An Open Optimization Problem in Statistical Quality Control

W. SEIDEL
Universität der Bundeswehr Hamburg, Postfach 70 08 22, 2000 Hamburg 70, Germany

(Received: 14 September 1990; accepted: 18 June 1991)

Abstract. We ask the experts in global optimization if there is an efficient solution to an optimization problem in acceptance sampling: Here, one often has incomplete prior information about the quality of incoming lots. Given a cost model, a decision rule for the inspection of a lot may then be designed that minimizes the maximum loss compatible with the available information. The resulting minimax problem is sometimes hard to solve, as the loss functions may have several local maxima which vary in an "unpredictable" way with the parameters of the decision rule.

Key words. Minimax problems, acceptance sampling, incomplete prior information, generalized moment conditions.

1. Introduction

A classical field of quality control is acceptance sampling, it deals mainly with the following problem: Given an "inspection lot" of items of the same product, one has to decide whether to accept or reject the lot on the basis of a random sample.

A good decision rule has to find some balance between the cost of wrong decisions and the cost of gaining more information about the lot. Let us assume that the quality of the product is judged by a single characteristic ξ, which is not a constant but a random variable with a probability distribution that depends on some parameter x. When a lot is inspected, the actual value of x is unknown; let X be the space of possible values of x. A conservative approach is to choose a strategy that minimizes the maximum value of the expected loss, where the maximum is taken over X.

Frequently such a minimax strategy is criticized because it protects against unlikely situations. If x itself varies from lot to lot according to a probability distribution π, and π is known, a "Bayes strategy" may be used that minimizes the integral of the loss with respect to π.

Whereas the minimax strategy is usually too pessimistic, this Bayes strategy may be too optimistic. It seems more realistic to assume that it is only known that π belongs to a certain set Π of probability measures on X. For instance, Krumbholz (1982) and v. Collani (1986) consider the set of all probability measures that put mass of (at least) γ to a certain area of X.

Given "incomplete information" Π, a strategy is appropriate that minimizes the maximum average loss, where the maximum is taken over all $\pi \in \Pi$. Using Choquet theory, maximization over Π may in certain cases be reduced to a maximization problem in several real variables. But the corresponding minimax problem may still be hard to solve, especially in realistic situations. The reason is, that the goal function often has several local maxima that vary in a complicated way with the sampling strategy.

In the following, a collection of such problems will be presented, coming from a particular stochastic model of acceptance sampling which is described in the appendix. The question is: Is there an efficient numerical algorithm which solves as many of these problems as possible? Applying such an algorithm should frequently help to reduce costs of acceptance sampling considerably.

2. The Regret Function

A loss function which is basic for the optimization problems is presented in this section; the details of the model may be found in the appendix.

The decision rules considered here will be indexed by a pair $(n, c) \in \mathbb{N} \times \mathbb{R}$: a sample of size n is taken from the lot, and a particular test statistic t_n is calculated. The lot is accepted iff t_n does not exceed the acceptance number c. Given n, t_n is assumed to be fixed, so (n, c) completely characterizes the decision rule. The loss is then a random variable, depending on the decision.

Let $x \in \mathbb{R}$ be the parameter of the stochastic model and let the fraction p of defective items be related to x according to $p = \Phi(x)$, where Φ denotes the cumulative distribution function of the standard normal distribution defined by

$$\Phi(x) = (2\pi)^{-1/2} \int_{-\infty}^{x} \exp(-t^2/2)\, \mathrm{d}t\,.$$

A variety of linear cost models may be reduced to a normalized version that depends only on two parameters, a "break even quality" $p_0 \in]0, 1[$ and the relative cost $q > 0$ of inspecting one item (see appendix). Let $x_0 := \Phi^{-1}(p_0)$. The expected value of the loss, when the lot is inspected by (n, c), depends on the model parameter x and is given by

$$R(x; n, c) = nq + \begin{cases} (\Phi(x) - \Phi(x_0))\Phi(\sqrt{n}(c - x))\,, & \text{if } x \geq x_0 \\ (\Phi(x_0) - \Phi(x))(1 - \Phi(\sqrt{n}(c - x)))\,, & \text{if } x < x_0\,. \end{cases}$$

Given (n, c), the function $R(\cdot\,; n, c)\colon \mathbb{R} \to [0, 1]$ is called a "regret function". Obviously, $R(x_0; n, c) = nq = \lim_{x\to-\infty} R(x; n, c) = \lim_{x\to\infty} R(x; n, c)$ and $R(x; n, c) > nq$ for $x \neq x_0$. Basler (1967/68) has proved that $R(\cdot; n, c)$ is strictly unimodal on the left and on the right of x_0; the maximum value of $R(\cdot; n, c)$ in $]-\infty, x_0[$ ($]x_0, \infty[$) will be denoted by $R_l(n, c)$ ($R_r(n, c)$). Finally, let $R_m(n, c) := \max\{R_l(n, c), R_r(n, c)\}$, the maximum value of the regret function. The graph of the regret function for $p_0 = 0.05$, $q = 10^{-4}$ and $(n, c) = (20, -1.6)$ is shown in Figure 1.

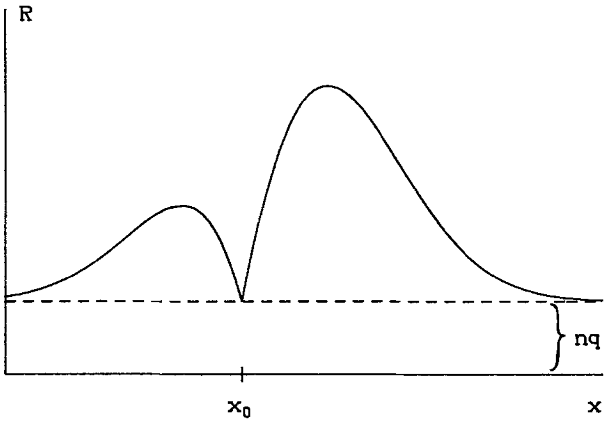

Fig. 1.

A FIRST MINIMAX PROBLEM

Let (n_0, c_0) be defined by

$$R_m(n_0, c_0) = \min_{n,\,c} R_m(n, c)\,.$$

The decision rule (n_0, c_0) minimizes the maximum loss; it is appropriate, if nothing is known about x in advance.

Basler (1967/68) proposes a solution which uses the fact that, for a fixed n, $R_l(n, c)$ is continuous and strictly antitonic and $R_r(n, c)$ is continuous and strictly isotonic in c. So for every $n \in \mathbb{N}$ there is exactly one c_n with

$$R_l(n, c_n) = R_r(n, c_n)\,;$$

it minimizes $R_m(n, c)$. Finally, (n_0, c_0) may be found by minimizing $R_l(n, c_n)$.

For example, let $p_0 = 0.05$ and $q = 10^{-4}$. Then $x_0 = -1.645$, $(n_0, c_0) = (20, -1.668)$ and $R_m(n_0, c_0) = 0.0059$.

Although calculating (n_0, c_0) in this way can be done very quickly, it would be advantageous to have a more efficient algorithm.

3. Prior Information Given by Conditions on Generalized Moments

In the following, $\mathbb{R}$ will always be endowed with the Borel σ-algebra. "Measurable" means Borel-measurable.

Let us assume that the model parameter varies from lot to lot according to a probability distribution π, then

$$r(\pi; n, c) := \int_{\mathbb{R}} R(x; n, c)\,\pi(\mathrm{d}x)$$

is called the "Bayes risk" of (n, c), given π. It is the expected value of the loss.

Let Γ denote the set of all probability measures on $\mathbb{R}$ and consider prior information $\Pi \subseteq \Gamma$ as discussed in the introduction. Then it is natural to choose a decision rule (n_Π, c_Π) that satisfies

$$\sup_{\pi \in \Pi} r(\pi; n_\Pi, c_\Pi) = \min_{n, c} \sup_{\pi \in \Pi} r(\pi; n, c) . \tag{1}$$

In the following, we shall assume that Π is given by a set of k "generalized moment conditions", i.e., for each $i = 1, \ldots, k$ there exists a measurable function f_i on $\mathbb{R}$ and a real number $c_i \in \mathbb{R}$ such that

$$\Pi = \{\pi \in \Gamma \colon f_i \text{ is } \pi\text{-integrable and} \int_{\mathbb{R}} f_i \, \mathrm{d}\pi \geq c_i , \quad i = 1, \ldots, k\} .$$

REMARK. Replacing f_i by $-f_i$ also allows us to treat constraints of the form $\int_{\mathbb{R}} f_i \, \mathrm{d}\pi \leq c_i$ and $c_i \leq \int_{\mathbb{R}} f_i \, \mathrm{d}\pi \leq d_i$, $c_i \leq d_i$.

Winkler (1982; 1988) proves that for Π given as above, the problem of maximizing the integral of a real function with respect to π, $\pi \in \Pi$, reduces to a maximization problem in $\mathbb{R}^l$, $l \leq 2k + 2$: Clearly, Π is convex; Winkler proves that the set of its extreme points is a subset of Π_0 defined by

$$\Pi_0 \colon = \Big\{ \pi \in \Pi \colon \pi = \sum_{i=1}^{m} t_i \delta_{x_i} , \quad t_i > 0 , \quad \sum_{i=1}^{m} t_i = 1, x_i \in \mathbb{R} , \quad 1 \leq m \leq k+1 ,$$
$$\text{the vectors } (1, f_1(x_i), \ldots, f_k(x_i)) , \quad 1 \leq i \leq m \text{ are linearly independent} \Big\} .$$

Here δ_x denotes the "Dirac measure" which puts mass 1 to $x \in \mathbb{R}$.

Given a function $g \colon \mathbb{R} \to \mathbb{R}$, integrable for every $\pi \in \Pi$, Winkler then proves that

$$\sup\Big\{\int_{\mathbb{R}} g \, \mathrm{d}\pi : \pi \in \Pi\Big\} = \sup\Big\{\int_{\mathbb{R}} g \, \mathrm{d}\pi : \pi \in \Pi_0\Big\} .$$

Application of Winkler's results shows, that $\sup_{\pi \in \Pi} r(\pi; n, c)$ is the solution to the following problem:

$$\begin{aligned} &\underset{t_j,\, x_j \in \mathbb{R}}{\text{Maximize}} \quad \sum_{j=1}^{k+1} t_j R(x_j; n, c) \\ &\text{subject to} \quad t_j \geq 0 , \quad 1 \leq j \leq k + 1 \\ &\qquad \sum_{j=1}^{k+1} t_j = 1 , \\ &\qquad \sum_{j=1}^{k+1} t_j f_i(x_j) \geq c_i , \quad 1 \leq i \leq k . \end{aligned}$$

Simplifications may be possible because of linear dependency.

EXAMPLE 3.1. Given $a \in \mathbb{R}$ and $0 < \gamma < 1$, let $k = 1$, $c_1 \colon = \gamma$ and f_1 be defined by $f_1(x) = 1$, if $x \leq a$, and $f_1(x) = 0$ otherwise. The Π corresponds to the inequality

$$\mathbb{P}(x \leqslant a) \geqslant \gamma \,. \tag{2}$$

The corresponding minimax problem may still be solved by using elementary methods, see Krumbholz and Schröder (1987). With the notation

$$ml(a; n, c) := \max_{x \leqslant a} R(x; n, c)$$

it holds true that

$$\sup_{\pi \in \Pi} r(\pi; n, c) = \gamma ml(a; n, c) + (1-\gamma) R_m(n, c) \,,$$

an expression which is easy to evaluate. Given n, this supremum is continuous and quasiconvex in c (Seidel, 1990b), so minimization in c is easy.

Let $R_\Pi := \sup_{\pi \in \Pi} r(\pi; n_\Pi, c_\Pi)$. For the parameters $p_0 = 0.05$, $q = 10^{-4}$ and $\gamma = 0.9$, we obtained for instance the following decision rules:

$$a = \Phi^{-1}(0.03): (n_\Pi, c_\Pi) = (16, -1.43),\ R_\Pi = 0.00369$$

$$a = \Phi^{-1}(0.05): (n_\Pi, c_\Pi) = (14, -1.405),\ R_\Pi = 0.00393 \,.$$

Again, an efficient method to solve the minimax problem is of interest, although the range of application of 3.1 is limited. The reason is, that one does not usually observe x directly, but an estimator of x with a different probability distribution, so the parameters of (2) are sometimes not easy to estimate from past inspections (Seidel, 1990a). A way out of this problem is to replace x in (2) by an observable estimator; this will be done in the next example.

EXAMPLE 3.2. Usually x is estimated by the mean value $\bar{x}_l$ of a sample of size l. In Seidel (1990a) it is shown that the information

$$\mathbb{P}(\bar{x}_l \leqslant a) \geqslant \gamma \tag{3}$$

can be described by a generalized moment condition with $k = 1$, $f_1(x) = \Phi(\sqrt{l}(a - x))$ and $c_1 = \gamma$. The parameters of (3) can easily be estimated from past inspections. To obtain typical values, let $p_0 = 0.05$ and $q = 10^{-4}$ and assume that a series of lots has been inspected by the minimax strategy $(n_0, c_0) = (20, -1.668)$. We can assume without loss of generality that the test statistic is $\bar{x}_l$ with $l = n_0 = 20$ (see appendix). Let $a = c_0 = -1.668$, then γ may be estimated by the fraction of accepted lots (possibly with a correction that protects against estimation errors). A reasonable value would be $\gamma = 0.9$.

In order to consider (3), we have to calculate

$$M := \max_{t,x,y \in \mathbb{R}} tR(x; n, c) + (1-t)R(y; n, c)$$

subject to

(i) $0 \leqslant t \leqslant 1$

(ii) $t\Phi(\sqrt{l}(a - x)) + (1-t)\Phi(\sqrt{l}(a - y)) \geqslant \gamma$.

Because of the symmetry in x and y, we may choose $x \leq y$. Obviously,

$$\Phi(\sqrt{l}(a-x))(\overset{>}{=})\gamma \Leftrightarrow x(\overset{<}{=})a(\Phi^{-1}(\gamma))/\sqrt{l} =: z.$$

If $y \leq z$, then (ii) is satisfied for any $t \in [0, 1]$.

If $(x, y) \in A := \{(x, y): x \leq z < y\}$ then (ii) is satisfied for any $t \in [t_{x,y}, 1]$, $t_{x,y}$ being defined as $t_{x,y} := (\gamma - \Phi(\sqrt{l}(a-y)))/(\Phi(\sqrt{l}(a-x)) - \Phi(\sqrt{l}(a-y)))$.

If $R(\cdot; n, c)$ takes its global maximum at some point less or equal to z, then $M = R_m(n, c)$.

Now assume that $R(x; n, c) < R_m(n, c)$ for all $x \leq z$. In this case,

$$M = \max_{(x,\,y) \in A} F(x, y)$$

holds, where $F(x, y)$ is defined by

$$F(x, y) := \max_{t_{x,y} \leq t \leq 1} t(R(x; n, c) - R(y; n, c)) + R(y; n, c).$$

Obviously,

$$F(x, y) = \begin{cases} R(x; n, c), & \text{if } R(x; n, c) \geq (y; n, c) \\ t_{x,y}(R(x; n, c) - R(y; n, c)) + R(y; n, c), & \text{otherwise}. \end{cases}$$

So the maximization problem reduces to an optimization problem in two real variables with simple constraints.

Choosing the parameters as in the beginning of this example, we obtained the decision rule $(n_{\Pi}, c_{\Pi}) = (13, -1.49)$ together with the corresponding maximum loss $M = 0.00269$. Figures 2–4 show typical surfaces of the goal function F (note that $z = -1.955$).

EXAMPLE 3.3. Let the i-th moment of π be defined by $\int_{\mathbb{R}} x^i \pi(\mathrm{d}x)$. The requirement, that the i-th moment of π lies within a certain closed real interval I_i, $i = 1, \ldots, k$, is a (classical) moment condition: in fact, generalized moment

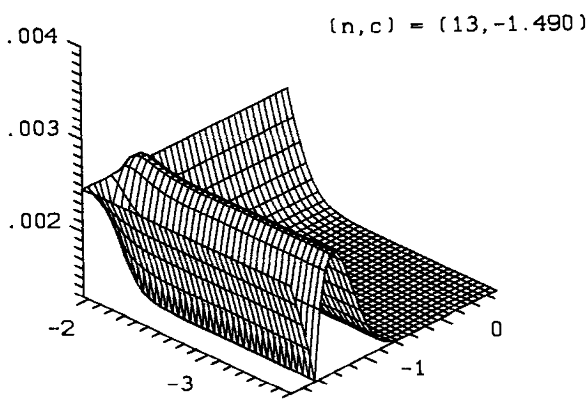

Fig. 2.

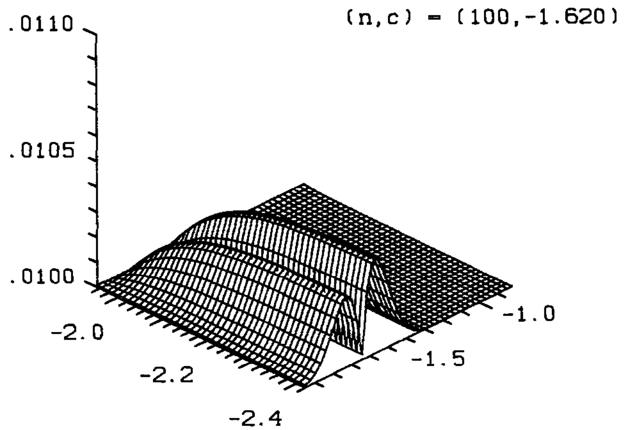

Fig. 3.

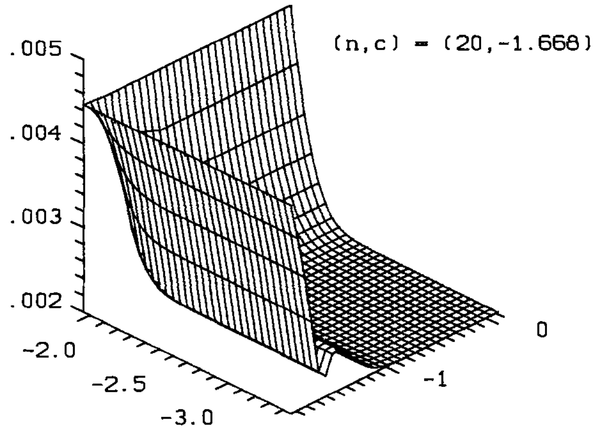

Fig. 4.

conditions are generalizations of it. For an estimator of moments of π, even if x is not observable, see Seidel (1990a).

In most applications, k will be less or equal to 3.

Even if lower *and* upper bounds for the moments are known, linear dependency yields that the corresponding maximization problem is

$$\begin{aligned} &\underset{t_j,\, x_j \in \mathbb{R}}{\text{Maximize}} \quad \sum_{j=1}^{k+1} t_j R(x_j; n, c) \\ &\text{subject to} \quad t_j \geq 0,\, 1 \leq j \leq k+1 \\ &\qquad\qquad \sum_{j=1}^{k+1} t_j = 1\,, \\ &\qquad\qquad \sum_{j=1}^{k+1} t_j x_j^i \in I_i\,, \quad 1 \leq i \leq k\,. \end{aligned}$$

Appendix: The Details of the Model

The stochastic model presented here can be found in nearly every textbook of acceptance sampling; for a version that also contains the cost model, see Stange (1964) or Basler (1967/68).

Assume that the quality of a certain product is measured by a one-dimensional quality characteristic ξ and that an item is acceptable iff $\xi \leq U$, an upper specification limit. Assume further that ξ is normally distributed with known variance σ^2 and unknown expectation μ. Then

$$p := \mathbb{P}_\mu(\xi > U) = \Phi((\mu - U)/\sigma)$$

is the fraction of defective items, it is bijectively related to μ.

In order to inspect a lot, a random sample of size n is taken and the sample mean $\bar{x}_n$ is calculated. The lot is accepted iff the test statistic $t_n = (\bar{x}_n - U)/\sigma$ does not exceed the acceptance limit c. The pair (n, c) completely characterizes the decision rule. The probability of accepting a lot of quality p is

$$W(p; ,n, c) := \mathbb{P}_p(t_n \leq c) = \Phi(\sqrt{n}(c - \Phi^{-1}(p))) .$$

Let us adopt the linear cost model introduced by Stange (1964), transformed by subtraction of the "unavoidable loss" and a suitable normalization: Given a "break even quality" $p_0 \in]0, 1[$ and relative costs $q > 0$ of inspecting one item, the loss when a lot of quality p is accepted is

$$L_a(p, n) := nq + \max\{p - p_0, 0\} ,$$

whereas the loss when the lot is rejected is

$$L_r(p, n) := nq + \max\{p_0 - p, 0\} .$$

The average loss per lot is

$$R(p; n, c) = L_a(p, n)W(p; n, c) + L_r(p, n)(1 - W(p; n, c))$$
$$= nq + \begin{cases} (p - p_0)\Phi(\sqrt{n}(c - \Phi^{-1}(p))) , & \text{if } p \geq p_0 \\ (p_0 - p)(1 - \Phi(\sqrt{n}(c - \Phi^{-1}(p)))) , & \text{if } p < p_0 . \end{cases}$$

The advantage of this representation of R is its clearness. For statistical purposes, a representation in terms of μ is better: Given a distribution π of μ, the distribution of the observable quantity $\bar{x}_n$ is the convolution on π and a (known) normal distribution with zero expectation and variance σ^2/n. This simple form makes parameters of π easy to estimate.

Finally, using the transformation $x = (\mu - U)/\sigma$, we may without loss of generality assume that $U = 0$ and $\sigma = 1$ holds. Writing $p = \Phi(x)$ ($x = \Phi^{-1}(p)$), we obtain the regret function of Section 2.

References

Basler, H. (1967/68), Bestimmung kostenoptimaler Prüfpläne mittels des Minimax-Prinzips, *Metrika* **12**, 115–154.

v. Collani, E. (1986), The α-Optimal Sampling Scheme, *Journal of Quality Technology* **18**, 63–66.

Krumbholz, W. (1982), Die Bestimmung einfacher Attributprüfpläne unter Berücksichtigung von unvollständiger Vorinformation, *Allgemeines Statistisches Archiv* **66**, 240–253.

Krumbholz, W. and J. Schröder (1987), Zur Ausnutzung unvollständiger Vorinformationen bei der Minimax-Regret-Methode, *Allgemeines Statistisches Archiv* **71**, 117–125.

Seidel, W. (1990a), Minimax Regret Sampling Plans Based on Generalized Moments of the Prior Distribution, Diskussionsbeiträge zur *Statistik und Quantitativen Ökonomie* **39**, Universität der Bundeswehr Hamburg.

Seidel, W. (1990b), On the Performance of a Sampling Scheme in Statistical Quality Control Using Incomplete Prior Information, *Statistical Papers* **31**, 119–130.

Stange, K. (1964), Die Berechnung wirtschaftlicher Pläne für die messende Prüfung, *Metrika* **8**, 48–82.

Winkler, G. (1982), Integral Representation and Upper Bounds for Stop-Loss Premiums under Constraints Given by Inequalities, *Scand. Actuarial J.*, 15–21.

Winkler, G. (1988), Extreme Points of Moment Sets, *Mathematics of Operations Research* **13**, 581–587.