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

Computing Bounds for the Solution of the Stochastic Optimization Problem with Incomplete Information on Distribution of Random Parameters

Gaivoronski, A.A.

IIASA Working Paper

WP-86-072

November 1986

COMPUTING BOUNDS FOR THE SOLUTION OF THE STOCHASTIC OPTIMIZATION PROBLEM WITH INCOMPLETE INFORMATION ON DISTRIBUTION OF RANDOM PARAMETERS

A. Gaivoronski

November 1986
WP-86-72

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

The paper deals with the solution of a stochastic optimization problem under incomplete information. It is assumed that the distribution of probabilistic parameters is unknown and the only available information comes with observations. In addition the set to which the probabilistic parameters belong is also known. Numerical techniques are proposed which allow to compute upper and lower bounds for the solution of the stochastic optimization problem under these assumptions. These bounds are updated successively after the arrival of new observations. The research reported in this paper was performed in the Adaptation and Optimization Project of the Systems and Decision Sciences Program.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

CONTENTS

COMPUTING BOUNDS FOR THE SOLUTION OF THE STOCHASTIC OPTIMIZATION PROBLEM WITH INCOMPLETE INFORMATION ON DISTRIBUTION OF RANDOM PARAMETERS

A. Gaivoronski

1. INTRODUCTION

During recent years considerable effort was made to develop numerical techniques for a solution of stochastic programming problems. These problems can be formulated quite generally as follows:

minimize

$$F(x) = E_\omega f(x, \omega) = \int f(x, \omega) \mathrm{d}H(\omega) \tag{1}$$

subject to constraints

$$x \in X$$

where $x \in R^n$ is a vector of decision variables, $\omega \in \Omega \subset R^m$ is a vector of random variables which belong to some probability space, H – some probability measure and E_ω denotes expectation with respect to ω. Algorithms for solving this problem as well as various applications to operations research and systems analysis can be found in Ermoliev [7], [8], Kall [13], Prekopa [21], Wets [23] [24] where one can find further references. It was assumed usually that probability distribution H is known. This distribution is obtained from the set of observations $\{\omega_1, \ldots, \omega_s, \ldots\}$ of random vector ω through application of appropriate statistical techniques.

However, information contained in observations is often insufficient to identify unique distribution H. Then it is often possible to define some set G to which distribution belong and apply minimax approach (worst case analysis) to the problem (1). This approach was studied by Dupačová [4] [5], Ermoliev [6], Golodnikov [12], Ermoliev, Gaivoronski and Nedeva [9], Birge and Wets [1], Gaivoronski [10] for the case when the set G of admissible distributions is defined by constraints of moment type and by Gaivoronski [11] for the case when the measure H is contained between some upper and lower measures. In this case however some unspecified statistical methods should be used to obtain these bounding measures or moment constraints. The resulting set G as well as the solution of the problem (1) and its accuracy will heavily depend on these techniques.

This paper presents the attempt to solve the problem (1) starting directly from the finite set of observations $\{\omega_1, \ldots, \omega_s\}$. In this case it is possible only to identify some upper and lower bounds on the optimal value of the objective function and point x^* which yields value within these bounds in some probabilistic sense. These bounds should possess the following properties:

- be valid for all s, nonasymptotical;
- be successive i.e. permit easy updates when new observations arrive, this will allow to compute them in real time;
- be independent from any particular parametric family of distributions such as normal, lognormal etc.

This approach, if successfully implemented, would allow to avoid such questions as "where do you get your distribution"? which are often confronted by systems analyst who applies model (1) to real systems. It could also present a "middle road" between a purely stochastic approach, when unknown parameters ω are considered to possess known probability distributions H and deterministic approach when it is assumed that all what is known about ω is that it belongs to some set Ω (for more details see Kurzhanski [17]).

This paper presents only first steps towards this direction. In the section 2 the precise formulation of the problem is developed using certain techniques of extracting information about distribution from observations. This leads to minimax problems with an inner problem which allows an explicit solution defined in the section 3. Finally in the section 4 numerical techniques for computing successive upper bounds is described. Results of some numerical experiments are presented in the course of exposition.

It should be noted that some techniques for computing bounds for solution of the problem (1) in different contexts was described in Birge and Wets [1], Cipra [2], Kall [14], Kankova [15].

2. PROBLEM FORMULATION

Informal statement. Solve the problem

$$\min_{x \in X} \int f(x, \omega) \, \mathrm{d}H(\omega) \tag{2}$$

when all that is known about random parameters ω is the set of observations $\{\omega_1, \ldots, \omega_s\}$.

We shall make two basic assumptions:

1. Distribution H of random parameters ω exists, but unknown
2. Observations $\omega_1, \ldots, \omega_s$ are mutually independent and form the sample from this distribution H.

In order to solve the problem (2) it is necessary to clarify what is considered as solution and learn how to extract information about distribution H from observations ω_i.

Let us assume that ω belongs to some set $\Omega \subset R^m$ with Borel field $\mathbf{B}$; probability measure H is defined on this field, thus we have a probability space $(\Omega, \mathbf{B}, H)$. For each fixed s let us consider the sample probability space $(\bar{\Omega}, \bar{\mathbf{B}}, \bar{\mathbf{P}})$ which is a Cartesian product of s spaces $(\Omega, \mathbf{B}, H)$. The space $(\bar{\Omega}, \bar{\mathbf{B}}, \bar{\mathbf{P}})$ is the smallest space which contain all $(\Omega^s, \mathbf{B}^s, \mathbf{P}^s)$. In what follows the "convergence with probability 1" will mean the "convergence with probability 1 in the space $(\Omega^s, \mathbf{B}^s, \mathbf{P}^s)$". With the set of observations $\{\omega_1, \ldots, \omega_s\}$ the set of distribution G_s will be associated in the following way.

Let us fix the confidence level $\alpha : 0 < \alpha < 1$. We shall consider events with probability $\mathbf{P}^s$ less than α "improbable" events and discard them. Let us consider arbitrary set $A \subset \mathbf{B}$. Among s observations $\{\omega_1, \ldots, \omega_s\}$ there are i_A observations which belong to set A, $0 \le i_A \le s$. The random variable i_A is distributed binominally and its values can be used to estimate $H(A)$ (Mainland [19]). To do this let us consider the following functions

$$\begin{aligned} \phi(s, k, z) &= \sum_{i=k}^{s} \frac{s!}{i!(s-i)!} z^i (1-z)^{s-i} \\ \Psi(s, k, z) &= \sum_{i=0}^{k} \frac{s!}{i!(s-i)!} z^i (1-z)^{s-i} \end{aligned} \tag{3}$$

observe that

$$\phi(s, k, z) = \Psi(s, s - k, 1 - z)$$

$$P^s(i_A \geq k) = \phi(s, k, H(A)) \tag{4}$$

$$P^s(i_A \leq k) = \Psi(s, k, H(A))$$

The function $\phi(s, k, z)$ is a monotonically increasing function of z on the interval $[0, 1]$, $\phi(s, k, 0) = 0$, $\phi(s, k, 1) = 1$, $k \neq 0$. Therefore the solution of equation $\phi(s, k, z) = c$ exist for any $0 \leq c \leq 1$. Let us take

$$a(s, k): \phi(s, k, a(s, k)) = \alpha,\ k \neq 0 \tag{5}$$

$$b(s, k): \Psi(s, k, b(s, k)) = \alpha,\ k \neq s$$

$$a(s, 0) = 0,\ b(s, s) = 1$$

The values $a(s, k)$ and $b(s, k)$ ar the lower and upper bounds for the probability $H(A)$ in the following sense.

LEMMA 1. *For any fixed set $A \subset \mathbf{B}$ the bound $a(s, k)$ defined in* (5) *possess the following properties*

1. $P^s\{a(s, i_A) > H(A)\} \leq \alpha$ *for any measure H.*
2. *If for some function $c(i)$, $i = 0:s$, $c(i+1) > c(i)$ we have $P^s\{c(i_A) > H(A)\} \leq \alpha$ for any H then $c(i) \leq a(s, i)$*

This lemma shows that $a(s, i_A)$ is in a certain sense the best lower bound for the probability $H(A)$. The similar result holds for the upper bound $b(s, i_A)$:

LEMMA 1′. *For any fixed set $A \subset B$ $b(s, k)$ defined in (5) possess the following properties:*

1. $P^s\{b(s, i_A) < H(A)\} \leq \alpha$
2. *If for some function $c(i)$, $i = 0:s$, $c(i+1) > c(i)$ we have $P^s\{c(i_A) < H(A)\} \leq \alpha$ for any H then $c(i) \geq b(s, i)$.*

PROOF of the Lemma 1

1. Statements (3)–(5) imply

$$P^s\{a(s, i_A) > H(A)\} = P^s\{\phi(s, i_A, H(A)) < \phi(s, i_A, a(s, i_A))\}$$

$$= P^s\{\phi(s, i_A, H(A)) < \alpha\}$$

$$= \sum_{j=0}^{s} \mathbf{P}^s \{i_A = j\} \mathbf{P}^s \{\phi(s, j, H(A)) < \alpha \mid i_A = j\}$$

$$= \phi(s, j(H, \alpha), H(A))$$

where $j(H,A) = \min_j \{j : \phi(s, j, H(A)) < \alpha\}$. Therefore $\mathbf{P}^s \{a(s, i_A) > H(A)\} \le \alpha$ for any H.

2. Consider now arbitrary function $c(i)$, $i = 0, \ldots, s$. We obtain:

$$\mathbf{P}^s(c(i_A) > H(A)) = \sum_{i=0}^{s} \mathbf{P}^s(i_A = i) \mathbf{P}^s \left[c(i) > H(A) |_{i_A = i} \right]$$

$$= \sum_{i=j(H,A)}^{s} \frac{s!}{i!(s-i)!} H(A)^i (1 - H(A))^{s-i}$$

where $j(H, A) = \min_j \{j : c(j) > H(A)\}$.

Assumption 2 of the lemma now implies

$$\sup_H \phi(s, j(H, A), H(A)) \le \alpha \tag{6}$$

Suppose now that $c(0) > 0$. Taking $H(A)$ such that $0 \le H(A) < c(0)$ we obtain $j(H, A) = 0$ and $\phi(s, j(H, A), H(A)) = 1$ which contradicts (6). Therefore $c(0) = 0$. In case if $c(i) > a(s, i)$ we can take $a(s, i) < H(A) < c(i)$. Then $j(H, A) = i$ and $\phi(s, j(H, A), H(A)) > \alpha$ which again contradicts (6). Thus, $c(i) \le a(s, i)$ for any i and the proof is completed.

The values $a(s, k)$ and $b(s, k)$ defined by (5) have the following useful property.

LEMMA 2. *There exists $\bar{\alpha}$ such that for all $\alpha < \bar{\alpha}$ the bounds $a(s, i)$ and $b(s, i)$ satisfy the following property:*

$$a(s, i+1) - a(s, i) > a(s, i) - a(s, i-1)$$

$$b(s, i+1) - b(s, i) < b(s, i) - b(s, i-1) \tag{7}$$

$$i = 1, \ldots, s-1$$

The proof of this lemma is very technical and is therefore omitted. The value of $\bar{\alpha}$ computed with four digits accuracy is $\bar{\alpha} = 0.4681$. This property is very important for further considerations and it will be assumed that $\alpha < \bar{\alpha}$.

Now it is possible to specify precisely the process of obtaining bounds for the solution of the problem (1):

Precise statement. The solution process evolves in discrete time $s = 0, 1, \ldots$. Before time interval s the set $\{\omega_1, \ldots, \omega_{s-1}\}$ of observations is available which define the set G_{s-1} of admissible distributions. At the time interval s the solution process consists of the following steps:

1. New observation ω_s arrives. The whole available set of observations $\{\omega_1, \ldots, \omega_s\}$ defines the set of admissible distributions G_s in the following way:

$$G_s = \{H : a(s, i_A) \le H(A) \le b(s, i_A)\}$$

for any measurable A, where $a(s, i_A)$ and $b(s, i_A)$ are defined in (5). Additional information about the actual distribution H of random parameters ω can be included in the definition of the set G_s. Some ways of doing this will be discussed later.

2. The solution of problem (1) at step s is defined as the pair (f_s^l, f_s^u) of lower and upper bounds

$$f_s^l = \min_{x \in X} \min_{H \in G_s} \int f(x, \omega)\, dH(\omega) \tag{8}$$

$$f_s^u = \min_{x \in X} \max_{H \in G_s} \int f(x, \omega)\, dH(\omega) \tag{9}$$

and the optimal point x_s^* is defined as follows:

$$\max_{H \in G_s} \int f(x_s^*, \omega)\, dH(\omega) = f_s^u, \; x_s^* \in X$$

3. The process is repeated next time interval $s + 1$ with arrival of new observation.

The bounds on solution obtained in this fashion are constructed involving the "best" and the "worst" admissible in the sense of lemma 1 distribution H and the point x_s^* yields the value of the objective within these bounds. In what follows we shall concentrate on the numerical aspects of the problems (8)–(10).

3. THE SOLUTION OF THE INNER PROBLEM

The minimax problems (8)–(9) look very difficult because the inner problems involve optimization over the set of probability measures defined by quite complicated constraints. It appears, however, that inner problems have explicit solution. Let us consider this problem in more detail and denote for simplicity $f(x, \omega) = g(\omega)$. We are interested in solving the following problem:

minimize (or maximize) with respect to H

$$\int g(\omega)\, dH(\omega) \tag{10}$$

subject to constraints

$$a(s, i_A) \le H(A) \le b(s, i_A),\ A \in \mathbf{B} \tag{11}$$

$$H(\Omega) = 1$$

Let us assume that $g(\omega^0) = \min_{\omega \in \Omega} g(\omega)$ and $g(\omega^{s+1}) = \max_{\omega \in \Omega} g(\omega)$ exist and arrange the set of observations $\{\omega_1, \ldots, \omega_s\}$ in order of increasing values of the function $g(\omega)$:

$$\omega^0, \omega^1, \ldots, \omega^s, \omega^{s+1}$$

Here and elsewhere the original order of observations is indicated by subscript and arrangement in increasing order of the values of g is indicated by superscript. The first element of new arrangement will always be the point with the minimal value of the objective function on the set Ω and the last element (with number $s + 1$) will be the point with maximal value. This arrangement depends on the number s of the time interval, but this dependence will not be explicitly indicated for the simplicity of notations.

The solution of the problem (10)–(11) is given by the following theorem:

THEOREM 1 *Suppose that exist points* ω^0 *and* ω^{s+1} *such that* $g(\omega^0) = \min_{\omega \in \Omega} g(\omega)$, $g(\omega^{s+1}) = \max_{\omega \in \Omega} g(\omega)$. *Then*

1. *The solution of the problem (10)–(11) exist and among extremal measures always exist discrete one which is concentrated in* $s + 1$ *points:*

$$\bar{g}_s = \max_{H \in G_s} \int g(\omega)\, dH(\omega) = \int g(\omega)\, d\bar{H}_s(\omega) = \sum_{i=1}^{s+1} p_s^i g(\omega^i) \tag{12}$$

$$\underline{g}_s = \min_{H \in G_s} \int g(\omega)\,dH(\omega) = \int g(\omega)\,d\underline{H}_s(\omega) = \sum_{i=0}^{s} q_s^i g(\omega^i) \tag{13}$$

$$\bar{H}^s = \{(\omega^0, p_s^0), \ldots, (\omega^{s+1}, p_s^{s+1})\}$$

$$\underline{H}_s = \{(\omega^0, q_s^0), \ldots, (\omega^{s+1}, q^{s+1})\}$$

$$p_s^i = a(s, i) - a(s, i-1),\ i = 1:s \tag{14}$$

$$q_s^i = a(s, s-i+1) - a(s, s-i),\ i = 1:s$$

$$p_s^0 = q_s^{s+1} = 0,\ p_s^{s+1} = q_s^0 = b(s, 0)$$

2.

$$\bar{g}_s - \underline{g}_s < \Delta_g \sqrt{\frac{2|\ln \alpha|}{s}}$$

where $\Delta_g = \max_{\omega \in \Omega} g(\omega) - \min_{\omega \in \Omega} g(\omega)$

3.

$$\bar{g}_s \longrightarrow \int g(\omega)\,dH(\omega)$$

$$\underline{g}_s \longrightarrow \int g(\omega)\,dH(\omega)$$

with probability 1 as $s \longrightarrow \infty$.

PROOF

1. Let us consider the upper bound $\bar{g}_s$ and define the sets

$$\Omega_i = \{\omega : \omega \in \Omega,\ g(\omega) > g(\omega^i)\},\ i = 0:s$$

and functions

$$g_i(\omega) = \begin{cases} g(\omega^{i+1}) - g(\omega^i) & \text{if } \omega \in \Omega_{i+1} \\ g(\omega) - g(\omega^i) & \text{if } \omega \in \Omega_i \setminus \Omega_{i+1} \\ 0 & \text{if } \omega \in \Omega \setminus \Omega_i \end{cases}$$

$i = 1:s-1$,

$$g_s = \begin{cases} g(\omega) - g(\omega^s) & \text{if } \omega \in \Omega_s \\ 0 & \text{if } \omega \in \Omega \setminus \Omega_s \end{cases}$$

$$g_0 = \begin{cases} g(\omega^1) - g(\omega^0) & \text{if } \omega \in \Omega_1 \\ g(\omega) - g(\omega^0) & \text{if } \omega \in \Omega \setminus \Omega_1 \end{cases}$$

Each set Ω_i contains $s - i$ points ω_i. Let us consider the problem

$$\max_{H \in \bar{G}_s} \int g(\omega)\, dH(\omega) \tag{15}$$

where

$$\bar{G}_s = \left\{ H : H(\Omega) = 1,\ a(s, s-i) \le H(\Omega_i) \le b(s, s-i),\ i = 0 : s \right\}$$

We have

$$\int g(\omega)\, dH(\omega) = g(\omega^0) + \sum_{i=0}^{s} \int_{\Omega_i} g_i(\omega)\, dH(\omega)$$

therefore

$$\max_{H \in G_s} \int g(\omega)\, dH(\omega) \le g(\omega^0) + \sum_{i=0}^{s} \max_{H \in G_s} \int_{\Omega_i} g_i(\omega)\, dH(\omega)$$

$$= g(\omega^0) + \sum_{i=0}^{s} (g(\omega^{i+1}) - g(\omega^i))\, b(s, s-i)$$

$$= \sum_{i=1}^{s+1} g(\omega^i)(b(s, s-i+1) - b(s, s-i)) = \bar{g}_s$$

where we took $b(N, -1) = 0$. Hence $\bar{H}_s$ is the solution of problem (15) with $\bar{g}_s$ being it's optimal value. We have

$$\{H : H(\Omega) = 1,\ a(s, i_A) \le H(A) \le b(s, i_A),\ A \subset \mathbf{B}\} = G_s \subset \bar{G}_s$$

and therefore

$$\max_{H \in G_s} \int g(\omega)\, dH(\omega) \le \bar{g}_s$$

It is left to prove that $\bar{H}_s \in G_s$. To do this let us consider arbitrary measurable subset A of Ω which contains k from $s+1$ points $\{\omega^1, \ldots, \omega^{s+1}\}$, say $\{\omega^{i_1}, \ldots, \omega^{i_k}\}$.

$$\bar{H}_s(A) = \sum_{j=1}^{k} (b(s, s-i_j+1) - b(s, s-i_j))$$

According to the lemma 2

$$b(s, s - i_j + 1) - b(s, s - i_j) \le b(s, k - j) - b(s, k - j - 1)$$

Therefore

$$\bar{H}_s(A) \le \sum_{j=1}^{k} (b(s, k - j) - b(s, k - j - 1)) \tag{16}$$

$$= b(s, k - 1)$$

On the other hand

$$b(s, s - i_j + 1) - b(s, s - i_j)$$

$$\ge b(s, s - j + 1) - b(s, s - j)$$

$$= a(s, j) - a(s, j - 1)$$

Therefore

$$\bar{H}_s(A) \ge \sum_{j=1}^{k} (a(s, j) - a(s, j - 1)) = a(s, k) \tag{17}$$

Inequalities (16) and (17) prove that $\bar{H}_s(A) \in G_s$. The proof of the first statement of the theorem is completed.

2. We have

$$\bar{g}_s - \underline{g}_s = \sum_{i=0}^{s+1} (p_s^i - q_s^i) g(\omega^i)$$

$$= \sum_{i=1}^{s} (p_s^i - q_s^i) g(\omega^i) + b(s, 0)(g(\omega^{s+1}) - g(\omega^0)) \tag{18}$$

Consider now the difference $p_s^i - q_s^i$

$$p_s^i - q_s^i = (a(s, i) - a(s, i - 1)) - (a(s, s - i + 1) - a(s, s - i))$$

The first term is monotonically increasing, the second is monotonically decreasing and for odd s we have

$$p_s^i - q_s^i < 0 \quad \text{for } 1 \le i < \frac{s + 1}{2}$$

$$p_s^i - q_s^i = 0 \quad \text{for } i = \frac{s+1}{2}$$

$$p_s^i - q_s^i > 0 \quad \text{for } \frac{s+1}{2} < i \le s$$

Therefore for odd s we can continue (18) as follows:

$$\sum_{i=1}^{s} (p_s^i - q_s^i) g(\omega^i) = \sum_{i=1}^{\frac{s+1}{2}} (p_s^i - q_s^i) q(\omega^i)$$

$$+ \sum_{i=\frac{s+1}{2}}^{s} (p_s^i - q_s^i) g(\omega^i)$$

$$\le g(\omega^0) \sum_{i=1}^{\frac{s+1}{2}} (p_s^i - q_s^i) + g(\omega^{s+1}) \sum_{i=\frac{s+1}{2}}^{s} (p_s^i - q_s^i)$$

According to the definition of p_s^i and q_s^i we have

$$\sum_{i=1}^{\frac{s+1}{2}} (p_s^i - q_s^i) = a\left[s, \frac{s+1}{2}\right] + a\left[s, \frac{s-1}{2}\right] - a(s, s)$$

$$\sum_{i=\frac{s+1}{2}} (p_s^i - q_s^i) = a(s, s) - a\left[s, \frac{s-1}{2}\right] - a\left[s, \frac{s+1}{2}\right]$$

Taking into account that $b(s, 0) = 1 - a(s, s)$ last equalities together with (18) will give

$$\bar{g}_s - \underline{g}_s \le \left[g(\omega^{s+1}) - g(\omega^0))(1 - a\left[s, \frac{s+1}{2}\right] - a\left[s, \frac{s-1}{2}\right]\right] \tag{20}$$

To proceed further we have to obtain the lower bound for $a(s, (s+1)/2)$ and $a(s, (s-1)/2)$. To do this remember that $a(s, (s+1)/2)$ is the solution of equation

$$\Pr\left[i_A \ge \frac{s+1}{2}\right] = \alpha$$

with respect to probability z in the binomial distribution, by Pr we denoted the binomial distribution with s trials. This can be rewritten as follows:

$$\Pr\left[\frac{i_A}{s} - z \ge \frac{1}{2} + \frac{1}{2s} - z\right] = \alpha$$

On the other hand we have the following Okamoto inequality [20]

$$\Pr\left[\frac{i_A}{s} - z \geq c\right] < e^{-2sc^2}$$

for any $c \geq 0$. Due to $1/2 + 1/2s - \alpha(s, (s+1)/2) > 0$ for all s we obtain from previous inequality

$$\Pr\left[\frac{i_A}{s} - z \geq \frac{1}{2} + \frac{1}{2s} - z\right] < e^{-2s(\frac{1}{2}+\frac{1}{2s}-z)^2}$$

for $z = \alpha(s, (s+1)/2)$.

Therefore the solution of the equation

$$e^{-2s(\frac{1}{2}+\frac{1}{2s}-z)^2} = \alpha$$

will give the lower bound for $\alpha(s, (s+1)/2)$ This gives $\alpha(s, (s+1)/2) \geq 1/2 + 1/2s - \sqrt{\frac{|\ln\alpha|}{2s}}$. In the same way we obtain

$$\alpha\left[s, \frac{s-1}{2}\right] \geq \frac{1}{2} - \frac{1}{2s} - \sqrt{\frac{|\ln\alpha|}{2s}}$$

substituting this in (20) we obtain

$$\bar{g}_s - \underline{g}_s \leq (g(\omega^{s+1}) - g(\omega^0))\sqrt{\frac{2|\ln\alpha|}{s}}$$

In case of the even s we obtain

$$p_s^i - q_s^i < 0 \quad \text{for } 1 \leq i \leq \frac{s}{2}$$

$$p_s^i - q_s^i > 0 \quad \text{for } \frac{s}{2} + 1 \leq i \leq s$$

This gives

$$\sum_{i=1}^{s/2} (p_s^i - q_s^i) = 2\alpha\left[s, \frac{s}{2}\right] - \alpha(s, s)$$

$$\sum_{i=s/2+1}^{s} (p_s^i - q_s^i) = \alpha(s, s) - 2\alpha\left[s, \frac{s}{2}\right]$$

and finally

$$\bar{g}_s - \underline{g}_s \le \Delta_g \left[1 - 2\alpha \left[s, \frac{s}{2} \right] \right] \tag{21}$$

$$= 2\,\Delta_g \left[\frac{1}{2} - \alpha \left[s, \frac{s}{2} \right] \right]$$

After applying Okamoto unequality we again obtain from (21) the desired unequality.

3. Observe that the empirical distribution $H_s^e = \{(\omega_1, 1/s), \ldots, (\omega_s, 1/s)\}$ always belongs to G_s. Therefore the last statement of the theorem follows from statement 2, boundedness of $g(\omega)$ on Ω and the law of large numbers.

The proof is completed.

Results of numerical experiments for computing bounds $\bar{g}_s$ and $\underline{g}_s$ are shown in Figures 1–6. These bounds were computed for confidence level $\alpha = 0.1$. They all exhibit similar behavior: rapid convergence for a small number of observations which slowed down as the number of observation grows in accordance with result of the theorem. Almost in all examples the actual value of $\int g(\omega) dH(\omega)$ always stayed within bounds, although the value of α was chosen 0.1. This happened because the bounds were computed for the worst distributions which are those concentrated in a finite number of points. When such distributions were taken the behavior of bounds worsened (Figures 2–4) and in some cases the bounds did not contain actual value (Figure 4). At the same time for smoother distributions convergence is faster (Figure 5, where distribution is close to normal).

Within the framework of nonparametric statistics [16] the bounds proposed in this section can be considered as a special type of L-estimates.

4. INCORPORATING ADDITIONAL INFORMATION

The method developed in the last section deals with the case when the only available information on the distribution H of random parameters ω comes from observations ω_i. In many cases, however, additional information is available which is drawn from observations on similar systems. One of the ways of using this information is considered in this section.

Suppose that additional information comes in the form of constraints on the values of moments like expectation, variance etc. This can be expressed in the following way:

$$\int v_i(\omega)\,dH(\omega) \leq 0 \quad i = 1:m_1$$

The problem of getting upper and lower bounds in this case can be expressed similarly to (10)–(11). Let us take the problem of getting the upper bound:

maximize with respct to H

$$\int g(\omega)\,dH(\omega) \tag{22}$$

subject to constraints

$$a(s, i_A) \leq H(A) \leq b(s, i_A) \quad \forall A \in \mathbf{B} \tag{23}$$

$$H(\Omega) = 1$$

$$\int v_i(\omega)\,dH(\omega) \leq 0 \quad i = 1:m_1 \tag{24}$$

The Lagrange multipliers are used to take account of constraint (24) and reduce the problem (22)–(24) to the problem (10)–(11).

Let us consider the function

$$L(\omega, u) = g(\omega) - \sum_{i=1}^{m_1} u_i\, g_i(\omega)$$

and assume that for each $u \geq 0$ exist $\omega^0(u)$ and $\omega^{s+1}(u)$ such that

$$L(\omega^0, u) = \min_{\omega \in \Omega} L(\omega, u),\ L(\omega^{s+1}, u) = \max_{\omega \in \Omega} L(\omega, u) \ . \tag{25}$$

For each fixed u arrange ω_i in increasing order:

$$\omega^0(u),\ \omega^1(u),\ \ldots,\ \omega^s(u),\ \omega^{s+1}(u)$$

where $L(\omega^i(u), u) \leq L(\omega^{i+1}(u), u)$. Here again we use superscript to indicate ordered observations, this time, however, ordering will depend on u.

$$\text{Construct } \bar{H}^s(u) = \{(\omega^0(u), p_s^0), \ldots, (\omega^{s+1}(u), p_s^{s+1})\} \tag{26}$$

and take

$$\Psi^s(u) = \sum_{i=0}^{s+1} p_s^i\, L(\omega^i(u), u) \tag{27}$$

where p_s^i are defined in (14).

THEOREM 2

Suppose that

1. *For* $u \geq 0$ *exist* $\omega^0(u)$, $\omega^{s+1}(u)$ *such that (25) is satisfied.*
2. *Exists* $\delta > 0$ *such that*

$$\int v_i(\omega)\, dH(\omega) < -\delta \quad i = 1 : m_1$$

Let us take $u^* : u^* \geq 0$, $\Psi^s(u^*) = \min_{u \geq 0} \Psi^s(u)$ *(if this* u^* *exists). Then the measure* $\bar{H}^s(u^*)$ *defined in (26) is the solution of the problem (22)–(24) with probability* $\mathbf{P}^s$ *in the space* $(\Omega^s, \mathbf{B}^s, H^s)$ *at least equal to* γ *where*

$$\gamma = 1 - \frac{1}{s\delta^2} \sum_{i=1}^{m_1} \left[\int v_i^2(\omega)\, dH(\omega) - \left(\int v_i(\omega)\, dH(\omega)\right)^2 \right]$$

$\Psi^s(u)$ *is defined in (27) and the corresponding upper bound is* $\Psi^s(u^*)$.

PROOF Observe that for empirical distribution H_s^e consisting of s points constraints (23) are satisfied. Let us estimate the probability with which $\int v_i(\omega)\, dH_s^e(\omega) = 1/s \sum_{i=1}^{s} v_i(\omega_s) < -\varepsilon$ with $\varepsilon > 0$.

According to generalization of Tchebyshev inequality

$$\mathbf{P}^s \left\{ \max_{i=1:m_1} \left| \frac{1}{s} \sum_{j=1}^{s} v_i(\omega_j) - \int v_i(\omega)\, dH(\omega) \right| > \delta \right\}$$

$$< \frac{1}{\delta^2} \sum_{i=1}^{m_1} E_{\mathbf{P}^s} \left[\frac{1}{s} \sum_{j=1}^{s} v_i(\omega_j) - E_{\mathbf{P}^s} \left[\frac{1}{s} \sum_{j=1}^{s} v_i(\omega_j) \right] \right]^2$$

$$= \frac{1}{s^2\delta^2} E_{\mathbf{P}^s} \left[\sum_{j=1}^{s} \left(v_i(\omega_j) - \int v_i(\omega)\, dH(\omega)\right) \right]^2$$

$$= \frac{1}{s\delta^2} \sum_{i=1}^{m_1} \left(\int v_i^2(\omega)\, dH(\omega) - \left(\int v_i(\omega)\, dH(\omega)\right)^2 \right)$$

Therefore

$$\int v_i(\omega)\, dH_s^e(\omega) < -\varepsilon \tag{28}$$

for some $\varepsilon > 0$ with probability at least γ. Convexity of the set, defined by (23) similarly to [9] implies now equivalence of the problem (22)–(24) and the problem

$$\min_{u \geq 0} \max_{H} \int L(\omega, u)\, dH(\omega) \tag{29}$$

subject to (23), which holds for all $\omega \in \Omega^s$ such that (28) is satisfied. The inner problem in (29) has explicit solution defined by theorem 1. This solution is described by relations (26)–(27).

The proof is completed.

A similar result holds for lower bound if we take $L(\omega, u) = g(\omega) + \sum_{i=1}^{m_1} u_i g(\omega)$ and $\Psi^s(\omega) = \sum_{i=0}^{s+1} q_s^i L(\omega^i(u), u)$.

The theorem 2 reduces the problem of getting upper bound to minimization of function $\Psi^s(u)$, defined in (27). This is a convex function with readily available values of subgradient. Therefore suitable nondifferentiable optimization techniques can be applied to get its minima [18, 22]. Different situations can occur during such computations.

1. It was found that $\Psi^s(u)$ is not bounded from below. This means that problem (22)–(24) is infeasible. To get an upper bound in this case it is necessary to drop constraint (24) and solve the problems (10)–(11) instead. According to theorem 2 the probability of this case tends to 0 as $s \longrightarrow \infty$.

2. The point u^* was found such that within prescribed accuracy $\Psi^s(u^*)$ will be the optimal value of the problem (22)–(24), i.e. the desired upper bound with probability that tends to 1 when $s \longrightarrow \infty$. For finite s, however, it is possible that $\Psi^s(u^*)$ exceeds the optimal value of (22)–(24). In both cases it will be the upper bound, but in the second case not the best one.

5. NUMERICAL TECHNIQUES FOR FINDING THE UPPER BOUNDS FOR THE SOLUTION OF OPTIMIZATION PROBLEMS

Let us return to the problems (8)–(9) now. Define

$$\bar{F}(s, x) = \max_{H \in G_s} \int f(x, \omega) \mathrm{d}H(\omega) \tag{30}$$

$$\underline{F}(s, x) = \min_{H \in G_s} \int f(x, \omega) \mathrm{d}H(\omega) \tag{31}$$

It is now possible to compute the values of these functions using results of the section 3. The problem of finding upper and lower bounds for the solution of (1) can be formulated as follows:

$$f_u^s = \min_{x \in X} \bar{F}(s, x);\ f_l^s = \min_{x \in X} \underline{F}(s, x)$$

Successive bounds for simple problem are given in Figure 7 to give a feeling how they evolve. For more complicated problems it is necessary to develop special numerical techniques. The problem of finding lower bound f_l^s is not convex because the function $\underline{F}(s, x)$ is not convex with respect to x. Therefore it needs special treatment in each particular case and general efficient techniques are out of reach so far. In this section we shall concentrate on the problem of finding the upper bound f_x^s which is much easier because the function $\bar{F}(s, x)$ is convex. This allows application of convex programming algorithms to find f_u^s. These algorithms however, need a substantial number of iterations to get close to a solution and with arrival of each new observation the process should be started anew. The main point of this section is that techniques can be developed that perform very limited, perhaps only one, iteration of convex programming algorithm with each new observation and still get reasonable upper bound. Generally speaking the process of obtaining bounds look like this

1. Start from some fixed number of observation r and point x^r.

2. Suppose that prior to step number s the observations $\omega_1, \ldots, \omega_{s-1}$ arrived, the point x^{s-1} was obtained and the current upper bound is taken equal to $\bar{F}(s-1, x^{s-1})$. The following calculations are performed at the step number s.

   2a. The observation ω_s arrives. The function $\bar{F}(s, x)$ and its subgradient is computed at the point x^{s-1} and possibly at some additional points.

   2b. These values are used to perform one step of minimization of the function $\bar{F}(s, x)$ which gives the new point x^s. The procedure goes to the next step.

We shall consider a particular method based on generalized linear programming [3]. This technique provides natural bounds for the solution of the problem $\min_x \bar{F}(s, x)$ which enables to control accuracy. With each new observation new supporting hyperplane is introduced.

1. Take an initial collection of points

$$Y^0 = \{y_1, \ldots, y_k\},\ k \geq n,\ y_i \in X \subset R^n \ .$$

For each $y \in Y^0$ compute $\omega(y)$:

$$f(y, \omega(y)) = \max_{\omega \in \Omega} f(y, \omega)$$

Take $\bar{F}(0, y) = f(y, \omega(y))$ and compute

$$\bar{F}_x(0, y) = f_x(y, \omega(y))$$

Solve the problem

$$\min_x \max_{y \in Y^0} [\bar{F}(0, y) + \langle x - y, \bar{F}_x(0, y)]$$

This is a linear programming problem which solution is x^1. This will be the initial solution of the original problem. Take the set of observations $\Omega^0 = \phi$.

2. At the beginning of step number s we have the set of observations Ω^{s-1}, the set of approximating points Y^{s-1}, current approximation to the minimum of the upper bound x^s and for each $y \in Y^{s-1}$ we have already computed $f(y, \omega)$ $\omega \in \Omega^{s-1}$, $f(y, \omega(y))$, $f_x(y, \omega)$, $f_x(y, \omega(y))$, and estimates

$$\bar{F}(\mu_{s-1}(y), y),\ \bar{F}_x(\mu_{s-1}(y), y),\ \mu_{s-1}(y) \leq s - 1$$

then we do the following:

(a) Obtain new observation ω_s and take $\Omega^s = \Omega^{s-1} \cup \{\omega_s\}$. The set Ω^s consists now of s points $\omega_1, \ldots, \omega_s$

(b) For $y = x^s$ compute $\omega(y)$:

$$f(y, \omega(y)) = \max_{\omega \in \Omega} f(y, \omega)$$

and compute $f(y, \omega)$, $f_x(y, \omega)$ for all $\omega \in \Omega^s$.

Arrange the set $\Omega^s \cup \{\omega(y)\}$ in the order of increasing values of $f(y, \omega)$:

$$\omega^1(y), \ldots, \omega^{s+1}(y),\ f(y, \omega^{i+1}(y)) \geq f(y, \omega^i(y))$$

and $\omega^{s+1}(y) = \omega(y)$

Assign $\mu_s(y) = s$ and take

$$\bar{F}(s, y) = \sum_{i=1} p_s^i f(y, \omega^i(y))$$

$$\bar{F}_x(s, y) = \sum_{i=1}^{s+1} p_s^i f_x(y, \omega^i(y))$$

For $y \in Y^{s-1}$ take $\mu_s(y) = \mu_{s-1}(y)$. Update the set Y^{s-1}:

$$Y^s = Y^{s-1} \cup \{x^s\}$$

(c) Find

$$\min_{x \in X} \max_{y \in Y^s} [\bar{F}(\mu_s(y), y) + \langle x - y, \bar{F}_x(\mu_s(y), y)\rangle]$$

This is linear programming problem and suppose that $\tilde{x}$ is its solution. Consider the set

$$Z^s = \{y : y \in Y^s, \bar{F}(\mu_s(y), y) + \langle\tilde{x} - y, \bar{F}_x(\mu_s(y), y)\rangle$$

$$= \min_{x \in X} \max_{y \in Y^s} [\bar{F}(\mu_s(y), y) + \langle x - y, \bar{F}_x(\mu_s(y), y)\rangle]$$

and $\mu_s(y) < s\}$

If $Z^s = \phi$ then take $x^{s+1} = \tilde{x}$ and go to the step number $s + 1$, otherwise proceed to (d).

(d) For each $y \in Z^s$ compute $f(y, \omega_i)$ and $f_x(y, \omega_i)$, $\mu_s(y) < i \leq s$. Take

$$\bar{F}(s, y) = \sum_{i=1}^{s+1} p_s^i f(y, \omega^i(y))$$

$$\bar{F}_x(s, y) = \sum_{i=1}^{s+1} p_s^i f_x(y, \omega^i(y))$$

Assign $\mu_s(y) = s$ and go to (c). The following theorem deals with the convergence of this method.

THEOREM 3. *Suppose that the following conditions are satisfied*

1. *Sets $X \subset R^n$ and $\Omega \subset R^m$ are compact sets.*
2. *The function $f(x, \omega)$ is convex on x and continuous on ω, $|f(x_1, \omega) - f(x_2, \omega)| \leq L|x_1 - x_2|$ for $x_1, x_2 \in X$, $\omega \in \Omega$ Then $\bar{F}(s, x^s) - \min_{x \in X} \bar{F}(s, x) \to 0$ as $s \to \infty$ with probability 1.*

PROOF Let us denote

$$\bar{F}^s(s, x) = \max_{y \in Y^s} [\bar{F}(\mu_s(y), y) + \langle x - y, \bar{F}_x(\mu_s(y), y)\rangle]$$

Then

$$\bar{F}^s(s, x^s) = \min_{x \in X} \bar{F}^s(s, x)$$

Let us prove that $\bar{F}(s, x^s) - \bar{F}^s(s, x^s) \to 0$ with probability 1. Note that

$\bar{F}(s, x^s) \geq \bar{F}^s(s, x^s)$ is always satisfied. Suppose that for some $\Delta > 0$ exist subsequence s_k such that

$$\bar{F}(s_k, x^{s_k}) - \bar{F}^{s_k}(s_k, x^{s_k}) > \Delta$$

Due to compactness of the set X we may assume without loss of generality that $x^{s_k} \rightarrow x^*$,

$$\begin{aligned}\bar{F}(s_k, x^{s_k}) - \bar{F}^{s_k}(s_k, x^{s_k}) = \\ & \bar{F}(s_k, x^{s_k}) - \bar{F}(s_k, x^*) + \bar{F}(s_k, x^*) - F(x^* \\ & - \bar{F}^{s_k}(s_k, x^{s_k}) + \bar{F}^{s_{k+1}}(s_{k+1}, x^{s_{k+1}}) \\ & - \bar{F}^{s_{k+1}}(s_{k+1}, x^{s_{k+1}}) + \bar{F}^{s_{k+1}}(s_{k+1}, x^{s_k}) \\ & - \bar{F}^{s_{k+1}}(s_{k+1}, x^{s_k}) + \bar{F}(\mu_{s_k}(x^{s_k}), x^{s_k}) \\ & - \bar{F}(\mu_{s_k}(x^{s_k}) + \bar{F}(\mu_{s_k}(x^{s_k}), x^*) - \bar{F}(\mu_{s_k}(x^{s_k}), x^*) + F(x^*) \end{aligned} \tag{32}$$

where $s_k \leq \mu_{s_k}(x^{s_k}) \leq s_{k+1}$. According to condition 2 we have

$$|F(s, x_1) - F(s, x_2)| \leq L\,|x_1 - x_2|\ ,$$

$$|F^s(s, x_1) - F^s(s, x_2) \leq L\,|x_1 - x_2|$$

for all s and $x_1, x_2 \in X$. Therefore

$$|\bar{F}(s_k, x^{s_k}) - \bar{F}(s_k, x^*) \leq L\,|x^{s_k} - x^*|$$

$$|\bar{F}^{s_{k+1}}(s_{k+1}, x^{s_{k+1}}) - \bar{F}^{s_{k+1}}(s_{k+1}, x^{s_k})| \leq L\,|x^{s_{k+1}} - x^{s_k}|$$

$$|\bar{F}(\mu_{s_k}(x^{s_k}), x^{s_k}) - \bar{F}(\mu_{s_k}(x^{s_k}), x^*)| \leq L\,|x^{s_k} - x^*| \tag{33}$$

The theorem 1 gives

$$|\bar{F}(s_k, x^*) - F(x^*)| \rightarrow 0$$

$$|\bar{F}(\mu_{s_k}(x^{s_k}), x^*) - F(x^*)| \rightarrow 0 \tag{34}$$

with probability 1.

Definition of the algorithm implies

$$\bar{F}^{s_{k+1}}(s_{k+1}, x^{s_k}) = \bar{F}(\mu^{s_k}(x^{s_k}), x^{s_k}) \tag{35}$$

Compactness of the sets X and Ω with continuity of $f(x, \omega)$ implies boundedness of $\bar{F}(s, x)$ uniformly on s. Therefore we may assume without loss of generality that

$$|\bar{F}^{s_k}(s_k, x^{s_k}) - \bar{F}^{s_{k+1}}(s_{k+1}, x^{s_{k+1}})| \longrightarrow 0 \tag{36}$$

combining (32)–(36) we obtain

$$\bar{F}(s_k, x^{s_k}) - \bar{F}^{s_k}(s_k, x^{s_k}) \longrightarrow 0$$

with probability 1.

This contradicts initial the assumption and therefore

$$\bar{F}(s, x^s) - \bar{F}^s(s, x^s) \longrightarrow 0$$

with probabiity 1.

Thus

$$\bar{F}(s, x^s) - \min_{x \in X} \bar{F}(s, x) \longrightarrow 0$$

because

$$\bar{F}(s, x^s) \geq \min_{x \in X} \bar{F}(s, x) \geq \bar{F}^s(s, x^s)$$

The proof is completed.

The theorem suggests that proposed techniques could be viable for computing upper bounds. The important question now is whether the speed of convergence to upper bound is faster than convergence of the bounds themselves. To find conditions which guarantee this is the objective of further study.

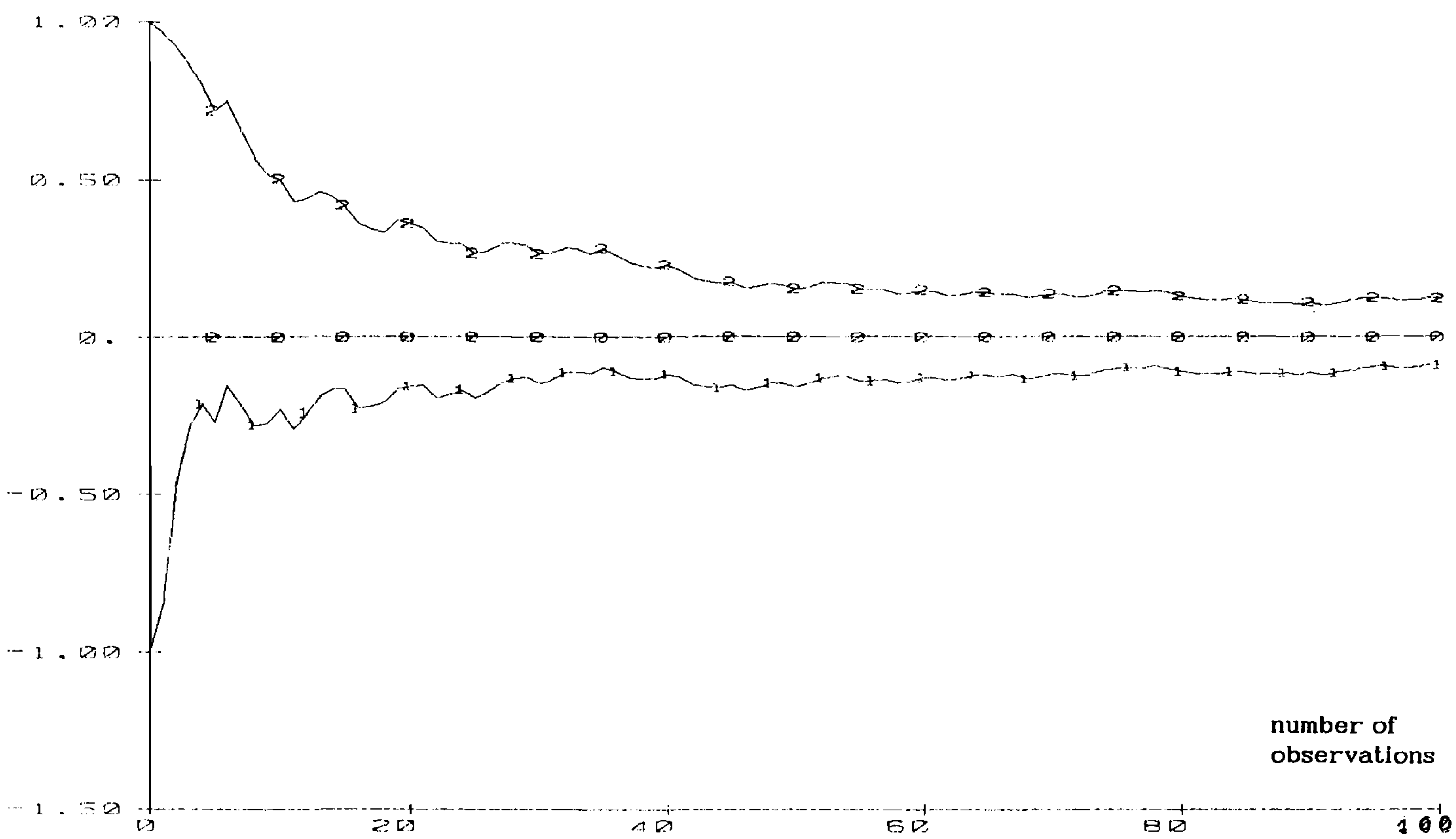

FIGURE 1. Bounds for $\int \omega \, dH(\omega)$ where H is uniform distribution on $[-1, 1]$.

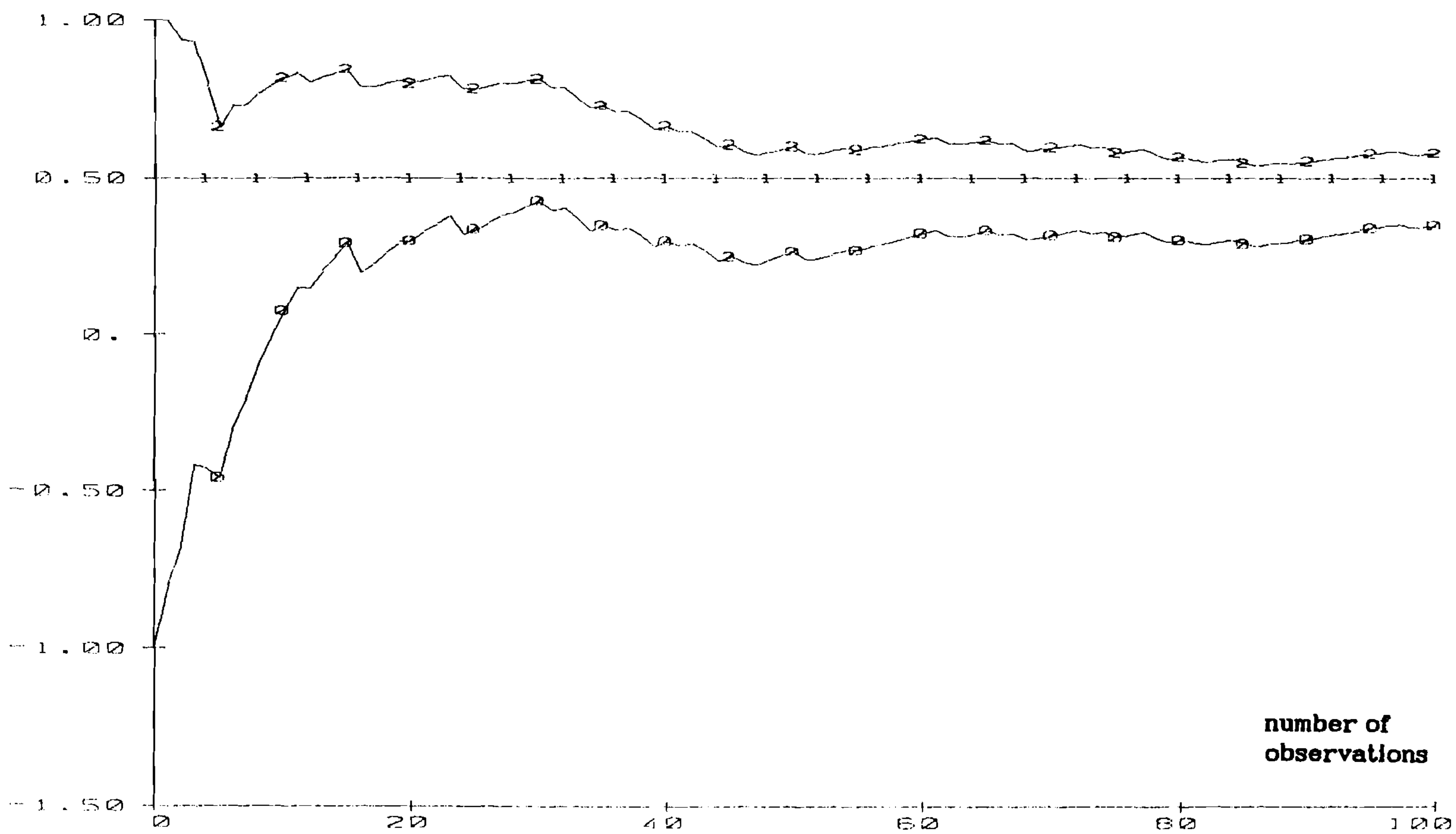

FIGURE 2. Bounds for $\int \omega \, dH(\omega)$ where $H = 0.5\,H_1 + 0.5\,H_2$, H_1 is uniform distribution on $[-1, 1]$, $H_2 = \{1\}$.

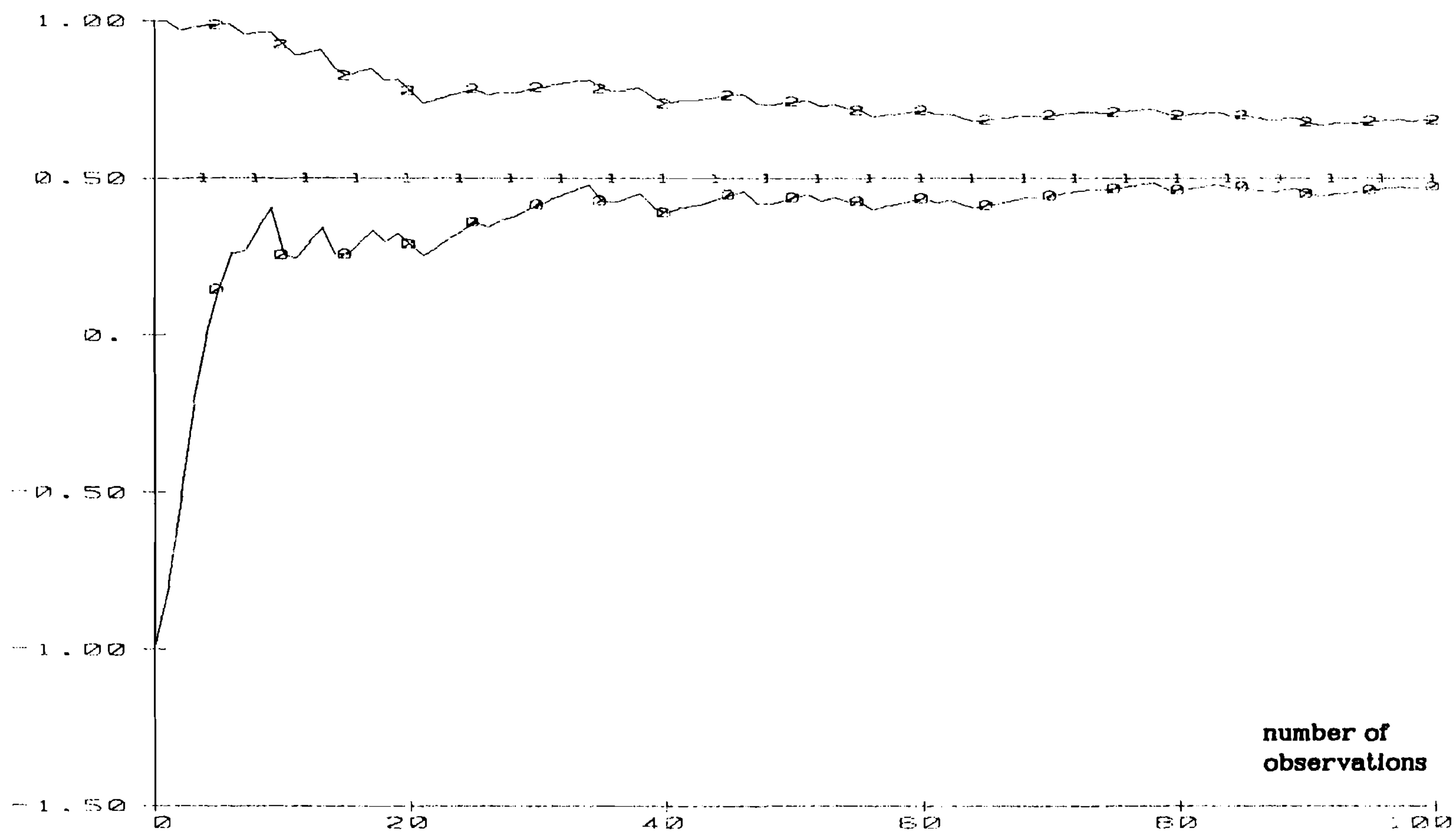

FIGURE 3. Bounds for $\int \omega \, dH(\omega)$ where $H = 0.5\, H_1 + 0.5\, H_2$, H_1 is uniform distribution on [−1, 1], $H_2 = \{1\}$ different set of observations.

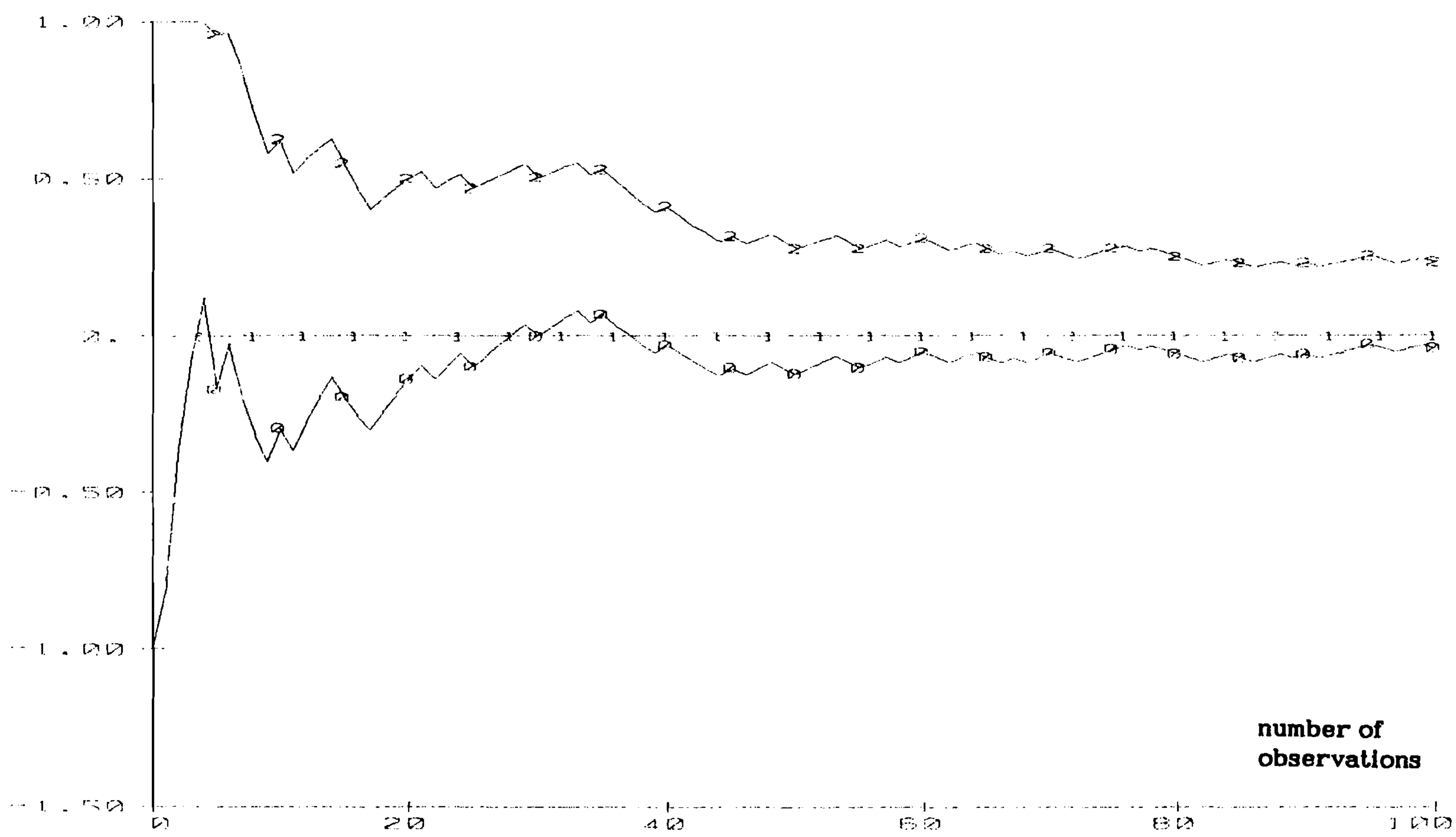

FIGURE 4. Bounds for $\int \omega \, dH(\omega)$ where $H = 0.5\, H_1 + 0.5\, H_2$, $H_1 = \{1\}$, $H_2 = \{1\}$.

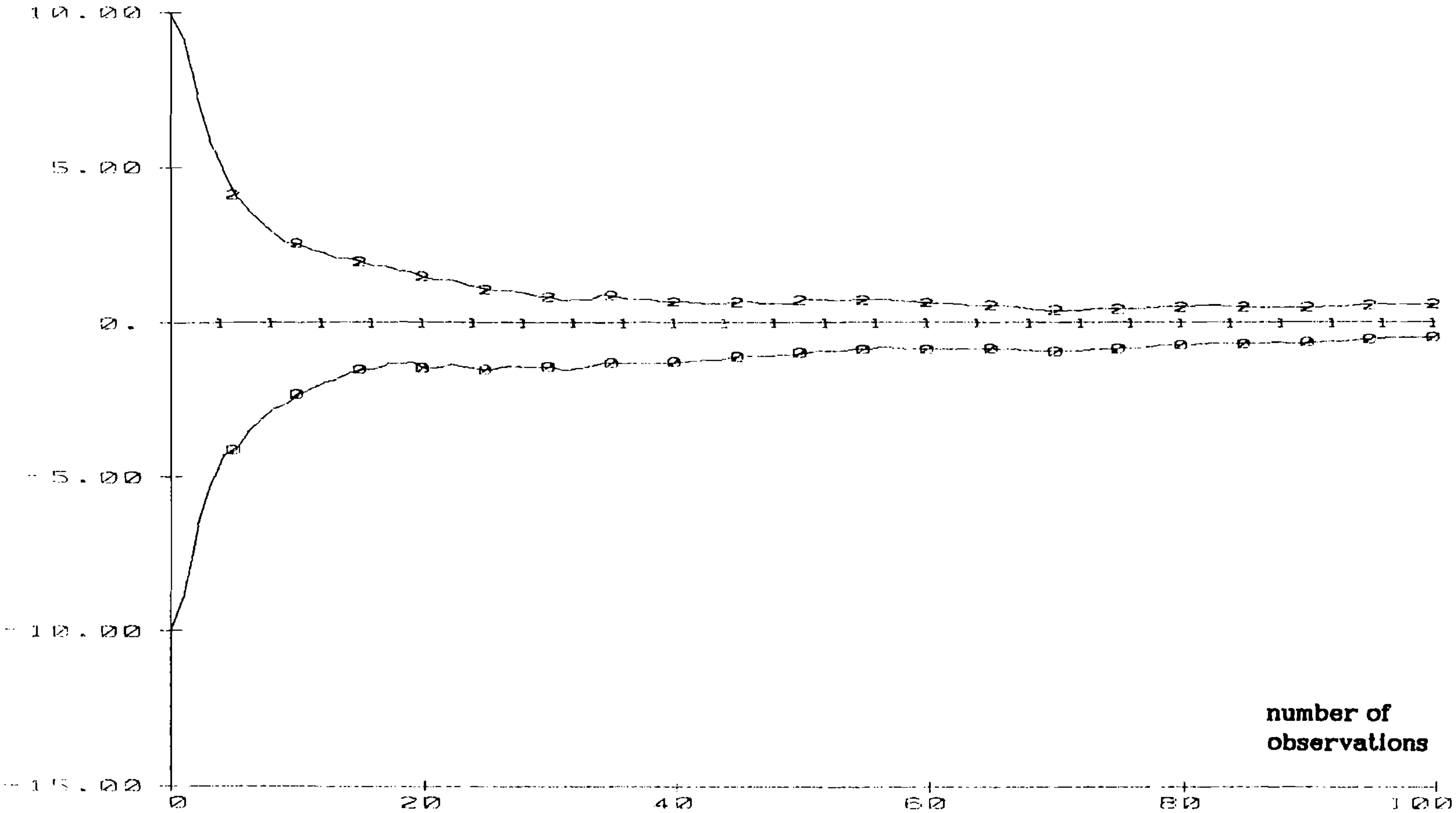

FIGURE 5. Bounds for $\int \omega \, dH(\omega)$ where H is distribution of the sum of 10 independent random variables, each distributed uniformly on [−1, 1].

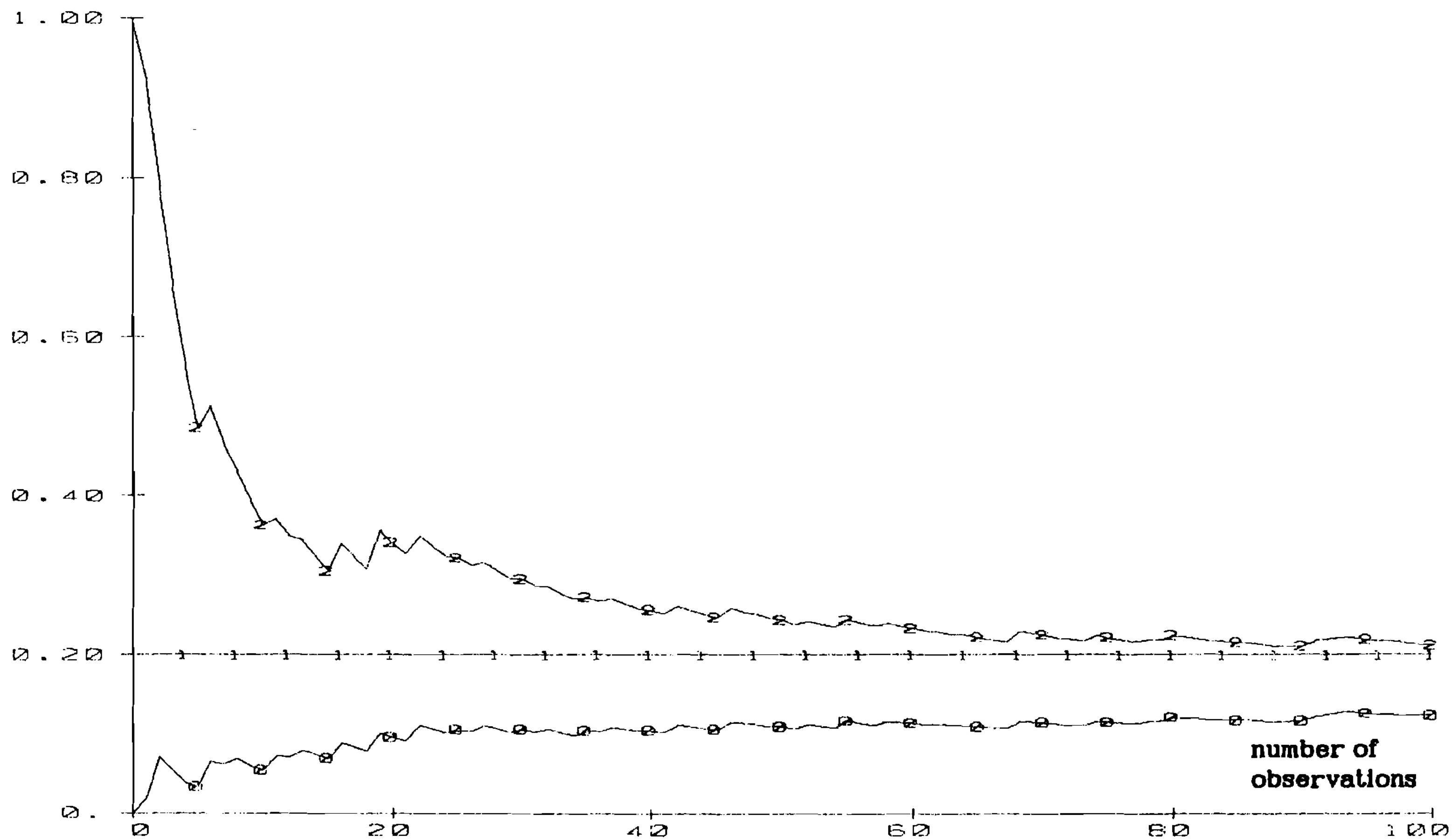

FIGURE 6. Bounds for $\int \omega^H \, dH(\omega)$ where H is uniform distribution on $[-1, 1]$.

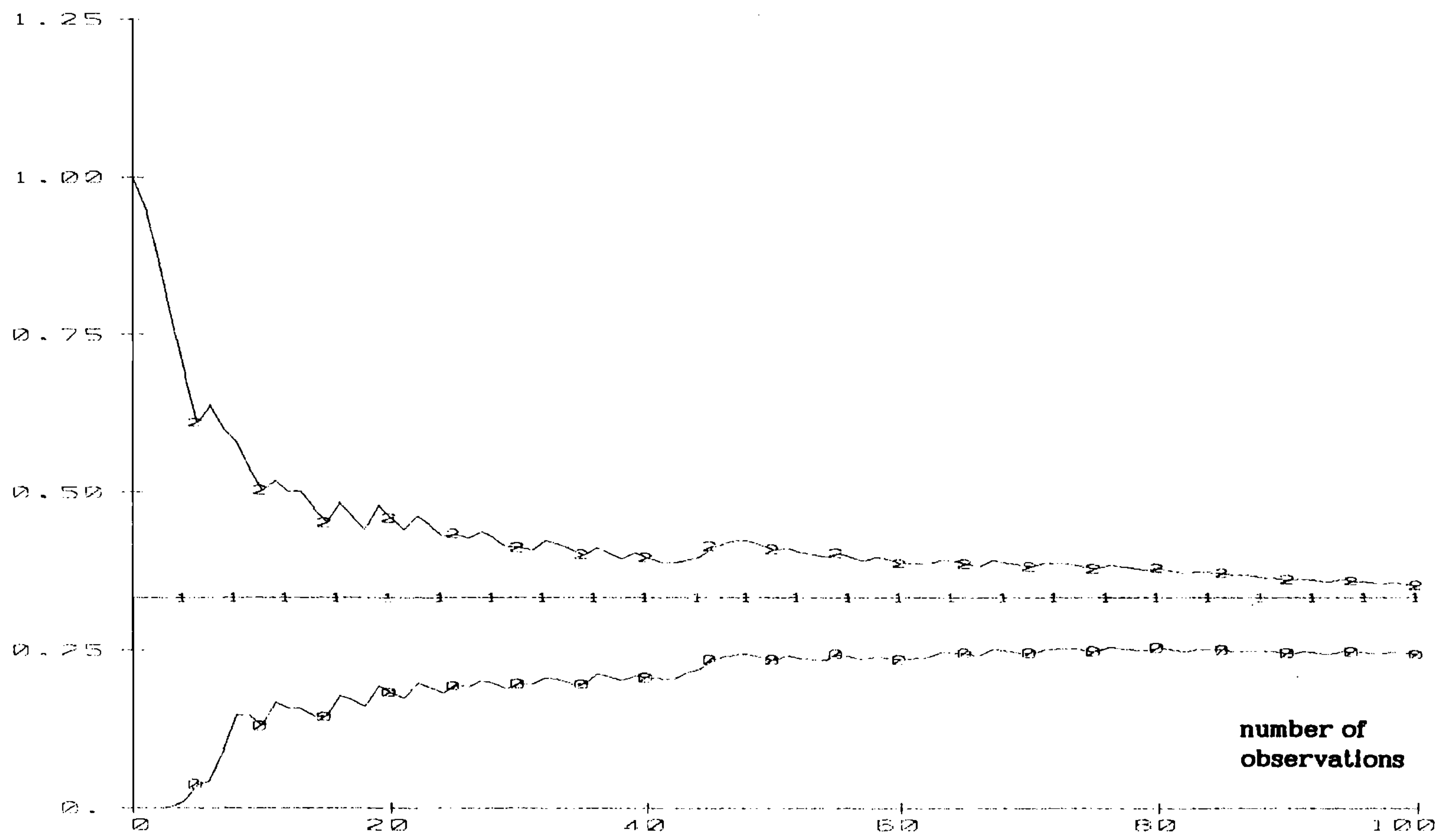

FIGURE 7. Bounds for solution of the problem $\min_x (x - \omega)^2$ where ω is uniformly distributed on $[-1, 1]$.

ACKNOWLEDGMENT

The author is grateful to Professors A. Kurzhanski and R. Wets for useful discussions.

REFERENCES

1 Birge, J. and R.J.-B. Wets: Computing bounds for stochastic programming problems by means of a generalized moment problem. Mathematics of Operations Research (forthcoming).

2 Cipra, T.:Confidence regions for linear programs with random coefficients. International Institute for Applied Systems Analysis, Working Paper WP-86-20, 1986.

3 Dantzig, G.B.: *Linear Programming and Extensions*, Princeton University Press (1980).

4 Dupačová, J.: Minimax Stochastic Programs with Nonconvex Nonseparable Penalty Functions, in: A. Prékopa (ed.): Coloquia Mathematica Societatis János Bolyai 12. Progress in Operations Research, Eger 1974 North-Holland (1976) pp. 303–316.

5 Dupačová, J.: Minimax Stochastic Programs with Nonseparable Penalties, in: K. Iracki, K. Malanowsky, S. Walukiewicz (eds.): Optimization Techniques, Part 1. LN Control and Information Sciences, Springer-Verlag, Berlin, vol. 22 (1980) pp. 157–163.

6 Ermoliev, Yu.: "Method for stochastic programming in randomized strategies," *Kibernetica* 1 (1970).

7 Ermoliev, Yu.: *Methods of Stochastic Programming* (in Russian), Nauka, Moscow (1976).

8 Ermoliev, Yu.: Stochastic quasigradient methods and their applications to system optimization, *Stochastics* **9**, 1–36, 1983.

9 Ermoliev, Yu., A Gaivoronski and C. Nedeva (1985): Stochastic optimization problems with incomplete information on distribution function. *SIAM J. Control and Optimization* **23**, 696–716.

10 Gaivoronski, A.: Linearization methods for optimization of functionals which depend on probability measures. Mathematical Programming Study 28 (1986) 157–181.

11 Gaivoronski, A.: Stochastic optimization techniques for finding optimal submeasures in: V.I. Aruin, A. Shiraev, R. Wets eds. Stochastic optimization. Proceedings of the International Conference, Kiev, 1984. Lecture Notes in Control and Information Sciences, 1986, Springer-Verlag, Berlin, Heidelberg.

12 Golodnikov, A.N.: *Finding of Optimal Distribution Function in Stochastic Programming Problems* (Dissertation abstract), Institute of Cybernetics Press, Kiev (1979).

13 Kall, P.: *Stochastic Linear Programming*, Springer Verlag, Berlin, 1976.

14 Kall, P.: Stochastic programs with recourse: an upper bound and the related moment problem. Institut für Operations Research der Universität Zürich. Working Paper, 1986.

15 Kankova, V.: An approximative solution of a stochastic optimization problem. Transaction of the eigth Prague conference on information theory, statistical decision functions, random processes v. A, P. Reidel Publishing Company, Boston, 1978.

16 Krishnaiah, P.R. and Sen, P.K. eds: Handbook of Statistics. v. 4, Nonparametric methods, North Holland, 1984.

17 Kurzhanskii, A.B.: *Control and Observation under Conditions of Uncertainty*, Nauka, Moscow, 1977.

18 Lemarechal, C.: "Bundle methods in nonsmooth optimization", in: C. Lemarechal and R. Mifflin, eds., Nonsmooth Optimization, IIASA Proceedings Series, Vol. 3, 79–102. Laxenburg, Austria (1977).

19 Mainland, D. (1948): Statistical methods in medical research, Canadian Journal of Research, 26, (section E), 1–166.

20 Okamoto, M. (1958): Some inequalities relating to the partial sum of binomial probabilities, Annals of the Institute of Statistical Mathematics, Tokyo, 10, 29–35.

21 Prékopa, A. (1973): Contributions to the theory of stochastic programming. *Mathematical Programming* **4**, 202–221.

22 Rockafellar, R.T.: "Monotone operators and the proximal point algorithm," SIAM Journal on Control and Optimization **14**(1976), 877–898.

23 Wets, R. Stochastic programming: solution techniques and approximation schemes, in *Mathematical Programming: The State of the Art*, eds., A. Bachem, M. Grötschel, and B. Korte. Springer Verlag, Berlin, 1983, pp. 566–603.

24 Wets, R. Large scale linar programming techniques in stochastic programming. IIASA Working Paper, WP-84-90, Laxenburg, Austria 1984.