(iii) $\gamma_i(\alpha) = \frac{\alpha_i}{\sigma_i^2}(F_i - m_i^*(\alpha))^2 + \sum_{j \neq i} \frac{\alpha_j}{\sigma_j^2}(F_j - m_i^*(\alpha))\, 1_{[F_j \leq m_i^*(\alpha)]}$.

Lemma 2 *For* $\gamma(\alpha) = \min_{i \neq i^*} \gamma_i(\alpha)$ *we have*

(i) $\gamma(\cdot)$ *is concave and nonnegative;*

(ii) $\gamma(\alpha) = 0$ *at the boundary of* $\mathcal{S}$.

Proof. We start by proving (i). From Remark 3 (i) it follows immediately that all γ_i are nonnegative and thus γ cannot be negative. Furthermore it may be seen immediately that each function γ_i is a minimum of linear functions (in α) and thus concave. Therefore γ – being the minimum of concave functions – is also concave.
Proof of (ii): Since we know that $\gamma(\alpha) \geq 0$, it is sufficient to show that $\gamma(\alpha) \leq 0$ at the boundary, i.e. for points α, where $\alpha_i = 0$ for at least one $i \in S$.

Assume first that $\alpha_i = 0$ for at least one $i \neq i^*$. By Remark 3 (i)

$$\gamma_i(\alpha) \leq g_i(\alpha, m) := \left[\frac{\alpha_i}{\sigma_i^2}(F_i - m)^2 + \sum_{j \neq i} \frac{\alpha_j}{\sigma_j^2}(F_j - m)^2 1_{[F_j \leq m]} \right], \tag{22}$$

for any m. Therefore we obtain that

$$\gamma(\alpha) \leq \gamma_i(\alpha) \leq g_i(\alpha, \min_j F_j) = 0.$$

Let us now assume $\alpha_i^* = 0$. Take i^+ as an index that satisfies $F_{i^+} = F_{[2]}$, where $F_{[2]}$ denotes the second smallest value of $\boldsymbol{F}$. Then, using again (22),

$$\gamma(\alpha) \leq \gamma_{i^+}(\alpha) = g_i(\alpha, F_{i^+}) = 0.$$

□

Lemma 3 *The function $\alpha \mapsto \gamma_i(\alpha)$ is differentiable with gradient*

$$[\nabla\gamma_i(\alpha)]_j = \begin{cases} \frac{1}{\sigma_j^2}(F_j - m_i^*(\alpha))^2 1_{[F_j \le m_i^*(\alpha)]} & j \neq i \\ \frac{1}{\sigma_i^2}(F_i - m_i^*(\alpha))^2 & j = i \end{cases},$$

where $m_i^*(\alpha)$ is calculated according to Remark 3 (ii).

Proof. Notice that $\gamma_i(\alpha)$ has representation

$$\gamma_i(\alpha) = \min_m \langle c_i(m), \alpha\rangle,$$

where $c_i(m)$ is defined as

$$c_i(m) = \begin{cases} \frac{1}{\sigma_j^2}(F_j - m)^2 1_{[F_j \le m]} & j \neq i \\ \frac{1}{\sigma_i^2}(F_i - m)^2 & j = i. \end{cases} \tag{23}$$

It is well known that the subgradient is given as,

$$\partial\gamma_i(\alpha) = \text{conv}\{ \text{ argmin }_{c_i(m)}\langle c_i(m), \alpha\rangle\}.$$

Since the argmin is unique it follows that

$$\nabla\gamma_i(\alpha) = c_i(m_i^*(\alpha)),$$

with m_i^* as in Remark 3 (ii). Therefore the gradient $\nabla\gamma_i(\alpha)$ is given by (23) with m replaced by m_i^*, i.e. $[\nabla\gamma_i(\alpha)]_j = [c_i(m_i^*)]_j$. □

Let, for any $c \in \mathbf{R}^k$, $\tilde{c} = c - \frac{1}{k}\langle c, \mathbb{1}\rangle\mathbb{1}$. Call $\widetilde{\nabla\gamma_i(\alpha)}$ the reduced gradient. The lemma below states the optimality condition for (15).

Lemma 4 *α^* solves (15), if and only if*

$$0 \in conv\{\widetilde{\nabla\gamma_i(\alpha^*)} : i \in I(\alpha^*)\}.$$

where $I(\alpha) = \{i : \gamma_i(\alpha) = \min_j \gamma_j(\alpha)\}$.

Proof. Let $\gamma'(\alpha, \beta) = \lim_{t\searrow 0}\frac{1}{t}[\gamma((1-t)\alpha + t\beta) - \gamma(\alpha)]$ be the directional derivative. It is easily verified that

$$\gamma'(\alpha, \beta) = \min\{\langle\nabla\gamma_i(\alpha), \beta - \alpha\rangle : i \in I(\alpha)\}.$$

α^* is optimal, if and only if there is no $\beta \in \mathcal{S}$ such that

$$\gamma'(\alpha^*, \beta) > 0 \quad \text{for all } i \in I(\alpha^*). \tag{24}$$

We will show that (24) is equivalent to our assertion. Let $c_i = \nabla\gamma_i(\alpha^*)$ and $I = I(\alpha^*)$. Assume first that

$$0 \in \text{conv}\{\tilde{c}_i : i \in I\}$$

and that (24) does not hold. Then there would be a $\beta \in \mathcal{S}$ such that $\langle c_i, \beta - \alpha^* \rangle > 0$ for all $i \in I$. Since

$$\langle c_i, \beta - \alpha^* \rangle = \langle \tilde{c}_i, \beta - \alpha^* \rangle$$

this implies

$$\langle \tilde{c}_i, \beta - \alpha^* \rangle > 0 \quad \text{for all } i \in I.$$

But since according to our assumption 0 is a convex combination of these $\tilde{c}_i$ it follows that

$$\langle 0, \beta - \alpha^* \rangle > 0,$$

which is a contradiction.

Assume now on the other hand that $0 \notin \text{conv}\{\tilde{c}_i : i \in I\}$. Then there exists a vector u such that $\langle \tilde{c}_i, u \rangle > 0$ for all $i \in I$. Since $\langle \tilde{c}_i, \mathbb{1} \rangle = 0$, we have that

$$\langle \tilde{c}_i, \tilde{u} \rangle > 0.$$

Choose now $\lambda > 0$ small enough such that

$$\beta = \lambda \tilde{u} + \alpha^*$$

is nonnegative. This is possible since α^* is in the interior of $\mathcal{S}$. With the above choice

$$\langle \tilde{c}_i, \beta - \alpha^* \rangle > 0 \text{ for all } i \in I,$$

which is a contradiction to (24).

□

The following Lemma states some facts about the function ϵ-min defined in (16).

Lemma 5 *(i)* $\epsilon\text{-min}(x_1, \ldots, x_\ell) \leq \min(x_1, \ldots, x_\ell)$.

(ii) $|\epsilon\text{-min}(x_1, \ldots, x_\ell) - \min(x_1, \ldots, x_\ell)| \leq \epsilon(\ell - 1)$.

(iii) $\epsilon\text{-min}(x_1, \ldots, x_\ell)$ *is monotone in each argument and concave.*

(iv) $\epsilon\text{-min}(x_1, \ldots, x_\ell)$ *is differentiable with the Lipschitz bound*

$$\|\nabla \epsilon\text{-min}(\boldsymbol{x}) - \nabla \epsilon\text{-min}(\boldsymbol{y})\| \leq \frac{2\ell}{\epsilon} \|\boldsymbol{x} - \boldsymbol{y}\|$$

for the gradient. The partial derivatives satisfy

$$\left|\frac{\partial}{\partial x_j} \epsilon\text{-min}(x_1, \ldots, x_\ell)\right| \leq 1$$

and

$$\lim_{\epsilon \to 0} \frac{\partial}{\partial x_j} \epsilon\text{-min}(x_1, \ldots, x_\ell) = \begin{cases} 0, & x_j > \min(x_1, \ldots, x_\ell) \\ b_i, & x_j = \min(x_1, \ldots, x_\ell) \end{cases}$$

where $b_i = (\#\{i : x_i = \min(x_1, \ldots, x_\ell)\})^{-1}$.

Proof. Statement (i) is obvious. To prove assertion (ii) let $x^* = \min(x_1, \ldots, x_\ell)$. Then

$$\begin{aligned} 0 &\leq \min(x_1, \ldots, x_\ell) - \epsilon\text{-}\min(x_1, \ldots, x_\ell) \\ &\leq \epsilon \left[\log \left(\sum_{i=1}^{\ell} e^{-x_i/\epsilon} \right) - \log \left(e^{-x^*/\epsilon} \right) \right] \\ &\leq \frac{\epsilon}{e^{-x^*/\epsilon}} \left[\sum_{i=1}^{\ell} e^{-x_i/\epsilon} - e^{-x^*/\epsilon} \right] \\ &= \epsilon \left[\sum_{i=1}^{\ell} e^{-(x_i - x^*)/\epsilon} - 1 \right] \\ &\leq \epsilon(\ell - 1). \end{aligned}$$

Proof of (iii). It is obvious that

$$x_j \mapsto -\epsilon \log \left(\sum_{i=1}^{\ell} e^{-x_i/\epsilon} \right)$$

is monotonically increasing. To prove concavity it is enough to consider the case $\epsilon = 1$ for the ϵ-min. Denote by $\mathcal{S}^\circ_{\ell-1} = \{\beta \in \mathbf{R}^\ell : \ \beta_i > 0, \sum_{i=1}^{\ell} \beta_i = 1\}$ the interior of the standard simplex. We show that

$$1\text{-}\min(x_1, \ldots, x_\ell) = \inf\{\sum_{i=1}^{\ell} (x_i \beta_i + \beta_i \log \beta_i) : \ \beta \in \mathcal{S}^\circ_{\ell-1}\} \tag{25}$$

which implies immediately concavity.

By Jensen's inequality, for all $\beta \in \mathcal{S}^\circ_{\ell-1}$

$$\begin{aligned} 1\text{-}\min(x_1, \ldots, x_\ell) &= -\log \left(\sum_{i=1}^{\ell} e^{-x_i} \right) = -\log \left(\sum_{i=1}^{\ell} \beta_i \frac{e^{-x_i}}{\beta_i} \right) \\ &\leq -\sum_{i=1}^{\ell} \beta_i \log \left(\frac{e^{-x_i}}{\beta_i} \right) \\ &= \sum_{i=1}^{\ell} (x_i \beta_i + \beta_i \log \beta_i). \end{aligned}$$

On the other hand, choosing $\bar{\beta}_i = \frac{e^{-x_i}}{\sum_{1 \leq j \leq \ell} e^{-x_j}}$ one sees that

$$1\text{-}\min(x_1, \ldots, x_\ell) = \sum_{i=1}^{\ell} (x_i \bar{\beta}_i + \bar{\beta}_i \log \bar{\beta}_i)$$

and this implies (25).

Proof of (iv): Since

$$\frac{\partial}{\partial x_j} \epsilon\text{-}\min(x_1, \ldots, x_\ell) = \frac{e^{-x_j/\epsilon}}{\sum_{i=1}^{\ell} e^{-x_i/\epsilon}},$$

the differentiability and the bound on the partial derivatives is obvious. The Lipschitz bound can be obtained by the mean value theorem applied to the gradient, since no second partial derivative is larger than $2/\epsilon$.

Noticing that the gradient vector $\nabla\epsilon\text{-min}(x_1, \ldots, x_\ell)$ is equal to the Gibbs distribution on $1, \ldots, \ell$, the limit $\lim_{\epsilon \to 0} \epsilon\text{-min}(x_1, \ldots, x_\ell)$ can e.g. be found in Corollary 2.1 of Aarts and Korst (1989).

□

References

[1] Aarts, E., Korst, J. (1989) Simulated annealing an boltzmann machines. Wiley, Chichester.

[2] Barndorff–Nielsen, O.E., Cox, D.R. (1989) Asymptotic techniques for use in statistics. Chapman and Hall, London.

[3] Bucklew, J. A. (1990). Large deviation techniques in decision, simulation, and estimation. Wiley, New York.

[4] Frank, M., Wolfe, P. (1956) An algorithm for quadratic programming. Naval Research Logistics Quarterly 3, 95–110.

[5] Gittins, J. C. (1989). Multi–armed Bandit Allocation Indices. Wiley, Chichester.

[6] Gupta, S. S. (1965). On some multiple decision (selection and ranking) rules. Technometrics 7, 225–245.

[7] Gupta, S. S., Huang, D.Y. (1976). Subset selection procedures for the means and variances of normal populations: Unequal sample sizes case. Sankhya, A 36, 112–128.

[8] Kleijnen, J. P. C. (1975) Statistical techniques in simulation. Part II. Dekker, New York.

[9] Ho, Y.C., Sreenivas, R.S., Vakili, P. (1992). Ordinal Optimization of DEDS. J. of Discrete Event Dynamical Systems 2 (2), 61–88.

[10] Lai, T. L., Robbins, H. (1985). Asymptotically efficient adaptive allocation rules. Adv. Appl. Math. 6, 4–22.

[11] Ruszczyński, A. (1987) A linearization method for nonsmooth stochastic programming problems. Mathem. of Oper. Res. 12, 32–49.

[12] Robertson, T., Wright, F. T., & Dykstra, R. L. (1988). Order Restricted Statistical Inference. Wiley, New York.