International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

Adaptive Design in Discrete Stochastic Optimization

Pflug, G.C.

IIASA Working Paper

WP-94-059

July 1994

Working Paper

Adaptive Designs in Discrete Stochastic Optimization

G. Gh. Pflug

WP-94-59
July, 1994

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 71521 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Adaptive Designs in Discrete Stochastic Optimization

G. Gh. Pflug

WP-94-59
July, 1994

International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 71521 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Adaptive Designs in Discrete Stochastic Optimization

G. Gh. Pflug
IIASA and University of Vienna

Abstract

We present adaptive assignment rules for the design of the necessary simulations when solving discrete stochastic optimization problems. The rules are constructed in such a way, that the expected size of confidence sets for the optimizer is as small as possible. .

1 Introduction

Suppose we have to find the best of m possible alternatives for a decision problem under uncertainity. Let ξ_x be a random variable (or a random vector or a random process) modelling the uncertainity. The probability law of ξ_x may depend on the decision x. If $H(x, \xi_x)$ measures the cost associated with decision x and the random outcome ξ_x, the problem is to minimize the expected costs $\mathbb{E}(H(x, \xi_x))$ over a finite set of alternatives:

$$(P) \quad \left\| \begin{array}{l} \text{Minimize } F(x) = \mathbb{E}(H(x, \xi_x)) \\ x \in \{x_1, \ldots, x_m\}. \end{array} \right.$$

If $F(\cdot)$ can be evaluated easily, problem (P) is a discrete optimization problem and techniques like Branch-and-Bound, Simulated Annealing etc. can be applied.

If the evaluation of F is impossible, one has to use replicated simulations $\xi_{x,k}$ of the random variable (or random vector or random process) ξ_x and approximate the problem (P) by the simulated problem (P_n):

$$(P_n) \quad \left\| \begin{array}{l} \text{Minimize } F_n(x) = \frac{1}{n} \sum_{k=1}^{n} H(x, \xi_{x,k}) \\ x \in \{x_1, \ldots, x_m\}. \end{array} \right.$$

For simplicity, we assume that the copies $\xi_{x,k}$ are i.i.d. A very similar approach as presented here applies to the case where $(H(x, \xi_{x,k}))_{k \geq 0}$ consitutes – for each x – an ergodic sequence which satisfies a Central Limit Theorem.

Set

$$X_{i,k} = H(x_i, \xi_{x_i,k}); \qquad \bar{X}_{i,n} = \frac{1}{n} \sum_{k=1}^{n} X_{i,k}.$$

Then

$$\begin{array}{ll} \mathbb{E}(X_{i,k}) = F(x_i) & \text{Var}\,(X_{i,k}) = \sigma_i^2 \text{ (say)} \\ \mathbb{E}(\bar{X}_{i,n}) = F(x_i) & \text{Var}\,(\bar{X}_{i,n}) = \frac{\sigma_i^2}{n}. \end{array}$$

(All variances are assumed to be finite.)

By the Law of Large Numbers $\bar{X}_{i,n} \to F(x_i)$ a.s. This and the finiteness of S imply that

$$\text{argmin}_x F_n(x) \overset{set}{\to} \text{argmin}_x F(x) \qquad a.s. \tag{1}$$

This illustrates that we may solve (P) by simulation of the empirical problem (P_n) for sufficiently large n.

However, the fact that we have consistency (i.e. a.s. convergence in (1)) does not tell anything about the efficiency of the method for small sample sizes and a restricted time budget.

In this paper, we discuss *experimental design plans* for the simulation experiment. It is clear, that it does not make sense to invest much time and effort in the simulation at a parameter point x_i, which is far from being optimal. On the other hand, we cannot exclude finally some parameter points in an early stage, since it may happen that due to random fluctuations our observations were much larger than their mean.

An experimental plan is called *non-adaptive*, if the same number n of replications are used for every alternative. In contrast, an *adaptive* plan determines the parameter points x_i, at which the next simulations are to be made, on the basis of the information gathered so far.

In the case of a continuous decision space, Stochastic Approximation methods and Response Surface Methods are well known adaptive design techniques.

We present here some adaptive assignment rules for a finite decision space, which exhibit the following properties:

(i) At each stage of the procedure, there is a confidence set available, which contains the argmin of F with a prespecified probability $1 - \alpha$.

(ii) No parameter is excluded from further investigation in a final manner. All of them are get an infinite number of observations, if the total number of observations tends to infinity. However the relative frequency of investigation of a particular parameter point can be very rather, if it shows itself not a good candidate for the minimizer.

(iii) If it is known that the minimizer is unique, there can be a nice stopping time based on the rules: The search may stop, if the confidence set contains only one point.

Recall that the theory of optimal experimental design used in statistical estimation is based on some measure of the information which is contained in a statistical sample. If we have already collected some data, we may calculate the amount of information we can expect, if we make an additional observation at x. This expected measure of information is often but not necessarily obtained by a Bayesian argument. The next observation is made at the point x, which maximizes the expected information.

We adopt here a similar approach. As a (negative) measure of information about the unknown minimizer, we consider the expected size of a confidence set for this minimizer. This is indeed a good indicator, since to obtain a small confidence set is all that we can expect from random optimization problems. It also makes sure that we never exclude parameters in a final manner, since the sizes of the confidence sets depend on the means *and* variances and therefore the size of the confidence set can only be small, if all variances are small.

In the next section, we describe methods for the construction of confidence sets. The pertaining assignment rules are presented in section 3. Section 4 contains an Example and the last section is devoted to the comparison with other methods.

2 Confidence sets

The construction of confidence sets in discrete stochastic optimization is called selection.

For simplicity, we treat only the case of normally distributed observations. In principle, every location-scale family may be treated in a similar manner, if the normal distribution function Φ is replaced by a general distribution function G.

Suppose that for each population, i.e. each $i \in \{1, \ldots m\}$ we may observe $Z_i \sim N(\mu_i, \sigma_i^2)$, where the mean values μ_i are unknown, but the variances σ_i^2 are known. We have to select candidates for the minimal μ_i by defining a confidence set $C = C(Z_1, \ldots, Z_m) \subseteq \{1, \ldots, m\}$.

Definition 1. $C = C(Z_1, \ldots, Z_m)$ is called *level α confidence set*, if for all minimizers μ_i, i.e. those who fulfill $\mu_i = \min_j \mu_j$:

$$\mathbb{P}\{i \in C\} \geq 1 - \alpha.$$

Those indices i, which are in C are called *selected*. *Selection Rule* is just a synonymous word for confidence set.

It is clear that among all level-α confidence sets, we prefer those with small size. The simplest selection rule uses the Bonferroni inequality. Recall that this inequality states that for arbitrary sets A_i

$$\mathbb{P}(\bigcup_{j=1}^{m} A_i) \leq \sum_{j=1}^{m} \mathbb{P}(A_j). \tag{2}$$

BONFERRONI RULE

Select population i if $Z_i \leq \min\{Z_j + d\sqrt{\sigma_i^2 + \sigma_j^2} : j \neq i\}$
where
$d = \Phi^{-1}(1 - \frac{\alpha}{m-1})$

Lemma 1. The Bonferroni-Rule defines a level α confidence set in the sense of Definition 1.

Proof. Suppose that $\mu_i = \min_j \mu_j$. Let $Z_j = Y_j + \mu_j$ with $Y_j \sim N(0, \sigma_j^2)$. We have to show that the probability that i is not selected is at most α.

$$\begin{aligned}
\mathbb{P}(Z_i > \min\{Z_j + d\sqrt{\sigma_i^2 + \sigma_j^2}\}) &= \mathbb{P}(\bigcup_{j \neq i}\{Z_i > Z_j + d\sqrt{\sigma_i^2 + \sigma_j^2}) \\
&\leq \sum_{j \neq i} \mathbb{P}(Z_i > Z_j + d\sqrt{\sigma_i^2 + \sigma_j^2}) \qquad \text{(by Bonferroni inequality)} \\
&= \sum_{j \neq i} \mathbb{P}(Y_i + \mu_i > Y_j + \mu_j + d\sqrt{\sigma_i^2 + \sigma_j^2}) \\
&= \sum_{j \neq i} \mathbb{P}\left(\frac{Y_j - Y_i}{\sqrt{\sigma_i^2 + \sigma_j^2}} < \frac{\mu_i - \mu_j}{\sqrt{\sigma_i^2 + \sigma_j^2}} - d\right) \\
&= \sum_{j \neq i} \Phi\left(\frac{\mu_i - \mu_j}{\sqrt{\sigma_i^2 + \sigma_j^2}} - d\right) \\
&\leq \sum_{j \neq i} \Phi(-d) = \sum_{j \neq i} \frac{\alpha}{m-1} = \alpha \qquad \text{(since } \mu_i \leq \mu_j\text{)}
\end{aligned}$$

□

The rule is based on the Bonferroni-inequality (2) which is only a rough estimate. It is also possible to use the exact value of $\mathbb{P}(\bigcup_{j \neq i}\{Z_i > Z_j + d\sqrt{\sigma_i^2 + \sigma_j^2})$. This leads to the Gupta Rule:

GUPTA RULE

Select population i **if** $Z_i \leq \min\{Z_j + d_i\sqrt{\sigma_i^2+\sigma_j^2} : j \neq i\}$
where
d_i **is the solution of**
$\int \prod_{j\neq i} \Phi\left(\frac{d_i\sqrt{\sigma_i^2+\sigma_j^2}-y}{\sigma_j}\right) \quad d\Phi\left(\frac{y}{\sigma_i}\right) = 1-\alpha$

Notice that the d_i defined in the Gupta-rule is the $1-\alpha$ quantile of

$$\max_{j\neq i}\left(\frac{Y_i - Y_j}{\sqrt{\sigma_j^2+\sigma_i^2}}\right)$$

where $Y_j \sim N(0,\sigma_j^2)$. These quantiles can be evaluated numerically. If $\sigma_i \equiv \sigma$, then

$$d_i = q_{1-\alpha,m}\sigma$$

where $q_{1-\alpha,m}$ are the Gupta quantiles, i.e. the solutions of

$$\int \Phi^{m-1}(q\sqrt{2}-y)\, d\Phi(y) = 1-\alpha.$$

A table of these quantiles can be found in Gibbons et al. (1977).

It is clear that the Gupta quantiles $q_{1-\alpha,m}$ are not larger than the Bonferroni quantiles $\Phi^{-1}(1-\frac{\alpha}{m-1})$. Here is again the argument: Let $Y_j \sim N(0,1)$.

$$\begin{aligned}\alpha &= \mathbb{P}\left(\max_{j>1}\left(\frac{Y_1-Y_j}{\sqrt{2}}\right) \geq q_{1-\alpha,m}\right) = \mathbb{P}\left(\bigcup_{j=2}^{m}\left\{\frac{Y_1-Y_j}{\sqrt{2}} \geq q_{1-\alpha,m}\right\}\right)\\ &\leq \sum_{j=2}^{m}\mathbb{P}\left(\frac{Y_1-Y_j}{\sqrt{2}} \geq q_{1-\alpha,m}\right) = (m-1)(1-\Phi(q_{1-\alpha,m}))\end{aligned}$$

which implies that

$$\Phi(q_{1-\alpha,m}) \leq 1-\frac{\alpha}{m-1}$$

and therefore

$$q_{1-\alpha,m} \leq \Phi^{-1}(1-\frac{\alpha}{m-1}).$$

In Figure 1 we have depicted a comparison of the Bonferroni-Quantiles $\Phi^{-1}(1-\frac{\alpha}{m-1})$ and the Gupta-Quantiles $q_{1-\alpha,m}$ for $m = 2,\ldots,50$ and $\alpha = 0.1$. The Gupta quantiles are always smaller and one may show that they are asymptotically (as $m \to \infty$) smaller by an order of magnitude.

Lemma 2. The Gupta-rule defines a level-α confidence set in the sense of Definition 1.

Proof. Let $Y_j = Z_j - \mu_j \sim N(0,\sigma_j^2)$ Suppose that $\mu_i = \min_j \mu_j$. It is easy to see that

$$\begin{aligned}\mathbb{P}(Z_i \leq \min\{Z_j + d_i\sqrt{\sigma_i^2+\sigma_j^2} : j\neq i\}) &\geq \mathbb{P}(Y_i \leq \min\{Y_j + d_i\sqrt{\sigma_i^2+\sigma_j^2} : j\neq i\})\\ &= \mathbb{E}[\mathbb{P}(y \leq \min\{Y_j + d_i\sqrt{\sigma_i^2+\sigma_j^2}\})|Y_i = y)]\\ &= \int\prod_{j\neq i}(1-\Phi\left(\frac{y-d_i\sqrt{\sigma_i^2+\sigma_j^2}}{\sigma_j}\right)) \quad d\Phi\left(\frac{y}{\sigma_i}\right)\\ &= \int\prod_{j\neq i}\Phi\left(\frac{d_i\sqrt{\sigma_i^2+\sigma_j^2}-y}{\sigma_j}\right) \quad d\Phi\left(\frac{y}{\sigma_i}\right) = 1-\alpha\end{aligned}$$

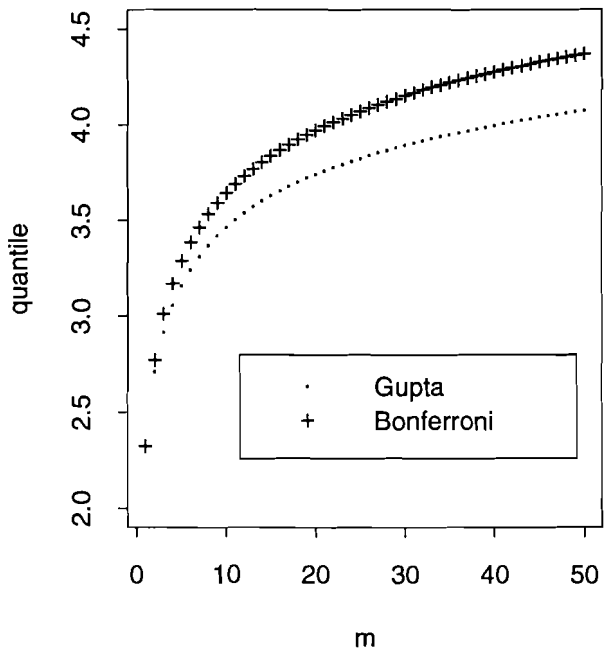


Figure 1: Bonferroni and Gupta quantiles for $\alpha = 0.1$

□

In the case of unequal variances, the calculation of the m different quantiles d_i needs m different numerical integrations. One may confine oneself to one numerical integration by noticing that the selection rule keeps its level, if d_i is replaced by $\max_j d_j$. Gupta and Huang (1976) have shown that the maximal d_i is obtained for that i which has minimal variance σ_i^2. We call the pertaining selection method the Gupta-Huang Rule.

GUPTA-HUANG RULE

Select population i if $Z_i \leq \min\{Z_j + d\sqrt{\sigma_i^2 + \sigma_j^2} : j \neq i\}$
where
d is the solution of
$\int \prod_{j \neq i^*} \Phi\left(\frac{d\sqrt{\sigma_{i^*}^2 + \sigma_j^2} - y}{\sigma_j}\right) \quad d\Phi\left(\frac{y}{\sigma_{i^*}}\right) = 1 - \alpha$
and $\sigma_{i^*}^2 = \min_j \sigma_j^2$.

The Bonferroni-Rule, the Gupta-Rule and the Gupta-Huang-Rule differ only in the specification of the quantiles d_i, the general structure is the same. Therefore, we may calculate by the same formula the expected size S of the confidence set.

Lemma. Let $C = \{i : Z_i \leq \min\{Z_j + d_i\sqrt{\sigma_i^2 + \sigma_j^2} : j \neq i\}\}$ be a confidence set. If $Z_j \sim N(\mu_j, \sigma_j^2)$, then the expected size S of C is

$$S(\mu_1, \ldots, \mu_m, \sigma_1^2, \ldots \sigma_m^2) = \sum_{i=1}^{m} \int \prod_{j \neq i} \Phi\left(\frac{d_i\sqrt{\sigma_i^2 + \sigma_j^2} + \mu_j - \mu_i - y}{\sigma_j}\right) \quad d\Phi\left(\frac{y}{\sigma_i}\right). \qquad (3)$$

An upper bound for this size is given by

$$\bar{S}(\mu_1,\ldots,\mu_m,\sigma_1^2,\ldots\sigma_m^2) = 1+\sum_{i\neq j^*}\Phi\left(\frac{\mu_{j^*}-\mu_i}{\sqrt{\sigma_i^2+\sigma_{j^*}^2}}+d_i\right). \tag{4}$$

where j^* is such that $\mu_{j^*} = \min_i \mu_i$.

Proof. Notice that the expected size is $\sum_{i=1}^m \mathbb{P}(Z_i \leq \min\{Z_j + d_i\sqrt{\sigma_i^2+\sigma_j^2} : j \neq i\})$. Formula (3) is the evaluation of this expression in terms of an integral.

The upper bound (4) is obtained by

$$\begin{aligned}\sum_{i=1}^m \mathbb{P}(Z_i \leq \min\{Z_j + d_i\sqrt{\sigma_i^2+\sigma_j^2}\}) &\leq \sum_{i=1}^m \mathbb{P}(Z_i \leq Z_{j^*} + d_i\sqrt{\sigma_i^2+\sigma_{j^*}^2})\\ &\leq 1+\sum_{i\neq j^*}\mathbb{P}(Y_i+\mu_i \leq Y_{j^*}+\mu_{j^*}+d_i\sqrt{\sigma_i^2+\sigma_{j^*}^2})\\ &= 1+\sum_{i\neq j^*}\Phi\left(\frac{\mu_{j^*}-\mu_i}{\sqrt{\sigma_i^2+\sigma_{j^*}^2}}+d_i\right).\end{aligned}$$

□

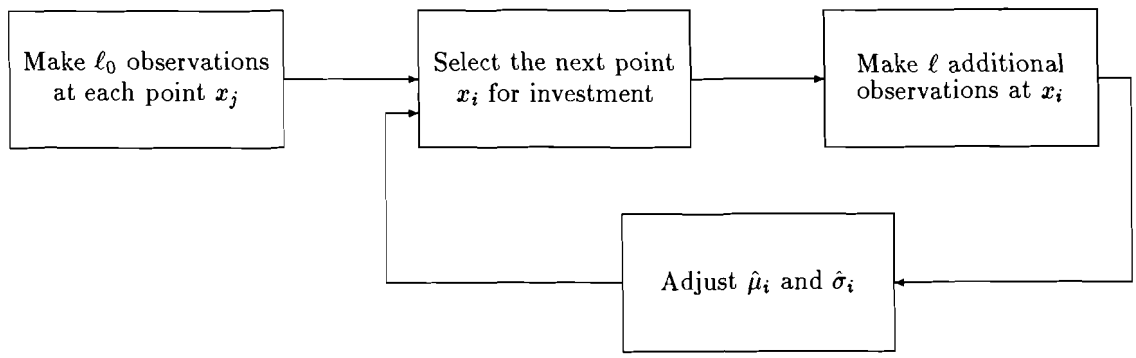
The bound is valid for any choice of j^*, however, the best bound is obtained if j^* is such that $\mu_{j^*} = \min_i \mu_i$.

3 Adaptive design of simulation experiments

We now turn back to the original problem of setting up an experimental plan: Based on a selection rule, we design an algorithm which decides in which parameter one should invest the next portion of simulations.

After an initial warmup phase, the assignment rule is repeated for each new portion of (say) ℓ simulations. After the new observations are obtained, the information is updated and the new assignment is based on the this new information.

The general structure of an adaptive rule is as follows:



Suppose that at each parameter point x_i we have already observed n_i i.i.d. random variables

$$X_{i,k} \sim N(F(x_i),\sigma_i^2); \qquad k = 1,\ldots,n_i.$$

We may estimate the unknown μ_i's and σ_i's by

$$\hat{\mu}_i = \frac{1}{n_i}\sum_{k=1}^{n_i} X_{i,k} \sim N(F(x_i), \frac{\sigma_i^2}{n_i}).$$

$$\hat{\sigma}_i^2 = \frac{1}{n_i}\sum_{k=1}^{n_i}(X_{i,k} - \hat{\mu}_i)^2.$$

Suppose we are allowed to make ℓ new observations. In which parameter should we invest and make the observations there? As was explained in the Introduction, we compare the expected sizes of confidence sets. Suppose $\hat{\mu}_j$ and $\hat{\sigma}_j$ are the *true* parameters: If we invest ℓ new observations in parameter x_i, then we could base the next confidence set on observations with distributions

$$\begin{array}{ll} N(\hat{\mu}_j, \frac{\hat{\sigma}_j^2}{n_j}) & \text{if } j \neq i \\ N(\hat{\mu}_i, \frac{\hat{\sigma}_i^2}{n_i+\ell}) & \text{if } j = i \end{array} \tag{5}$$

The assignment rule selects x_i, if the expected size of the confidence set calculated with parameters (5) is smallest.

GENERAL ASSIGNMENT RULE

Let $\hat{\mu}_i$, $\hat{\sigma}_i$, n_i be given.
Calculate for every i the expected size of the confidence set according to formula (3):
$S_i = S\left(\hat{\mu}_1, \ldots, \hat{\mu}_i, \ldots, \hat{\mu}_m, \frac{\hat{\sigma}_1^2}{n_1}, \ldots, \frac{\hat{\sigma}_i^2}{n_i+\ell}, \ldots, \frac{\hat{\sigma}_m^2}{n_m}\right)$
Make the next ℓ observations for that i which gives the smallest value S_i.

This application of this rule relatively complicated, since many numerical integrations have to be carried out.

A simplified rule works with the upper bound of the expected size $\bar{S}$ instead of S itself and avoids numeric integration. By using the upper bound we are always on the safe side: If $\bar{S}$ is small than necessarily S is small too.

Instead of minimizing $\bar{S}\left(\hat{\mu}_1, \ldots, \hat{\mu}_i, \ldots, \hat{\mu}_m, \frac{\hat{\sigma}_1^2}{n_1}, \ldots, \frac{\hat{\sigma}_i^2}{n_i+\ell}, \ldots, \frac{\hat{\sigma}_m^2}{n_m}\right)$ one might as well minimize the difference $\bar{S}\left(\hat{\mu}_1, \ldots, \hat{\mu}_i, \ldots, \hat{\mu}_m, \frac{\hat{\sigma}_1^2}{n_1}, \ldots, \frac{\hat{\sigma}_i^2}{n_i+\ell}, \ldots, \frac{\hat{\sigma}_m^2}{n_m}\right) - \bar{S}\left(\hat{\mu}_1, \ldots, \hat{\mu}_m, \frac{\hat{\sigma}_1^2}{n_1}, \ldots, \frac{\hat{\sigma}_m^2}{n_m}\right)$. This leads to the following rule:

SIMPLIFIED ASSIGNMENT RULE 1

Let $\hat{\mu}_i$, $\hat{\sigma}_i$, n_i be given.
Set $\hat{\mu}_{j^*} = \min_j \hat{\mu}_j$. Calculate for every $i \neq j^*$

$$D_i = \Phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i+\ell} + \frac{\hat{\sigma}_{j^*}^2}{n_i+\ell}}} + d_i\right) - \Phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_i}}} + d_i\right)$$

and

$$D_{j^*} = \sum_{i \neq j^*} \Phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}+\ell}}} + d_i\right) - \Phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_i}}} + d_i\right)$$

Assign the next ℓ observations to the population with minimal D_i.

Instead of calculating the difference of two expected sample sizes, one could as well work with derivatives w.r.t. ℓ and therefore avoid the use of the normal d.f. Φ, which is not available on every computer. Notice however that by basing the decision only on derivatives, we find out the parameter x_i in which we should invest in tendency and therefore the approximation is good only for small values of ℓ.

Let, for $i \neq j^*$

$$D_i' = \frac{\partial}{\partial \ell}\Phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i+\ell} + \frac{\sigma_{j^*}^2}{n_{j^*}}}} + d_i\right)\Bigg|_{\ell=0} = \phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}}} + d_i\right)(\hat{\mu}_{j^*} - \hat{\mu}_i)\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}\right)^{-3/2}\frac{\hat{\sigma}_i^2}{n_i^2}$$

and

$$\begin{aligned} D_{j^*}' &= \frac{\partial}{\partial \ell}\sum_{i \neq j^*}\Phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\sigma_{j^*}^2}{n_{j^*}+\ell}}} + d_i\right)\Bigg|_{\ell=0} \\ &= \sum_{i \neq j^*}\phi\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}}} + d_i\right)(\hat{\mu}_{j^*} - \hat{\mu}_i)\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}\right)^{-3/2}\frac{\hat{\sigma}_j^2}{n_j^2} \end{aligned}$$

where ϕ is the normal density $\phi(u) = \frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}u^2)$.

SIMPLIFIED ASSIGNMENT RULE 2

Let $\hat{\mu}_i$, $\hat{\sigma}_i$, n_i be given.
Set $\hat{\mu}_{j^*} = \min_j \hat{\mu}_j$. Calculate for every $i \neq j^*$

$$D_i' = \exp\left(-\frac{1}{2}\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}}} + d_i\right)^2\right)(\hat{\mu}_{j^*} - \hat{\mu}_i)\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}\right)^{-3/2}\frac{\hat{\sigma}_i^2}{n_i^2}$$

and

$$D_{j^*}' = \sum_{i \neq j^*}\exp\left(-\frac{1}{2}\left(\frac{\hat{\mu}_{j^*} - \hat{\mu}_i}{\sqrt{\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}}} + d_i\right)^2\right)(\hat{\mu}_{j^*} - \hat{\mu}_i)\left(\frac{\hat{\sigma}_i^2}{n_i} + \frac{\hat{\sigma}_{j^*}^2}{n_{j^*}}\right)^{-3/2}\frac{\hat{\sigma}_j^2}{n_j^2}$$

Assign the next ℓ observations to the population with minimal D_i'.

4 An Example

Consider 10 populations with means μ_i and standard deviations σ_i given by Figure (2).

The double triangles indicate the mean μ_i as the center and $\mu_i + \sigma_i$ resp. $\mu_i - \sigma_i$ as upper resp. lower peak of the triangles. We see that population 1 has the lowest mean, but largest variance.

Initially, we have made $\ell_0 = 20$ observations for each parameter value. Then portions of $\ell = 10$ observations are assigned according to the following rules: Non-adaptive Rule (equal portion of ℓ/m observations for each population – here $\ell/m = 1$), Simplified Assignment Rule 1 and Simplified Assignment Rule 2. Then we have compared the expected sizes of the confidence sets for $\alpha = 0.1$. Here are the results:

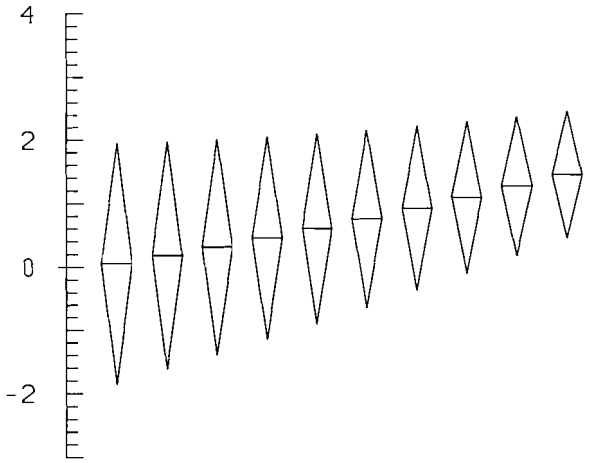


Figure 2: A μ-σ-diagram

mean size of the 0.1-level confidence set	Non-adaptive design	Assignment Rule 1	Assignment Rule 2
after 200 obs.	6.77	6.77	6.77
after 400 obs.	5.77	5.02	5.00
after 600 obs.	5.15	4.28	4.34
after 800 obs.	4.78	3.85	3.92
after 1000 obs.	4.56	3.49	3.60

Table 1.

One sees that both adaptive assignment rules are about of the same quality and are considerably better than the non-adaptive design.

The number of observations taken at each parameter point by the adaptive assignment rules are very different. The following table shows the mean number of observations for the simplified rule 2:

mean number of observ. in group i	1	2	3	4	5	6	7	8	9	10
among first 200 obs.	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0	20.0
among first 400 obs.	79.6	59.5	50.3	42.2	38.2	34.1	29.3	24.5	21.6	20.6
among first 600 obs.	128.9	97.7	84.9	68.4	60.8	49.5	37.7	28.4	22.7	21.0
among first 800 obs.	189.5	134.9	117.9	100.0	80.0	60.2	42.3	30.2	23.7	21.2
among first 1000 obs.	253.7	174.5	155.3	133.0	95.8	66.2	45.1	31.0	24.0	21.3

Table 2.

Notice that more and more observations are taken from the "better" points. As a by-product, one gets therefore a 10 times more accurate estimate for the minimal value $\min_x F(x)$ as compared to the non-adaptive design with the same total number of observations.

We remark here that it might be advisable to use different α's for the assigment rule and for the final estimate of the confidence interval. The α for the assignment rule may be chosen quite large (as 0.3), the α for the confidence level should not be larger than 0.1.

5 Discussion

Let us compare our approach to some seemingly related approaches, the *multi-armed bandit problem* and Ho's *ordinal optimization.*

The name multi-armed bandit comes from the problem of finding the best slot machine out of a finite number of such machines. The player does not know which machine provides the highest payoff. Therefore, he has to try all machines to find out the one with the highest winning chance. However, it is advisable not to play all machines with the same frequency: The machine with the presumably highest payoff should be played as often as possible.

Multi-armed bandit processes are also successfully applied in biostatistics. Here the problem is to find the best out of a finite number of drugs. The drugs are applied to patients and the experimental plan should be designed such that the best drug is found w.pr.1 and the number of patients who do not get the best drug should be minimized.

The theory of multi-armed bandit processes shows that there is a method of designing the experiment in such a manner that for $N \to \infty$

- The best machine is found almost surely.
- The relative number of experiments conducted with the best machine converges to 1.

In particular, Lai and Robbins (1985) show the following: Let m populations be given, each with probability density $f(x, \theta_i)$; $i = 1, \ldots m$. Let μ_i be the mean of density $f(x, \theta_i)$ and let μ_{i^*} be the best mean $\mu_{i^*} = \max_j \mu_j$. Suppose the sum of outcomes according to some sampling plan after n observations is $S_n = \sum_{i=1}^{n} X_i$. Call $R_n = n\mu_{i^*} - \mathbb{E}(S_n)$ the *regret.*

If $R_n = o(n^a)$ for every $a > 0$, then necessarily the regret obeys the following lower bound

$$\liminf_{n\to\infty} \frac{R_n}{\log n} \geq \sum_{\mu_j < \mu_{i^*}} \frac{\mu_{i^*} - \mu_j}{I(\theta_j, \theta_{i^*})},$$

where $I(\theta_j, \theta_{i^*})$ is the Kullback-Leibler entropy.

Lai and Robbins construct an allocation rule that achieves the lower bound and call this rule *efficient.*

Notice that efficiency is only an asymptotic property. By changing the rule for finitely many observations, the efficiency is not changed.

If the criterion to be minimized is not the regret, but the discounted sum $\sum \beta^i X_i$, the the best allocation rule must be based on dynamic allocation indices (*Gittins indices*) – see Gittins (1979).

It is important to see that our discrete optimization problem is different in nature from the multi-armed bandit problem. In our problem every observation costs the same and the objective is only to find the minimizer with as few observations as possible. We can make many observations at non-optimal points if we consider them as necessary with no extra costs.

In contrast, the costs of the multi-armed bandit are proportional to the payoff and therefore it is important to play the presumably best machine as often as possible. It is not advisable to use rules from bandit-processes for the optimal design of simulation experiments.

Ho's ordinal optimization method (Ho, Sreenivas and Vakili (1992)) is more in the spirit of our approach. They however consider only confidence sets which are made up from the best estimates, i.e.

$$\tilde{C} = \{Z_i : Z_i \text{ is among the } k \text{ smallest of the } Z_j's\} \tag{6}$$

The method consists in monitoring those populations which enter or leave the sets $\tilde{C}$ as the simulation goes on.

Notice that the correct confidence sets as described in section 2 depend both on the means and the variances of the populations. If all variances are equal, then all confidence sets are of the form (6). If however not all variances are equal, then the correct confidence sets are typically not of this form, since it is likely that a population with some mean and small variance is excluded, whereas another population with larger mean but large variance is not. Thus order intervals of the form (6) are not confidence sets in the sense of Definition 1 unless k is very large.

References

[1] Anantharam ., Varaiya P. (1986). Asymptotically Efficient Rules in Multiarmed Bandit Problems. LN in Control and Information Sciences 105, Springer Verlag

[2] Berger R.L., Gupta S.S. (1979). Minimax subset selection rules with applications to unequal variance (unequal sample size) problems. Scand. J. Statist 7, 21 – 26

[3] Berry D.A., Fristedt B. (1985). Bandit problems. Chapman and Hall, London

[4] Gibbons J.D., Olkin I., Sobel M. (1977). Selecting and Ordering Populations. Wiley, New York

[5] Gittins J.C. (1979). Bandit processes and dynamic allocation indices. J. Roy. Statist. Soc. 41, 148 – 177

[6] 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

[7] Gupta S. S., Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and methodology of selecting and ranking populations. Wiley, New York.

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

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

[10] Norkin V., Ermoliev Yu., Ruszczynski A. (1994). Optimal allocation of indivisibles under uncertainty. IIASA working paper WP-94-21